Spring 2017

Surface Display for Metabolic Engineering of Industrially Important Acetic Acid Bacteria

Marshal Allen Blank

SURFACE DISPLAY FOR METABOLIC ENGINEERING OF INDUSTRIALLY IMPORTANT ACETIC ACID BACTERIA

A Masters Thesis

Presented to

The Graduate College of

Missouri State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science, Biology

By

Marshal Blank

May 2017

SURFACE DISPLAY FOR METABOLIC ENGINEERING OF INDUSTRIALLY IMPORTANT ACETIC ACID BACTERIA

Biology

Missouri State University, May 2017

Master of Science

Marshal Blank

ABSTRACT

Acetic acid bacteria (AAB) have unique metabolic characteristics that suit them for a variety of applications. They possess an arsenal of membrane-bound dehydrogenases in the periplasmic space that are capable of regiospecific and enantioselective partial oxidation of sugars, alcohols, and polyols. The resulting products are deposited directly into the medium where they are easily recovered for use as pharmaceutical precursors, industrial chemicals, food additives, and consumer products. Unfortunately, there has been little research towards improving AAB, as few molecular tools exist for metabolic engineering of these microbes. To this end, an original surface display system was developed to express recombinant enzymes at the outer membrane of the model acetic acid bacterium *Gluconobacter oxydans*. Three anchor proteins (OprF188, INPNC, and PgsA) were tested for the ability to deliver alkaline phosphatase enzyme, PhoA, to the cell surface. The OprF188 system was demonstrated for biocatalysis in whole-cell assays, and PhoA was proteolytically cleaved from the cell surface, suggesting proper delivery to the outer membrane. A linker library was also constructed to optimize surface display. The $(EAAAK)_1$ rigid linker led to the greatest improvement, increasing PhoA activity by 70%. Surface display could be used both to extend the capabilities of AAB in current biotechnological processes, and to broaden the potential of these microbes in the production of value-added products.

KEYWORDS: *Gluconobacter oxydans*, surface display, biocatalysis, fusion linkers, outer membrane proteins

This abstract is approved as to form and content

_____

Paul Schweiger, PhD
Chairperson, Advisory Committee
Missouri State University

SURFACE DISPLAY FOR METABOLIC ENGINEERING OF INDUSTRIALLY IMPORTANT ACETIC ACID BACTERIA

By

Marshal Blank

A Masters Thesis
Submitted to the Graduate College
Of Missouri State University
In Partial Fulfillment of the Requirements
For the Degree of Master of Science, Biology

May 2017

Approved:

_____
Paul Schweiger, PhD: Chairperson

_____
Kyoungtae Kim, PhD: Committee Member

_____
Laszlo Kovacs, PhD: Committee Member

_____
Julie Masterson, PhD: Dean, Graduate College

ACKNOWLEDGEMENTS

First of all, I would like to express my gratitude to Dr. Schweiger. He has done much in the past few years to help me become a better microbiologist and to further my professional development. I wish him the best at the University of Wisconsin. I would also like to acknowledge Dr. Kim and Dr. Kovacs for devoting their time as my committee members and as my professors. Dr. Kim has always taken an interest in my research and has had many helpful suggestions (and questions) along the way. Dr. Kovacs' genetics course got me interested in molecular biology to begin with, and he has aided in my graduate research as well. Of course, I would also like to recognize the Biology Department and the Graduate College for financial support. Finally, I owe a huge thanks to my family, friends, and girlfriend, Ashlynne, for their support and encouragement during this time.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The Unique Metabolism of *Gluconobacter oxydans*

Acetic acid bacteria (AAB) belong to the *Acetobacteriaceae*, a family of gram-negative, strictly-aerobic, acidophilic *Alphaproteobacteria* that produce acetic acid as their signature metabolite (De Ley et al. 1984). In nature, these microorganisms dominate carbohydrate-rich environments including fruits and flowers and serve an ecological role as secondary symbionts of insects such as fruit flies and honeybees (Crotti et al. 2010; De Ley et al. 1984; Gupta et al. 2001; Raspor and Goranovič 2008). AAB have unique metabolic characteristics when compared to other well-known prokaryotes, especially other aerobes. Most strikingly, many AAB do not fully oxidize growth substrates as part of their normal metabolism. While genera such as *Acetobacter* can oxidize acetic acid to carbon dioxide, *Gluconobacter* species cannot (Deppenmeier et al. 2002). Consequently, AAB are characterized by their ability to partially oxidize sugars, alcohols, polyols, and aldehydes to form metabolic byproducts that are both diverse and abundant. These aldehyde, ketone, and organic acid byproducts are excreted almost entirely into the growth medium (Figure 1) (Deppenmeier et al. 2002; Deppenmeier and Ehrenreich 2009; Matsushita et al. 2004).

The model acetic acid bacterium, *Gluconobacter oxydans* is particularly adept at incomplete oxidation. *G. oxydans* is both the best-characterized member of the *Acetobacteriaceae* and one of the most important microbes used in biotechnology. Its genome sequence is known and confirmed previous biochemical studies while revealing new insight into the metabolism of AAB (Prust et al. 2005). Among these insights is that

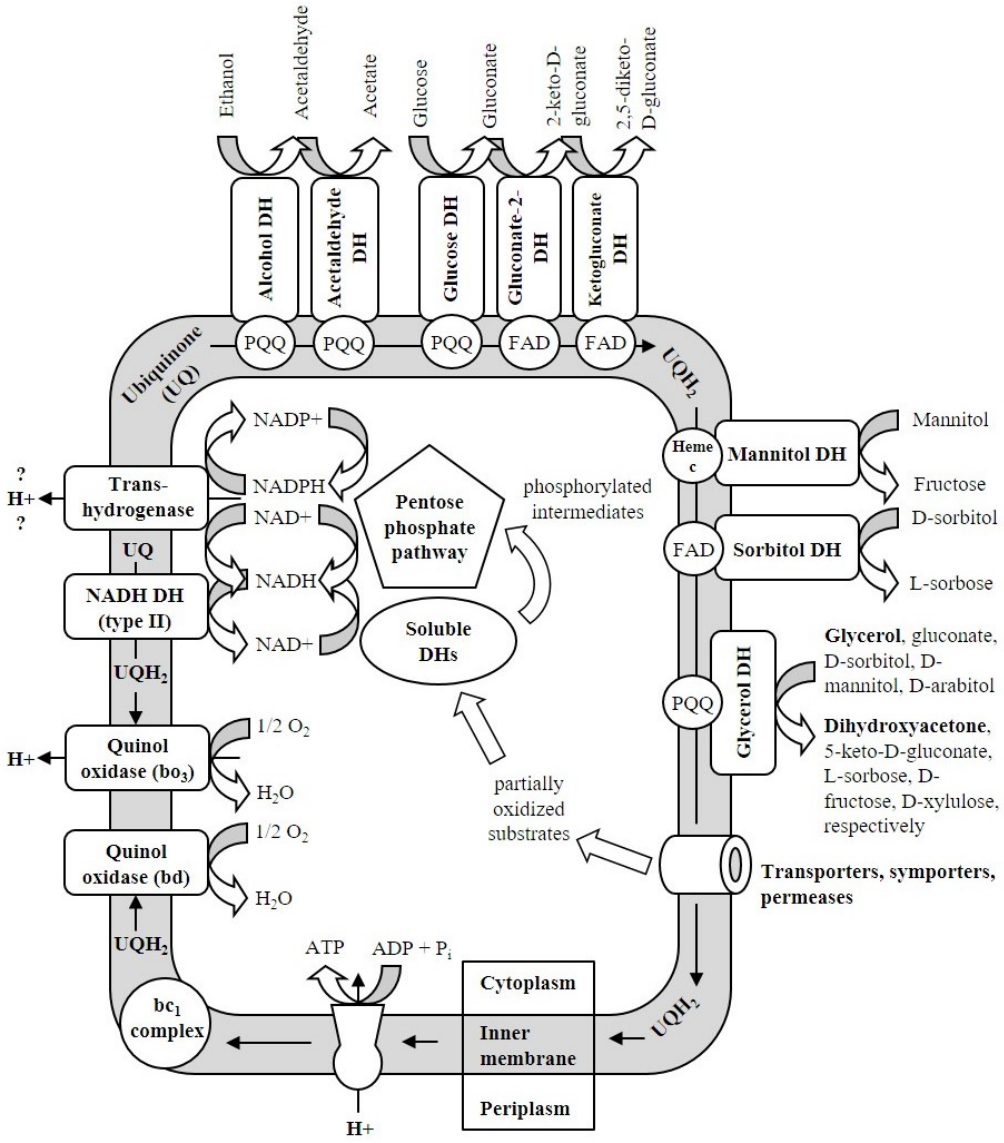


Figure 1. The Metabolism of *Gluconobacter oxydans*. Acetic acid bacteria such as *Gluconobacter oxydans* possess membrane-bound dehydrogenases (DHs) in the periplasm that allow them to partially oxidize sugars, alcohols, polyols, and aldehydes. The resulting products are excreted fully into the medium, and the membrane-bound dehydrogenases donate electrons directly to ubiquinone (UQ) in the respiratory chain, as they contain PQQ, FAD, and heme c prosthetic groups. Additionally, soluble dehydrogenases carry out $NAD(P)^+$-dependent oxidations in the cytoplasm, and the resulting phosphorylated intermediates are incorporated into the oxidative pentose phosphate pathway for biosynthesis. The respiratory chain contains a type II NADH dehydrogenase, which does not translocate protons; and two terminal ubiquinol (UQH_2) oxidases, one that establishes proton-motive force (bo_3 type) and one that cannot pump protons (bd type). *G. oxydans* does not possess cytochrome c oxidase (complex IV); therefore, the bc_1 complex (complex III) does not serve as a coupling site. However, *G. oxydans* does possess a membrane-bound transhydrogenase that may serve to regenerate $NADP^+$ and to extrude protons.

G. oxydans oxidizes substrates via an arsenal of membrane-bound dehydrogenases that is unparalleled by other bacterial taxa (Figure 1) (Deppenmeier and Ehrenreich 2009; Matsushita et al. 2004). These enzymes possess active sites in the periplasm, permitting interaction with substrates from the medium, and allowing for rapid deposition of products (Matsushita et al. 1994; Prust et al. 2005). What is more, the membrane-bound dehydrogenases channel electrons directly into the respiratory chain via pyrroloquinoline quinone (PQQ), flavin adenine dinucleotide (FAD), and heme c prosthetic groups, which donate electrons to ubiquinone in the cell membrane (Deppenmeier et al. 2002; Deppenmeier et al. 2009; Matsushita et al. 2004). It may seem disadvantageous that *G. oxydans* does not fully oxidize substrates, but its metabolism is befitting to its natural habitat. AAB occur where carbon sources are abundant, and they can oxidize them faster than other microorganisms via their periplasmic dehydrogenases. The resulting byproducts are not readily usable by competitors and organic acids accumulate to acidify the environment, inhibiting the growth of less acidophilic bacteria (Deppenmeier and Ehrenreich 2009; Matsushita et al. 2004; Prust et al. 2005).

The myriad of excreted metabolites does not go to waste; when primary nutrients become scarce, *G. oxydans* enters a secondary growth phase in which the partially oxidized byproducts are transported into the cytoplasm and further oxidized by soluble dehydrogenases in a $NAD(P)^+$-dependent manner. These secondary oxidations are more complex and mainly serve to provide intermediates for biosynthesis rather than for energy production (Deppenmeier and Ehrenreich 2009; Matsushita et al. 1994; Prust et al. 2005; Rauch et al. 2010). Interestingly, *G. oxydans* lacks several enzymes required for typical central carbon metabolism. For example, the genes for phosphofructokinase,

phosphoenolpyruvate synthase, succinate dehydrogenase, and ketoglutarate dehydrogenase are missing. Consequently, glycolysis (the Embden-Meyerhof-Parnas pathway), gluconeogenesis, and the tricarboxylic acid cycle are incomplete (Deppenmeier and Ehrenreich 2009; Prust et al. 2005; Raspor and Goranovič 2008; Rauch et al. 2010). However, *G. oxydans* does produce all enzymes for the Entner-Doudoroff pathway and the oxidative pentose phosphate pathway, the latter of which is most important for central metabolism in this bacterium (Prust et al. 2005; Rauch et al. 2010; Richhardt et al. 2012).

In *G. oxydans*, the soluble dehydrogenases produce phosphorylated intermediates that are assimilated into the pentose phosphate pathway (Figure 1) (Deppenmeier and Ehrenreich 2009; Prust et al. 2005). This pathway is responsible for generating NADPH for biosynthesis and some NADH for energy production (Rauch et al. 2010). Ultimately, electrons derived from nonphosphorylative oxidations in the periplasm and phosphorylative oxidations in the cytoplasm reduce ubiquinone in the respiratory chain (Matsushita et al. 2004). Electrons from the cytoplasm enter the respiratory chain through a type II NADH dehydrogenase that does not translocate protons (Deppenmeier and Ehrenreich 2009; Prust et al. 2005). *G. oxydans* also produces a nicotinamide dinucleotide transhydrogenase. It is predicted that this enzyme serves to regenerate $NADP^+$ by coupling the oxidation of NADPH with the reduction of NAD^+. Based on homology with other bacterial transhydrogenases, this protein may also contribute to proton motive force (Deppenmeier and Ehrenreich 2009; Prust et al. 2005).

Finally, ubiquinol shuttles electrons to one of two terminal oxidases. The genome sequence confirmed that *G. oxydans* produces both a bo_3 and bd type ubiquinol oxidase,

with only the bo_3 type contributing to proton motive force (Prust et al. 2005). The respiratory chain of this microbe lacks a proton-translocating type I NADH dehydrogenase (complex I), as well as cytochrome c oxidase (complex IV). Therefore, the ability of *G. oxydans* to generate proton motive force for ATP synthesis is limited (Deppenmeier and Ehrenreich 2009; Prust et al. 2005). Oddly, the fact that *G. oxydans* cannot produce much ATP may give it a competitive edge in nature. The two quinol oxidases provide a quick sink for removing electrons from the respiratory chain by channeling them to the terminal electron acceptor—oxygen—without generating high membrane potential, which would otherwise hinder rapid oxidation of carbon substrates (Deppenmeier and Ehrenreich 2009; Matsushita et al. 2004; Prust et al. 2005).

Industrial Applications of *G. oxydans*

Acetic acid bacteria have been used to make foods and beverages for centuries, although their very existence was unknown until the advent of microbiology (Raspor and Goranovič 2008). AAB are used worldwide to produce vinegar in processes that often parallel the production of alcoholic beverages. Interestingly, vinegar simply refers to acetic acid that is produced from the microbial oxidation of ethanol, and this form of acetic acid is widely used in food preparations. Vinegar is currently produced by various methods ranging from artisanal fermentation in wooden barrels to industrial-scale production in bioreactors (Deppenmeier et al. 2002; Raspor and Goranovič 2008). Another ancient use of AAB in food preparation is in the production of cocoa. Before cacao beans are dried and roasted, cacao pods are harvested and the pulp and beans within are fermented together. The various organic acids and other metabolites produced

by AAB and other microbes are responsible for the characteristic flavor profile of cocoa and derived products (Raspor and Goranovič 2008).

G. oxydans is an especially important acetic acid bacterium, as it is employed for a variety of applications in biotechnology (Figure 2). Most of these applications take advantage of its arsenal of membrane-bound dehydrogenases that is capable of rapid regiospecific and enantioselective partial oxidation of sugars, alcohols, polyols, and aldehydes to produce organic acids and ketones at high yields (Deppenmeier et al. 2002; Gupta et al. 2001). One of the most important modern biotransformations performed by *G. oxydans* is the regiospecific enantiopure conversion of D-sorbitol to L-sorbose in the periplasmic space (Figure 1) (De Muynck et al. 2007). L-Sorbose is a vital intermediate in the industrial production of L-ascorbic acid, or vitamin C (Deppenmeier et al. 2002; Macauley et al. 2001; Pappenberger and Hohmann 2014; Raspor and Goranovič 2008; Yang and Xu 2016). *G. oxydans* is also used to produce the antidiabetic drug Miglitol (Schedel 2000). This microbe carries out the biotransformation of 1-amino-1-deoxy-D-sorbitol to 6-amino-6-deoxy-L-sorbose, which is an intermediate in Miglitol synthesis (Schedel 2000). Additionally, *G. oxydans* is involved in the production of antivirals. For example, it converts ribitol to L-ribulose, which is a precursor to nucleoside analogue drugs. *G. oxydans* also holds promise to produce shikimate, which is a difficult to synthesize precursor to the anti-flu drug Oseltamivir, as well as some antibiotics (De Muynck et al. 2007; Raspor and Goranovič 2008).

Dihydroxyacetone is another important compound produced by *G. oxydans* through the incomplete oxidation of glycerol by the membrane-bound glycerol dehydrogenase (Figure 1) (De Muynck et al. 2007; Deppenmeier et al. 2002).

Dihydroxyacetone has applications as a tanning agent in the cosmetics industry, and as a precursor for industrial chemicals and pharmaceuticals (De Muynck et al. 2007; Deppenmeier et al. 2002). *G. oxydans* also generates several useful byproducts as part of glucose metabolism (Figure 1). Glucose is first oxidized to D-gluconate. Thereafter, 2-ketogluconate, 2,5-diketogluconate and 5-ketogluconate are formed. These compounds find uses as chelating/sequestering agents, as pharmaceuticals and food additives, and as precursors for other products (De Muynck et al. 2007; Deppenmeier et al. 2002; Gupta et al. 2001).

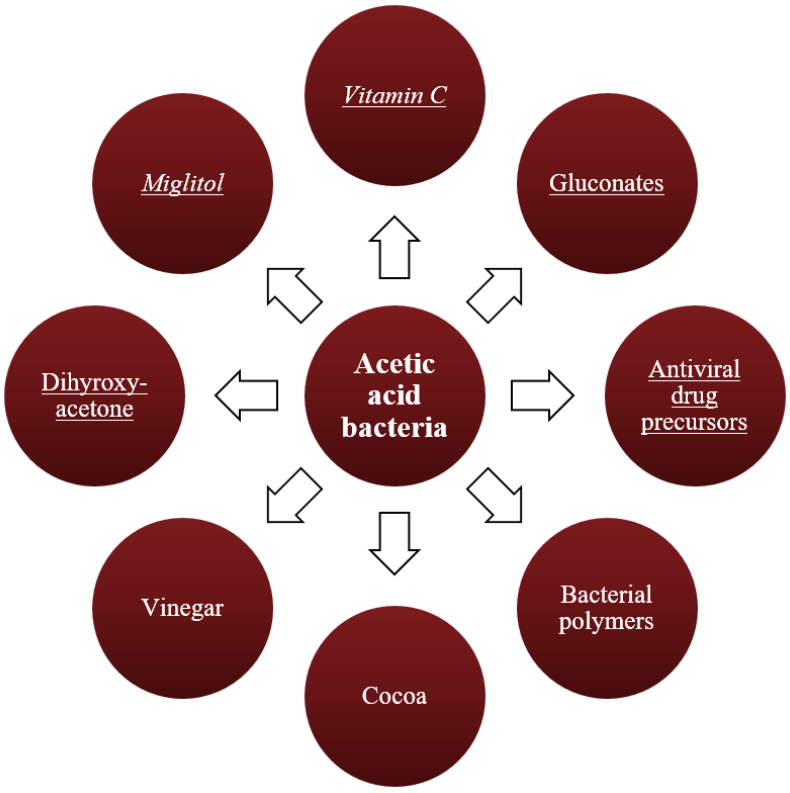


Figure 2. Industrial Applications of Acetic Acid Bacteria. Acetic acid bacteria are widely used in biotechnology, and *Gluconobacter oxydans* is used to make several important products (underlined). Not only does *G. oxydans* produce vital intermediates for the synthesis of Miglitol and vitamin C (italicized), but it also produces completed compounds such as gluconate and gluconate derivatives, dihydroxyacetone, and antiviral drug precursors such as L-ribulose. Other AAB are involved in the production of vinegar, cocoa, and polysaccharides such as bacterial cellulose.

While AAB belonging to the genus *Gluconobacter* specialize in the oxidative modification of small sugar molecules, some members of this family are capable of building sugar polymers such as bacterial cellulose (De Muynck et al. 2007; Raspor and Goranovič 2008). Several genera of bacteria can produce cellulose (Römling and Galperin 2015), but the acetic acid bacterium *Komagataeibacter xylinus* (formerly *Gluconacetobacter xylinus* and *Acetobacter xylinum*) is considered one of the best (Keshk 2014; Raspor and Goranovič 2008). Bacterial cellulose production is complex and not fully understood. Briefly, glucan strands are synthesized in the periplasm and excreted across the outer membrane. Thereafter, individual strands combine to form microfibrils, which in turn form larger bundles that eventually crystalize to form a superfine matrix (De Muynck et al. 2007; Raspor and Goranovič 2008).

Bacterial cellulose is pure, strong, flexible, and has an extremely high water capacity, among other positive attributes, making it superior to plant cellulose (De Muynck et al. 2007; Raspor and Goranovič 2008). Microbial cellulose is a highly desired material, and there has been much research into its production and applications. It can be used as an emulsifier, a bioadsorbant, a drug carrier, as a healing agent for burns and other soft tissue injuries, and could even be used to make artificial blood vessels (Czaja et al. 2006; De Muynck et al. 2007; Keshk 2014; Römling and Galperin 2015; Raspor and Goranovič 2008). Unfortunately, large-scale production remains troublesome. Optimized bioprocesses need to be developed to improve cellulose production in AAB (Czaja et al. 2006; Keshk 2014; Römling and Galperin 2015; Raspor and Goranovič 2008). Additionally, metabolic engineering could improve their ability to efficiently produce important polysaccharides, including cellulose.

Development of Molecular Tools for Acetic Acid Bacteria

AAB have unique metabolic characteristics that suit them for many applications in biotechnology. Namely, their ability to partially oxidize substrates in the periplasmic space is advantageous for the industrial production of important sugar, alcohol, and polyol derivatives. Additionally, AAB are osmotolerant, acidophilic, and produce low biomass, all of which are desirable characteristics (Deppenmeier et al. 2002; Olijve and Kok 1979). Nonetheless, AAB do have some limitations. For example, *Gluconobacter* species are incapable of growth on most disaccharides and all polysaccharides because they do not produce the hydrolytic enzymes required for the degradation of such compounds (Kosciow et al. 2014; Kosciow et al. 2016). Therefore, *G. oxydans* relies on relatively expensive monomeric growth substrates. Despite their widespread use in industry, little progress has been made towards metabolic improvement of AAB. This is partly because few molecular tools exist for genetic manipulation of these microbes (Kallnik et al. 2010). Molecular tools are needed to enable metabolic engineering of AAB to improve their ability to make important nutraceuticals, pharmaceuticals, industrial chemicals, food additives, and consumer products.

It is more pragmatic to improve AAB to account for their weaknesses rather than attempt to replicate their positive attributes in other bacteria. The arsenal of membrane-bound dehydrogenases possessed by *G. oxydans* is not a phenotype that can be easily recreated in other microorganisms. Expression of these enzymes in microbes such as *E. coli* is possible, but they are rendered nonfunctional because *E. coli* cannot produce the necessary prosthetic groups such as PQQ (Meyer et al. 2013). Even if prosthetic groups were to be supplied exogenously or produced endogenously through significant

engineering of *E. coli*, recombinant periplasmic dehydrogenases would have to compete with the natural metabolism of the host for access to substrates. *E. coli* imports growth substrates and oxidizes them in the cytoplasm (Matsushita et al. 1994), and the resulting metabolic intermediates are retained within the cell where they are not readily available for isolation. Under aerobic conditions, these intermediates are fully oxidized to carbon dioxide. In contrast, *G. oxydans* rapidly deposits metabolic byproducts into the medium as part of its normal metabolism (Deppenmeier and Ehrenreich 2009; Matsushita et al. 2004).

Development of molecular tools for AAB is possible, as Kallnik et al. (2010) created expression vectors specifically for *G. oxydans*. These vectors were derived from the broad-host-range plasmid, pBBR1MCS-2, originally produced by Kovach et al. (1995). The resulting plasmids, pBBR1p264 and pBBR1p452, contain *G. oxydans* promotor sequences (p264 and p452), as well as a multiple cloning site and a kanamycin resistance cassette. The promotor sequences were derived from the 5' untranslated regions of genes encoding ribosomal proteins in *G. oxydans*. Thus, these promotors were expected to provide strong constitutive expression of recombinant proteins. Surprisingly, p264 was shown to be a strong promoter of protein expression, while p452 was a moderate-strength promotor (Kallnik et al. 2010).

Two subsequent studies used these expression vectors for metabolic engineering of *G. oxydans* to enable growth on more economical, renewable feedstocks. Kosciow et al. (2014) developed a recombinant strain of *G. oxydans* that could grow on the disaccharide, trehalose, which can enter the periplasm through outer membrane porins. To do this, they created a derivative of the pBBR1p264 vector for expression of the

periplamic trehalase, TreA, from *E. coli*. Additionally, Kosciow et al. (2016) recently developed a *G. oxydans* strain expressing endoxylanase, XynA, from *Bacillus subtilis*. This strain was demonstrated for growth on xylan. Because xylan is a polysaccharide, it must be degraded in the medium by exoenzymes such as endoxylanases. For XynA to be excreted into the medium, the gene for TolB had to be deleted, resulting in a leaky outer membrane phenotype (Kosciow et al. 2016). These studies demonstrate that *G. oxydans* has untapped potential, but the latter study also highlights the fact that *G. oxydans* is not known to possess machinery for export of proteins across the outer membrane. Thus, alternative molecular tools are needed to express recombinant enzymes at the cell surface and extracellular space of AAB. One such tool is surface display.

Surface Display for Biocatalysis

Surface display is the expression of useful proteins at the surface of microbial cells. This involves translational fusion of a protein-of-interest—the passenger protein—with an anchor protein that naturally localizes to the cell surface (Figure 3). Surface display has been demonstrated in gram-positive and gram-negative bacteria, bacterial endospores, and some eukaryotes (mainly yeast) (Schüürmann et al. 2014). The anchor protein is the most important component of a surface display system. In gram-negative microorganisms, anchor proteins either insert into the outer membrane or, less commonly, into the inner membrane (van Bloois et al. 2011). The anchor protein serves two purposes: 1) it contains a signal peptide necessary for secretion across the cell membrane and, 2) it serves as a stable platform for presentation of the passenger protein at the cell surface—thus the term, anchor protein. There are several types of proteins that

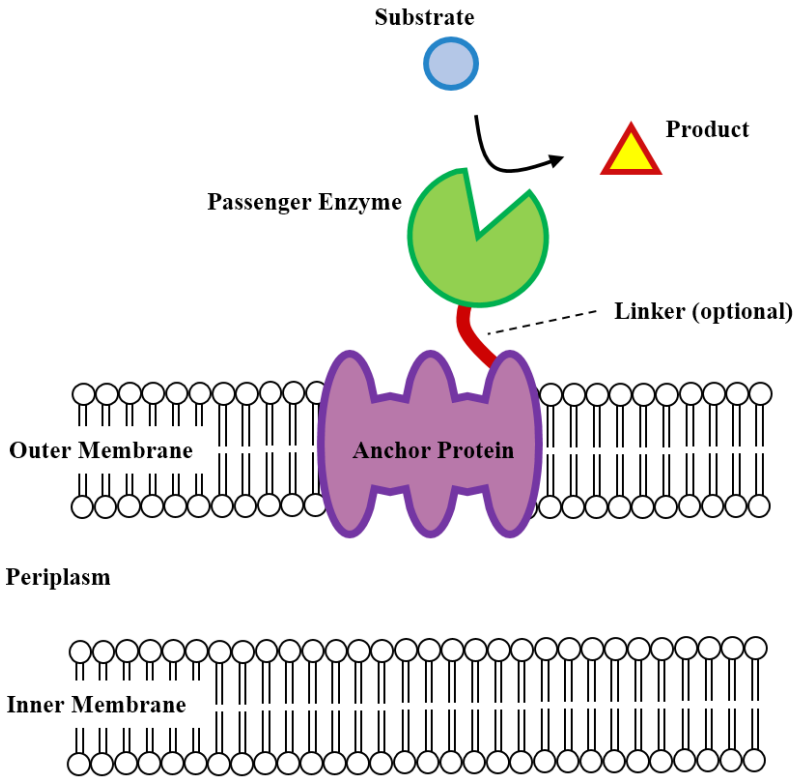


Figure 3. Surface Display in Gram-Negative Bacteria. In gram-negative surface display systems, a recombinant passenger enzyme or protein-of-interest is expressed at the outer membrane via translational fusion to an anchor protein. Anchor proteins are native outer membrane proteins that are secreted across the inner membrane, passed through the periplasm, and then inserted into the outer membrane. Surface display provides a stable platform for biocatalysis because substrates do not have to cross membrane barriers to interact with passenger enzymes, and products are excreted directly into the medium for easy extraction. Enzymes anchored to whole cells can be used for multiple rounds of biocatalysis, as cells can be removed from medium containing the product by centrifugation and transferred to medium containing new substrate.

can be used as anchors in bacterial surface display systems. For example, outer membrane proteins, which fold to form transmembrane β-barrel structures within the outer membrane, are often used (Lee et al. 2003; Schüürmann et al. 2014; van Bloois et al. 2011). Lipoproteins can also function as anchor proteins and components from outer membrane proteins and lipoproteins can be fused to create hybrid anchors. Additionally, extracellular structures such as flagella, fimbriae, and pili can be used to display small

passengers (van Bloois et al. 2011). In nature, some bacteria express monogenic proteins consisting of a C-terminal anchor domain and an N-terminal passenger domain containing a secretory signal. These proteins are named autotransporters and utilization of them for surface display is often referred to as autodisplay (Jose 2006; Jose and Meyer 2007; van Bloois et al. 2011).

Surface display systems have found many uses including, but not limited to, bioremediation, biodetection, drug discovery, vaccine development, and biocatalysis of everything from biofuels to pharmaceuticals (van Bloois et al. 2011; Lee et al. 2003; Jose and Meyer 2007; Saleem et al. 2007; Schüürmann et al. 2014). For removal of environmental contaminants, functional proteins can be expressed at the cell surface that adsorb and sequester toxic heavy metals, while recombinant enzymes can degrade recalcitrant organic compounds such as pesticides (Saleem et al. 2008). For example, Wei et al. (2014) expressed the lead binding protein, PbrR, from *Cupriavidus metallidurans*, at the surface of *E. coli* and demonstrated its ability to detoxify the medium through removal of Pb^{+2} ions. In another study, Yin et al. (2016) developed a system to both detect and sequester elemental mercury contamination by expressing the carboxylesterase E2 enzyme from *Pseudomonas aeruginosa* at the surface of *E. coli*. Carboxylesterase E2 adsorbs Hg^{+2} ions, removing them from the environment. In doing so, its ability to hydrolyze substrates such as *p*-nitrophenyl acetate is blocked, allowing mercury contamination to be detected via a colorimetric assay (Yin et al. 2016).

Surface display is also useful for biomedicine. For example, bacteria expressing antigens at the cell surface can be used as customizable live vaccines (Lee et al. 2003). Epitopes from pathogens such as *Salmonella* can be expressed in resident

microorganisms such as *E. coli*, which can be safely administered (Nhan et al. 2011; Gustavsson et al. 2015). Likewise, lactic acid bacteria have been identified as possible agents for immunization against bacterial toxins, viruses such as HPV, and even parasitic infections through use of surface display (Michon et al. 2016). Similarly, bacterial surface display systems can be used for industrial production of antibodies by serving both as antigen production platforms and to enhance the immune response in animals used for antibody production (Martineau et al. 1991).

The potential applications of surface display are diverse, but perhaps the greatest appeal of surface display is that it can be used as a platform for biocatalysis. To this end, surface display offers several advantages. Expression of enzymes at the cell surface is advantageous over production in the cytoplasm, because it eliminates the need for translocation of substrates and products across membrane barriers and because the product is produced directly in the medium. Therefore, the product can be recovered by centrifugation rather than extraction by cell lysis. Consequently, cells can be recycled for multiple rounds of biocatalysis, decreasing overall costs. Furthermore, anchoring enzymes to the cell surface can increases their stability and catalytic rate, leading to improved yields (Schüürmann et al. 2014). Surface display even facilitates multimerization of complex proteins owing to the fluid mosaic nature of the lipid bilayer (Jose and Meyer 2007; Schüürmann et al. 2014).

In a more general sense, biocatalysis offers significant advantages when compared to traditional chemical synthesis. Most important is the fact that many enzymes have high fidelity: they can modify functional groups at specific sites on substrate molecules (regiospecific), and they can target and produce desired enantiomers of chiral

compounds (enantioselective) (Alcade et al. 2006; Woodley 2008). Thus, biocatalysis can be used to make value-added products—particularly pharmaceuticals and pharmaceutical intermediates/precursors—that are otherwise very difficult to make (Alcade et al. 2006; Schüürmann et al. 2014; Woodley 2008). Biocatalysis is also appealing because it is a form of green chemistry. First, microorganisms are both renewable and biodegradable. Second, whole cells and purified microbial enzymes require relatively benign conditions, while traditional chemical synthesis of pharmaceuticals and other fine chemicals often requires extreme pH, high temperature and pressure, toxic metal catalysts, and large quantities of organic solvents that are often harmful the environment. Third, biocatalysis can reduce the number of steps required to modify substrates containing multiple identical functional groups due to their regio- and enatioselectivity, eliminating the need for protection group chemistry. Fourth, enzymes can be engineered and optimized using mutagenesis (targeted or random) (Alcade et al. 2006; Schüürmann et al. 2014; Woodley 2008). Additionally, advances in metagenomics and computational methods allow us to mine for novel enzymes without the need to culture microbes (Ferrar et al. 2008; Steele et al. 2009; Ufarté et al. 2015).

Localization of Outer Membrane Proteins in Bacteria

In bacteria, proteins destined to be secreted from the cytoplasm require active transport across the inner membrane. For most of these proteins (about 96% in *E. coli*), this transport is facilitated by the Sec translocon (SecYEG), while a minority of exported proteins rely on the twin-arginine translocation pathway (Saleem et al. 2008; Tsirigotaki et al. 2017). Generally, proteins bound for the cell membrane, periplasm, outer

membrane, or extracellular milieu are produced in the cytoplasm as unfolded precursors containing amino-terminal signal peptides. Many integral cell membrane proteins are exported co-translationally, as the nascent protein interacts with a signal recognition particle that routes the ribosomal assembly to the cell membrane. Alternatively, most periplasmic, outer membrane, and extracellular protein precursors interact with chaperones for stabilization and targeting to the cell membrane, where they are post-translationally exported (Tsirigotaki et al. 2017).

Proteins that rely on SecYEG for translocation across or into the cell membrane first bind to a protein called SecA in the cytoplasm. SecA not only associates with SecYEG, but it is also an ATPase that coverts chemical energy into the kinetic energy needed to feed proteins through the SecYEG transmembrane channel (Tsirigotaki et al. 2017). As secreted proteins arrive on the other side of the cell membrane, signal peptidases remove the signal peptides before release into the periplasmic space (De Geyter et al. 2016; Tsirigotaki et al. 2017). Compared to the cytoplasm, the periplasm is generally oxidative rather than reductive and is a relatively unstable compartment that varies in accordance to the extracellular environment. Nonetheless, the periplasm is gel-like as it is densely packed with proteins. A variety of chaperones, isomerases and disulfide-forming enzymes are present to facilitate folding, guide transport, and form intramolecular covalent bonds, among other important tasks (De Geyter et al. 2016).

While transport across the cell membrane is relatively well-characterized (Tsirigotaki et al. 2017), the exact mechanisms of protein transport to the outer membrane remain elusive (Han et al. 2016; Noinaj et al. 2017). The outer membrane of gram-negative bacteria is composed of an inner leaflet that faces the periplasm and an

outer leaflet that faces the extracellular milieu. In contrast to the cell membrane, the outer membrane is asymmetrical because the inner and outer leaflets are composed of different materials (De Geyter et al. 2016). The inner leaflet contains phospholipid and is fluid, while the outer leaflet contains lipopolysaccharide and is considerably less fluid by comparison (De Geyter et al. 2016; Knowles et al. 2009; Mahoney et al. 2016). Unlike inner membrane proteins, OMPs occur almost exclusively as β-barrel structures in their mature form (Knowles et al. 2009). Like cytoplasmic membrane proteins, OMPs initially contain amino-terminal signal peptides that permit secretion across the inner membrane via the aforementioned general secretory SecYEG translocation complex (Gu et al. 2016; Knowles et al. 2009; Tsirigotaki et al. 2017). Next, the signal peptide is cleaved and periplasmic chaperones ferry precursors across the periplasm and peptidoglycan cell wall to the β-barrel assembly machinery complex at the outer membrane (Knowles et al. 2009; Noinaj et al. 2017). Among the many periplasmic chaperones, SurA and Skp are considered to be most important for OMPs (Mahoney et al. 2016; Noinaj et al. 2017).

The β-barrel assembly machinery (BAM) complex is required for insertion of OMPs into the outer membrane. While several studies have found that some OMPs can insert into membranes independently of BAM proteins *in vitro*, spontaneous insertion is unlikely in living cells due to the kinetic barrier imposed by native phospholipid head groups (Gessmann et al. 2014). In *E. coli*, the BAM complex consists of five proteins, BamA-E, the most important of which is BamA, which is conserved in all gram-negative bacteria and has a mitochondrial homologue, Sam50 (Han et al. 2016; Knowles et al. 2009; Noinaj et al. 2017). BamA is a multidomain protein consisting of five periplasmic polypeptide transport-associated (POTRA) domains and its own outer membrane β-barrel

domain (Gu et al. 2016; Han et al. 2016; Knowles et al. 2009; Noinaj et al. 2017). The structure of the BAM complex has been determined (Gu et al. 2016; Han et al. 2016). While the β-barrel component of BamA spans the outer membrane, its POTRA domains interact with some of the BAM lipoproteins in the periplasm to form a ring-like structure beneath the inner leaflet of the outer membrane (Han et al. 2016; Noinaj et al. 2017). Two distinct BamA conformations have been described: an inward-open conformation and an outward-open or lateral-open conformation (Gu et al. 2016; Noinaj et al. 2017). The latter conformation has a periplasmic opening within the aforementioned ring structure and is closed to the extracellular space at the opposite end. The former conformation has a lateral opening on the side of the BamA β-barrel as well as an opening to the outside of the cell at the top of BamA (Gu et al. 2016; Noinaj et al. 2017).

There are several proposed mechanisms to explain OMP biogenesis by the BAM complex, but there is yet to be enough evidence to reach consensus on any particular hypothesis (Noinaj et al. 2017). BamA is a 16-stranded β-barrel with a hydrophobic exterior and a hydrophilic interior, and when it is in the inward-open conformation β-strand 1 and 16 only partially interact with each other (Noinaj et al. 2017). One leading proposed mechanism is that OMPs may feed into the BamA β-barrel through the ring structure when the complex is in the inward-open state (Gu et al. 2016; Noinaj et al. 2017). When BamA switches to the lateral-open confirmation, all interaction is lost between strands 1 and 16, resulting in an opening into the membrane. Essentially, BamA may serve as a template for forming OMPs. This conformational change also causes local destabilization of the membrane, which may provide an entry point for OMPs as they thread out of BamA and into the bilayer. In the lateral-open state, there is also an opening

at the top of BamA which may serve as a portal for extrusion of OMP extracellular loops. In the last step, the first β-strand of the forming β-barrel binds to the last strand, causing the new OMP to detach from BamA and into the outer membrane (Gu et al. 2016; Noinaj et al. 2017).

Most OMPs contain a characteristic amino acid sequence at their C-terminal end that is thought to be a recognition signal for the BAM complex. This signal is identified by a conserved phenylalanine or tryptophan residue at the end of the last β-strand, and hydrophobic residues at positions three, five, seven, and nine from the terminus (Robert et al. 2006). This sequence is often referred to the as the β-signal, and while there is some variation in the specific amino acids comprising this sequence between species, β-signals are probably not species-specific as once thought (Paramasivam et al. 2012).

Development of an Original Surface Display System in *G. oxydans*

Surface display is a potentially valuable molecular tool for AAB because their inner membrane is already crowded with membrane-bound dehydrogenases (Prust et al. 2005). Essentially, surface display would allow extra-cytoplasmic expression of recombinant enzymes without competing for inner membrane space. Furthermore, enzymes expressed at the cell surface would have access to substrates that could not otherwise cross the outer membrane to interact with enzymes in the periplasm, including polymers such as polysaccharides. To this end, the primary goal of this study was to develop an original surface display system for the model acetic acid bacterium, *G. oxydans*. Three potential anchor proteins, INPNC, OprF188, and PgsA were tested for their ability to deliver the alkaline phosphatase enzyme PhoA to the surface of *G.*

oxydans, and the activity of the reporter enzyme was quantified in whole-cell assays. The next goal was to validate the localization of recombinant enzymes at the surface of *G. oxydans* cells expressing the surface display system. Finally, this new molecular tool was optimized by creating the first linker library for a surface display system.

MATERIALS AND METHODS

Bacterial Strains, Culturing, and Storage

Escherichia coli 10-beta (herein *E. coli*) was obtained from New England BioLabs and *Gluconobacter oxydans* 621H was obtained from the Leibnez Institute DSMZ in Germany (DSM No. 2343). *E. coli* was grown in lysogeny broth (LB), composed of 1% NaCl, 1% tryptone, 0.5% yeast extract, and streptomycin (100 µg/mL) to maintain pure cultures. Agar was added to 1.5% when solid media was needed. Cultures were incubated in standard culture tubes or baffled flasks at 37°C with shaking at 200 rpm. Kanamycin was added to 50 µg/mL for plasmid maintenance in recombinant strains. Super optimal broth with catabolite repression (SOC), composed of 2% tryptone 0.5%, yeast extract, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 2.5 mM NaCl, and 20 mM glucose, was used for post-transformation growth of *E. coli* following electroporation.

G. oxydans was grown in yeast mannitol (YM) broth, composed of 2% mannitol, 0.6% yeast extract, and cefoxitin (50 µg/mL) to maintain pure cultures. Kanamycin was added to 50 µg/mL for plasmid maintenance in recombinant strains. For preparation of electrocompetent cells, *G. oxydans* was grown in either YM broth or electroporation (EP) medium, composed of 8% mannitol, 1.5% yeast extract, 0.5 mL/L glycerol, and 10.1 mM $MgSO_4$ and 13.5 mM $CaCl_2$ (added after autoclaving) (Kallnik et al. 2010). Cultures were incubated in 50–500 mL baffled flasks filled 10–20% volume at 30°C with shaking at 200 rpm. For post-transformation growth of *G. oxydans* following electroporation, either YM broth or EP medium was used. For selection of positive clones after transformation, *G. oxydans* was plated on yeast glucose calcium carbonate (YGC) agar, composed of 2%

glucose, 1.5% agar, 0.7% $CaCO_3$, 0.6% yeast extract, and containing cefoxitin and kanamycin. For long-term storage, bacterial suspensions were supplemented with 15% (v/v) glycerol and stored at -80ºC.

Molecular Techniques

Plasmid DNA was extracted from 5–10 mL *E. coli* cultures using a Thermo Fisher Scientific GeneJET Plasmid Miniprep Kit according to the product literature. Genomic DNA was extracted from *E. coli* using a Sigma Aldrich GenElute Bacterial Genomic DNA Kit according to manufacturer's instructions. Restriction-digested vector DNA was excised from agarose gels following electrophoresis and purified using a Promega Wizard SV Gel and PCR Clean-Up Kit following to the manufacturer's guidelines. Primers were purchased from either Integrated DNA Technologies or Eurofins Genomics (Table 1). The Thermo Fisher Scientific Melting Temperature Calculator Web Tool was used to determine the annealing temperature for each primer pair. Alternatively, the annealing temperature was determined empirically using gradient PCR.

For cloning, DNA was amplified using Thermo Scientific Phusion High-Fidelity DNA polymerase. Phusion PCR mixtures were composed and reactions were performed using the two-step protocol per the manufacturer guidelines. The structure of amplicons was confirmed by agarose gel electrophoresis. Putative recombinant colonies were screened by colony PCR using Thermo Scientific DreamTaq DNA polymerase. Briefly, a small amount of cells from an isolated colony was collected with a toothpick and suspended in a small volume (20–50 µL) of sterile, nuclease-free water, of which 5 µL was used as a template for PCR. The reactions were performed according to the

manufacturer's recommended thermocycler settings. Two primer pairs were used for colony PCR in this study: either pBBR1_F and pASK_R, or oprF_F and pASK_R. The annealing temperature was 50°C and 53°C, respectively. Colony PCR products were analyzed by agarose gel electrophoresis to identify positive transformants.

Table 1. Primers

Name	Sequence 5'→3'	Cut site
pBBR1_F	actcactatagggcgaattg	-
pASK_R	cgcagtagcggtaaacg	-
oprF_F	ATGGAATTC**AGGAGG**TAATATTTatgaaactgaagaacaccttaggc	EcoRI
phoA_F	AATTTACGTAcctgttctggaaaaccggg	Eco105I
phoA_R2	ATATAAGCTTtcatttcagccccagagcggc	HindIII
oprF-FL1_R	AGA*GGATCCGCCGCCGCC*gacgttgtcgcaaacgcc	BamHI
oprF-FL2_R	TAT*GGATCCGCCGCCGCCCGAGCCGCCGCCGCC*gacgttgtcgcaaacgcc	BamHI
oprF-RL1_R	*CTTGGCGGCCGCTTC*gacgttgtcgcaaacgcc	NotI
oprF-RL2_R	*CTTGGCGGCCGCTTCCTTCGCCGCGGCTTC*gacgttgtcgcaaacgcc	NotI
phoA-FL1_F	*GGCGGCGGC*GGATCCcctgttctggaaaaccgg	BamHI
phoA-FL2_F	*GGCGGCGGC*GGATCC*GGCGGCGGCGGCTCG*cctgttctggaaaaccgg	BamHI
phoA-RL1_F	*GAAGCGGCCGC*CAAGcctgttctggaaaaccgg	NotI
phoA-RL2_F	*GAAGCGGCCGCCAAGGAAGCCGCGGCGAAG*cctgttctggaaaaccgg	NotI
phoA-CL_F	ATATTACGTA*ATCGACGGCCGC*GGCTCCcctgttctggaaaaccgg	Eco105I
phoA_IR	gagcgttaagcagctgttcg	-

Lowercase nucleotides denote the annealing portion of the primers with homology to their respective templates, while capitalized nucleotides are synthetic additions to the amplicon. Underlined nucleotides indicate the location of restriction enzyme sites, italicized nucleotides encode for linker sequences, and bold nucleotides denote a ribosomal binding site.

Vectors were cut with Thermo Scientific FastDigest restriction enzymes. Briefly, vector DNA was added in the largest possible quantity (preferably 1–5 µg) to the reaction and incubated at 37ºC, then heat-inactivated following the manufacturer's recommendations, except that the incubation time was doubled for reactions containing

more than 1µg of vector. The digest was analyzed and purified by fractionation in a electrophoretic agarose gel. The desired fragment was then excised from the gel and purified. DNA inserts were cut in a similar manner, but resulting products from these reactions were used directly in ligation reactions without prior electrophoresis.

DNA was fractionated and analyzed via gel electrophoresis using a Tris-acetate-EDTA (TAE) buffer, composed of 2 M Tris-base, 57 mL/L glacial acetic acid, and 100 mL/L 0.5 M EDTA (previously adjusted to pH 8.0). PCR products up to 3 kb were analyzed with 1.0% (w/v) agarose gels and restriction digestions were validated with 0.7% gels. Ethidium bromide was added to a final concentration of 0.5 µg/mL. When necessary, Thermo Fisher Scientific 6X DNA Gel Loading Dye was added to the DNA to a 1X concentration before gel loading. The Thermo Fisher Scientific GeneRuler 1 kb DNA ladder was used as a standard in all gels.

Ligations were performed using Thermo Scientific T4 DNA Ligase according to product literature, except that the suggested incubation time of 10 minutes was extended. Typically, 100 ng of vector was added to ligation reactions, and each insert was added to an insert:vector molar ratio of 3:1. Ligations were incubated at 22°C for 1–2 hours followed by heat inactivation at 65°C for 10 minutes. Alternatively, ligations were incubated at 16°C overnight followed by heat inactivation at 65°C for 10 minutes.

Electrocompetent *E. coli* cells were prepared according to the Bio-Rad MicroPulser manual. Briefly, LB broth was inoculated with 1/100 volume of an overnight *E. coli* culture and incubated at 37°C with shaking at 200 rpm. Cells were grown to an optical density ($OD_{600\ nm}$) of 0.5–0.7 and then placed on ice for 20 minutes. Chilled cells were transferred to centrifuge tubes and centrifuged at 4,000 × *g* for 15 minutes at 4°C.

The resulting supernatant was removed by decanting and the remaining cell pellet was then gently resuspended in a volume of sterile, ice-cold 10% glycerol solution equal to that of the culture added to the tube (100% volume). The cells were washed and centrifuged at 4,000 × *g* for 15 minutes at 4°C. The wash was then repeated with at least 50% volume of the glycerol wash solution. Next, each pellet was resuspended in a small amount of wash solution (<2 mL) and transferred to a microcentrifuge tube. Cells were concentrated by centrifugation and the resulting cell pellets were resuspended with 10% glycerol and combined so that the total volume of electrocompetent cells was approximately equal to 0.5% of the volume of the original culture. Cells were used immediately for electroporation or divided into 200-µL aliquots, flash-frozen in liquid nitrogen, and stored at -80°C. When stored, glycerol was added to a final concentration of 15% (v/v).

Electrocompetent *G. oxydans* cells were prepared using a modified method from Kallnick et al. (2010). At least 100 mL of YM broth or EP medium was inoculated to an optical density of 0.1 with an overnight culture of *G. oxydans* and incubated at 30°C with shaking at 200 rpm. Cells were grown to an optical density of 0.8–1.2 and then placed on ice for 20 minutes. Chilled cells were transferred to centrifuge tubes and centrifuged at 2,000 × *g* for 10 minutes at 4°C. The resulting supernatant was removed by decanting and the remaining cell pellet was gently resuspended in a volume of sterile, ice-cold HEPES buffer (1 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) equal to that of the culture added to the tube (100% volume). The cells were washed via centrifugation at 4,000 × *g* for 10 minutes at 4°C. This wash step was repeated at least twice more with at least 50% volume of wash solution. Finally, one cell pellet was resuspended with 250 µL

of the wash solution, and the resulting suspension was used to resuspend additional cell pellets. Electrocompetent *G. oxydans* cells were used immediately for electroporation. Alternatively, glycerol was added to a final concentration of 15% (v/v) and 200-µL aliquots were flash frozen in liquid nitrogen and stored at -80°C.

Electrocompetent cells were transformed according to the Bio-Rad MicroPulser manual. Briefly, 1 µL of plasmid DNA was added to 40 µL of electrocompetent cells and mixed together by gentle pipetting and transferred to a sterile 0.1 cm electroporation cuvette. *E. coli* was electroporated with a Bio-Rad MicroPulser using the Ec1 setting which delivers a 1.8 kV pulse (field strength: 18 kV/cm), whereas *G. oxydans* was electroporated using the Agr setting which delivers a 2.2 kV pulse (field strength: 22 kV/cm). Immediately, 1 mL of SOC medium was added to pulsed *E. coli* and either YM or EP medium was added to pulsed *G. oxydans* cells. The cell suspension was then mixed via pipetting and transferred to a culture tube or baffled flask. Cells were incubated at 37°C for 1 hour for *E. coli* or 30°C for 6–16 hours for *G. oxydans* with shaking at 200 rpm. Next, the culture was pelleted at 2,000 × *g* for 5 minutes and cells were resuspended in ~100 µL of the supernatant and ~10–20 µL and ~80–90 µL of the suspension was plated on selective media: LB containing streptomycin and kanamycin agar for *E. coli* or YGC agar containing cefoxitin and kanamycin for *G. oxydans*.

Sanger sequencing was done by Eurofins Genomics (Eurofins MWG Operon), and samples were prepared according to their recommendations. Briefly, purified plasmid was added to a final concentration of 100 ng/µL, primer was added to 8.3 pmol/µL, and nuclease-free water was added to a final volume of 12 µL. Sequence data was aligned to

predicted sequences and analyzed using BioEdit (Ibis Biosciences), Vector NTI (Thermo Fisher Scientific), and CLC Main Workbench (QIAGEN Bioinformatics) software.

Construction of Plasmids

Constructs were designed *in silico* using Clone Manager Basic 9 software (Scientific and Educational Software). SnapGene Viewer (GSL Biotech) was used to generate plasmid maps. An anchor library for surface display in *G. oxydans* was reconstructed from plasmids originally made by Pearson (2014) (see Table 2). The gene encoding PhoA was amplified without its native signal sequence from the *E. coli* genome using primers phoA_F and phoA_R2, containing extended Eco105I and HindIII restriction sites, respectively (Table 1). Vectors pBBR1p264-INPNC-phoA-ST, pBBR1p452-INPNC-phoA-ST, pBBR1p264-pgsA-phoA-ST, and pBBR1p452-pgsA-phoA-ST were cut with Eco105I and HindIII to remove the *pho-ST* sequence, and vectors pBBR1p264-oprF-ST and pBBR1p452-oprF-ST were cut with Eco105I and HindIII to remove the Strep-tag sequence. Similarly-cut *phoA* amplicon was ligated into the linearized vectors in frame with either *INPNC*, *pgsA*, or *oprF188* to produce the plasmids, pBBR1p264-INPNC-phoA, pBBR1p452-INPNC-phoA, pBBR1p264-pgsA-phoA, pBBR1p452-pgsA-phoA, pBBR1p264-oprF-phoA, and pBBR1p452-oprF-phoA (Table 2).

To incorporate a cleavable linker into the OprF-PhoA fusion protein, a 5'-extended version of the PhoA gene, *CL-phoA*, was amplified using primers phoA-CL_F and phoA_R2 (Table 1). Primer phoA-CL_F encodes the amino acid sequence Ile-Asp-Gly-Arg, which comprises the cleavable linker (CL), and contains an Eco105I site. Next,

Table 2. Plasmids used in this study

Plasmid	Name	Description	Source
P1	pBBR1p264-ST	pBBR1MCS-2 derivative containing the 5' UTR of *gox0264*, a multiple cloning site (MCS), Kan^R, and a Strep-tag (ST) sequence	1
P2	pBBR1p452-ST	pBBR1MCS-2 derivative containing the 5' UTR of *gox0452*, MCS, Kan^R, and *ST*	2
P3	pBBR1p264-INPNC-ST	P1 derivative containing *INPNC*	3
P4	pBBR1p452-INPNC-ST	P2 derivative containing *INPNC*	3
P5	pBBR1p264-INPNC-phoA-ST	P3 derivative containing *phoA*	3
P6	pBBR1p452-INPNC-phoA-ST	P4 derivative containing *phoA*	3
P7	pBBR1p264-oprF-ST	P1 derivative containing *oprF188*	3
P8	pBBR1p452-oprF-ST	P2 derivative containing *oprF188*	3
P9	pBBR1p452-oprF-phoA-ST	P8 derivative containing *phoA*	3
P10	pBBR1p264-pgsA-ST	P1 derivative containing *pgsA*	4
P11	pBBR1p452-pgsA-ST	P2 derivative containing *pgsA*	4
P12	pBBR1p264-pgsA-phoA-ST	P11 derivative containing *phoA*	4
P13	pBBR1p452-pgsA-phoA-ST	P12 derivative containing *phoA*	4
P14	pBBR1p264-INPNC-phoA	P5 with *phoA-ST* removed, replaced by *phoA*	5
P15	pBBR1p452-INPNC-phoA	P6 with *phoA-ST* removed, replaced by *phoA*	5
P16	pBBR1p264-oprF-phoA	P7 with *ST* removed, replaced by *phoA*	5
P17	pBBR1p452-oprF-phoA	P8 with *ST* removed, replaced by *phoA*	5
P18	pBBR1p264-pgsA-phoA	P13 with *phoA-ST* removed, replaced by *phoA*	5
P19	pBBR1p452-pgsA-phoA	P14 with *phoA-ST* removed, replaced by *phoA*	5
P20	pBBR1p264-oprF-FL1-phoA	P7 with *oprF-ST* removed, replaced by *oprF-FL1-phoA* encoding a $(GGGGS)_1$ flexible linker	5
P21	pBBR1p452-oprF-FL1-phoA	P8 with *oprF-ST* removed, replaced by *oprF-FL1-phoA* encoding a $(GGGGS)_1$ flexible linker	5
P22	pBBR1p264-oprF-FL2-phoA	P7 with *oprF-ST* removed, replaced by *oprF-FL2-phoA* encoding a $(GGGGS)_2$ flexible linker	5
P23	pBBR1p452-oprF-FL2-phoA	P8 with *oprF-ST* removed, replaced by *oprF-FL2-phoA* encoding a $(GGGGS)_2$ flexible linker	5
P24	pBBR1p264-oprF-FL3-phoA	P7 with *oprF-ST* removed, replaced by *oprF-FL3-phoA* encoding a $(GGGGS)_3$ flexible linker	5
P25	pBBR1p452-oprF-FL3-phoA	P8 with *oprF-ST* removed, replaced by *oprF-FL3-phoA* encoding a $(GGGGS)_3$ flexible linker	5
P26	pBBR1p264-oprF-RL1-phoA	P7 with *oprF-ST* removed, replaced by *oprF-RL1-phoA* encoding a $(EAAAK)_1$ rigid linker	5
P27	pBBR1p452-oprF-RL1-phoA	P8 with *oprF-ST* removed, replaced by *oprF-RL1-phoA* encoding a $(EAAAK)_1$ rigid linker	5
P28	pBBR1p264-oprF-RL2-phoA	P7 with *oprF-ST* removed, replaced by *oprF-RL2-phoA* encoding a $(EAAAK)_2$ rigid linker	5
P29	pBBR1p452-oprF-RL2-phoA	P8 with *oprF-ST* removed, replaced by *oprF-RL2-phoA* encoding a $(EAAAK)_2$ rigid linker	5
P30	pBBR1p264-oprF-RL3-phoA	P7 with *oprF-ST* removed, replaced by *oprF-RL3-phoA* encoding a $(EAAAK)_3$ rigid linker	5
P31	pBBR1p452-oprF-RL3-phoA	P8 with *oprF-ST* removed, replaced by *oprF-RL3-phoA* encoding a $(EAAAK)_3$ rigid linker	5
P32	pBBR1p452-oprF-CL-phoA	P18 derivative with *phoA* removed, replaced by *CL-phoA* encoding a Factor Xa cleavable linker	5

Sources: 1. Zeiser et al. 2014; 2. Schweiger et al., unpublished results; 3. Pearson 2014; 4. Pearson, unpublished results; 5. This study.

vector pBBR1p452-oprF-phoA was digested with Eco105I and HindIII to remove the *phoA* sequence. Finally, the *CL-phoA* amplicon was cut with Eco105I and HindIII and ligated into the linearized vector in frame with *oprF188* to produce the plasmid, pBBR1p452-oprF-CL-phoA (Table 2).

A library of 12 constructs for expression of six OprF-PhoA fusion proteins with varying linkers was assembled using the technique described by Li et al. (2016). Construction of this library required eight linker system inserts: four *oprF188* derivatives containing linker sequences at their 3' ends and four *phoA* derivatives containing linker sequences at their 5' ends. The inserts, *oprF-FL1*, *oprF-FL2*, *oprF-RL1*, and *oprF-RL2*, were amplified from the plasmid, pBBR1p264-oprF-ST, using the forward primer, oprF_F, and the respectively-named reverse primers, oprF-FL1_R, oprF-FL2_R, oprF-RL1_R, and oprF-RL2_R (Table 1). Here, FL and RL are abbreviations for flexible linker and rigid linker, respectively, and the number denotes the number of times the corresponding flexible (i.e. GGGGS) or rigid (i.e. EAAAK) motif is repeated in the encoded linker. Primer oprF_F contains an EcoRI site as well as the prokaryotic ribosomal binding site and a start codon. Primers oprF-FL1_R and oprF-FL2_R contain BamHI sites and primers oprF-RL1_R and oprF-RL2_R contain NotI sites. The inserts, *FL1-phoA*, *FL2-phoA*, *RL1-phoA*, and *RL2-phoA* were amplified from the *E. coli* genome using the reverse primer, phoA_R2, containing a stop codon and a HindIII site; and the respectively-named forward primers, FL1-phoA_F and FL2-phoA_F containing BamHI sites and RL1-phoA_F and RL2-phoA_F containing NotI sites (Table 1). These forward primers were also designed to exclude the sequence for the native signal peptide.

Next, the linker system inserts were cut with their corresponding restriction enzymes—all *oprF188* inserts were cut with EcoRI, all *phoA* inserts were cut with HindIII, all flexible linker (*FL*) inserts were cut with BamHI, and all rigid linker (*RL*) inserts were cut with NotI. In parallel, vectors pBBR1p264-oprF-ST and pBBR1p452-oprF-ST were digested with EcoRI and HindIII to remove the *oprF-ST* sequence. Finally, the inserts were ligated into the linearized vectors in a combinatorial fashion to produce the plasmids, pBBR1p264-oprF-FL1-phoA, pBBR1p452-oprF-FL1-phoA, pBBR1p264-oprF-FL2-phoA, pBBR1p452-oprF-FL2-phoA, pBBR1p264-oprF-FL3-phoA, pBBR1p452-oprF-FL3-phoA, pBBR1p264-oprF-RL1-phoA, pBBR1p452-oprF-RL1-phoA, pBBR1p264-oprF-RL2-phoA, pBBR1p452-oprF-RL2-phoA, pBBR1p264-oprF-RL3-phoA, and pBBR1p452-oprF-RL3-phoA (Table 2).

Assays

Alkaline phosphatase assays were conducted by combining whole cells with substrate-buffer in standard 96-well plates using an optimized method of Kosciow et al. (2014). Experimental cultures were seeded from overnight cultures and cells were grown to mid-late exponential phase (approximately $OD_{600\ nm}$ 0.8–1.0 for *E. coli* and 0.6–0.9 for *G. oxydans*), and exact optical density measurements were recorded. Immediately prior to the assays, the substrate, *p*-nitrophenylphosphate, was added to a concentration of 1.25 mM to alkaline phosphatase buffer (1 M Tris-base, 10 mM $MnSO_4$, 10 mM $ZnSO_4$, pH 8.0) to form the substrate-buffer. To set up the plates, 40 µL of cells were added to each well. Next, 160 µL of substrate-buffer, pre-warmed to 30°C, was added to each well. The plate was then placed in a BioTek EL808 plate reader (pre-warmed to 30°C) and

incubated for 60 minutes with shaking, with measurements taken at 405 nm every 5 minutes.

Phosphatase activity was measured by following the absorbance change at 405 nm resulting from the conversion of the substrate, pNPP, to the product, *p*-nitrophenol. These data are reported as the change in absorbance per hour (ΔA_{405}), normalized by the optical density (OD_{600}) of the bacterial cultures, denoted by the unit, $\Delta A_{405}/(hr \times OD_{600})$. Because this relative unit is only a proxy for enzymatic activity and product yield, volume activity was also calculated. Volume activity is reported as nanomoles of product (*p*-nitrophenol) produced per milliliter of cells per minute, corrected for the optical density of the cultures. Thus, the unit for volume activity is $nmol/(mL \times min \times OD_{600})$, which is equal to $mU/(mL \times OD_{600})$. Volume activity was calculated and normalized to cell number by optical density using the equation, $\frac{\Delta A_{405} \times 5}{(0.018\ nmol^{-1} \cdot mL \cdot cm^{-1}) \times L \times T \times OD_{600}}$, wherein ΔA_{405} is the measured change in absorbance at 405nm, multiplied by 5 to account for the 5-fold dilution of bacterial cells in the assay buffer solution; $0.018\ nmol^{-1} \cdot mL \cdot cm^{-1}$ (equivalent to $18{,}000\ M^{-1} \cdot cm^{-1}$) is the extinction coefficient for *p*-nitrophenol (New England Biolabs product manual); L is the path length of light through the sample, which was 0.56 cm in this study; T (time) is the length of the assay, which was 60 minutes in this study; and OD_{600} is the optical density of each culture.

To conduct the cleavable linker assay, YM broth was inoculated from an overnight culture of *G. oxydans* 621H harboring the plasmid, pBBR1452-oprF-CL-phoA. Cultures were grown to mid-late exponential phase, and the optical density of cultures was recorded. Two 500-µL aliquots were gently centrifuged at 2,000 × *g* for 5 minutes and the supernatant was removed. Pellets were resuspended in 200 µL of Factor Xa (FX)

buffer, composed of 100 mM NaCl, 20 mM Tris-base, 2 mM $CaCl_2$, with pH adjusted to 8.0. Four microliters of Factor Xa Protease from New England Biolabs were then added to the treatment sample to a final concentration of 0.02 µg/µL. Next, control and treatment samples were incubated for 64 hours at 23°C, with shaking. After incubation, the samples were pelleted at 2,000 × *g* for 5 minutes, and the supernatants were transferred to new microcentrifuge tubes. To ensure that cells were not present in the supernatants, those tubes were centrifuged at 16,100 × *g* for 2 minutes. The cell pellets were then resuspended in 200 µL of FX buffer, and four 40-µL aliquots of whole cells were transferred to four wells in a 96-well plate. In parallel, four 40-µL aliquots of supernatant were transferred to the plate. Finally, a phosphatase assay was conducted as described previously, except that the incubation time was extended to 3 hours with readings taken every 15 minutes.

To determine growth rates of *G. oxydans* strains, cultures were grown in standard 24-well plates and optical density was measured using a FLUOstar Optima (BMG Labtech GmbH) plate reader. To prepare samples, aliquots of YM broth containing appropriate antibiotics were inoculated to $OD_{600\ nm}$ 0.05 from overnight cultures of *G. oxydans* strains. Next, 1 mL of inoculated broth was added to each well in a 24-well plate. The plate was then incubated at 30°C with shaking at 150 rpm for about 23 hours, and absorbance readings were taken at 595 nm approximately every 5 minutes.

R Studio was used to perform statistical analyses and to generate box-and-whisker plots, strip charts, and growth curve graphics (R Core Team 2017). Data were analyzed by performing an analysis of variance (ANOVA) and a *post-hoc* Tukey's HSD test

(q=0.05). The R packages used in this study were dplyr, ggplot2, growthcurver, plyr, multcomp, and reshape2. See the Appendix for R code.

RESULTS

Surface Display in Acetic Acid Bacteria

To enable targeted expression of recombinant enzymes at the cell surface of AAB, an anchor library was constructed to test three anchor proteins for their ability to localize a reporter enzyme to the cell surface of *G. oxydans*. This was achieved by translational fusion of PhoA from *E. coli* to the C-terminal end of these anchor proteins. The three anchor proteins were INPNC, a truncated version of the ice nucleation protein from *Xanthomonas campestris*; OprF188, a truncated version of outer membrane porin F from *Pseudomonas aeruginosa*; and PgsA, the poly γ-glutamate synthesis protein A from *Bacillus subtilis*. The sequences encoding INPNC, OprF188, and PgsA were previously translationally fused to PhoA and a C-terminal Strep-tag (Pearson 2014) (Table 2). Unfortunately, the presence of the Strep-tag appeared to interfere with either outer membrane export or enzymatic function (Pearson 2014), leading to low phosphatase activity. To investigate the influence of the Strep-tag on export and enzyme activity, the Strep-tag sequence was removed from these constructs. Plasmids were digested with Eco105I and HindIII to excise the Strep-tag fragment (Figure 4). Next, the gene encoding PhoA was amplified using a forward primer designed to exclude the coding region for its native periplasmic signal peptide (Figure 5) (Ehrmann et al. 1990; Kosciow et al. 2014; Peterson et al. 2011; Tran et al. 2005). The resulting *phoA* amplicon was cloned into the six linearized vectors to create a plasmid library containing one of two promotors (high-strength p264 and moderate-strength p452), each encoding one of the three anchor proteins fused to PhoA. The resulting constructs (Table 2) were then transformed into *E.*

coli, screened via colony PCR (Figure 6), and confirmed by sequencing. The anchor library is illustrated in Figure 7.

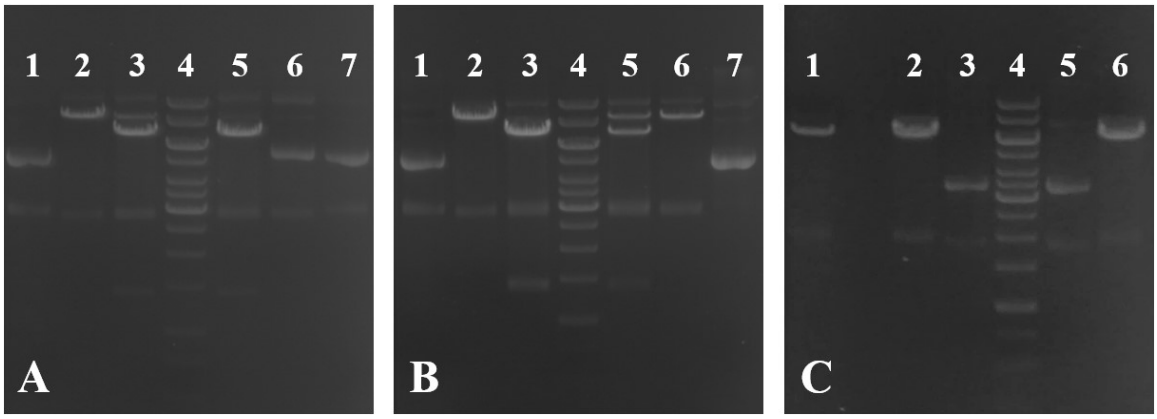


Figure 4. Restriction Digestion of Vectors for the Anchor Library. (A) Vectors p264-INPNC-phoA-ST and 452-INPNC-phoA-ST were cut with Eco105I and HindIII to remove the 1,391 bp *phoA-ST* sequence prior to ligation with similarly-cut *phoA* amplicon to create plasmids for expression of an INPNC-PhoA fusion protein. Lane 1: uncut p264-INPNC-phoA-ST; 7,930 bp circular. Lane 2: p264-INPNC-phoA-ST cut with Eco105I; 7,390 bp linear. Lane 3: p264-INPNC-phoA-ST cut with Eco105I and HindIII; 6,539 and 1,391 bp linear. Lane 4: GeneRuler 1 kb DNA Ladder; 10,000, 8,000, 6,000, 5,000, 4,000, 3,500, 3,000, 2,500, 2,000, 1,500, 1,000, 750, 500, and 250 bp fragments. Lane 5: p452-INPNC-phoA-ST cut with Eco105I and HindIII; 6,462 and 1,391 bp linear. Lane 6: p452-INPNC-phoA-ST cut with HindIII; 7,853 bp linear. Lane 7: uncut p452-INPNC-phoA-ST; 7,853 bp circular. **(B)** Vectors p264-pgsA-phoA-ST and p452-pgsA-phoA-ST were cut with Eco105I and HindIII to remove the 1,391 bp *phoA-ST* sequence prior to ligation with similarly-cut *phoA* amplicon to create plasmids for expression of a PgsA-PhoA fusion protein. Lane 1: uncut p264-pgsA-phoA-ST; 8,113 bp circular. Lane 2: p264-pgsA-phoA-ST cut with Eco105I; 8,113 bp linear. Lane 3: p264-pgsA-phoA-ST cut with Eco105I and HindIII; 6,722 and 1,391 bp linear. Lane 4: GeneRuler 1 kb DNA Ladder. Lane 5: p452-pgsA-phoA-ST cut with Eco105I and HindIII; 6,645 and 1,391 bp linear. Lane 6: p452-pgsA-phoA-ST cut with HindIII; 8,036 bp linear. Lane 7: uncut p452-pgsA-phoA-ST; 8,036 bp circular. **(C)** Vectors p264-oprF-ST and p452-oprF-ST were cut with Eco105I and HindIII to remove the 56 bp Strep-tag sequence prior to ligation with similarly-cut *phoA* amplicon to create plasmids for expression of an OprF188-PhoA fusion protein. Lane 1: p264-oprF-ST cut with Eco105I; 6,280 bp linear. Lane 2: p264-oprF-ST cut with Eco105I and HindIII; 6,224 and 56 bp linear. Lane 3: p264-oprF-ST uncut; 6,280 bp circular. Lane 4: GeneRuler 1 kb DNA Ladder. Lane 5: p452-oprF-ST uncut; 6,203 bp circular. Lane 6: p452-oprF-ST cut with Eco105I and HindIII; 6,147 and 56 bp linear.

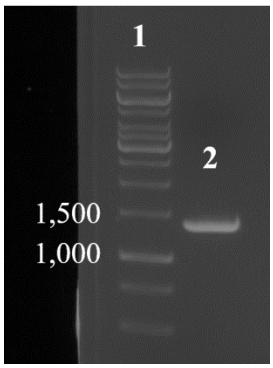


Figure 5. Amplification of *phoA*. A truncated version of the gene encoding PhoA was amplified from the *E. coli* genome using primers phoA_F and phoA_R2, which contain extended Eco105I and HindIII restriction sites, respectively. The coding sequence for the first 26 amino acids was excluded, as those residues constitute a periplasmic signal peptide in the native PhoA translation product. Lane 1: GeneRuler 1 kb DNA Ladder. Lane 2: *phoA* amplicon; 1,358 bp.

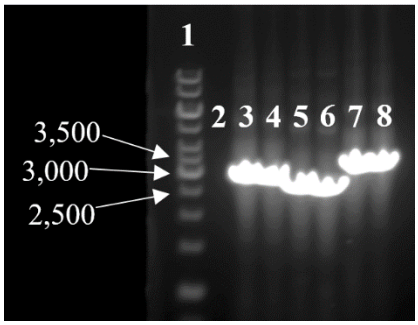


Figure 6. Colony PCR Screen of the Anchor Library. The six constructs constituting the anchor library were transformed into *E. coli* and putative recombinant colonies were screened via colony PCR using primers pBBR1_F and pASK_R. Lane 1: GeneRuler 1 kb DNA Ladder. Lane 2: negative-control reaction (no template). Lane 3: p264-INPNC-phoA; 2,919 bp. Lane 4: p452-INPNC-phoA; 2,919 bp. Lane 5: p264-oprF-phoA; 2,604 bp. Lane 6: p452-oprF-phoA; 2,604 bp. Lane 7: p264-pgsA-phoA; 3,102 bp. Lane 8: p452-pgsA-phoA; 3,102 bp.

As a preliminary test, the anchor library was expressed in E. coli and phosphatase activity was measured in a whole-cell assays (Figure 8, Table 3). Regardless of expression level in *E. coli*, both INPNC-PhoA fusion systems failed to produce a statistically significant absorbance change compared to their respective negative controls ($q=0.976$ using p264 and $q=1.000$ using p452). Likewise, the strain expressing the p264-

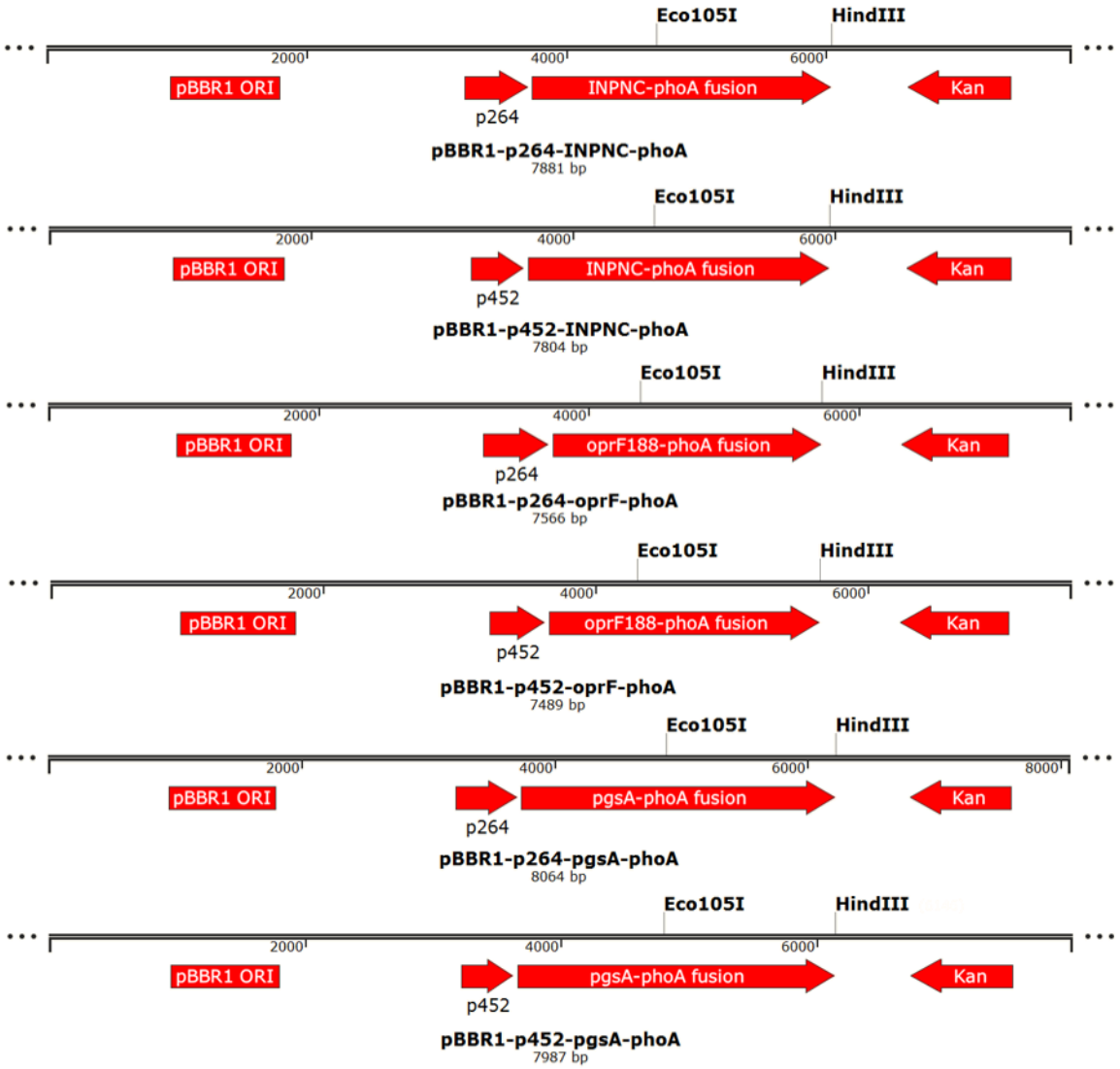


Figure 7. Plasmid Maps of the Anchor Library. Partial plasmid maps of the six constructs constituting the anchor library. The gene encoding the reporter enzyme PhoA was cloned in frame and downstream of genes encoding the anchor proteins, INPNC, OprF188, and PgsA to create *INPNC-phoA*, *oprF188-phoA*, and *pgsA-phoA* gene fusions. Each gene fusion was expressed via two *G. oxydans* expression vectors; one containing a high-strength promotor (p264), and the other containing a moderate-strength promotor (p452). The restriction sites used to insert *phoA*, Eco105I and HindIII, are shown. pBBR1 ORI refers to the origin of replication, and Kan denotes the kanamycin resistance gene. Not shown: a ribosomal binding site precedes each gene fusion.

pgsA-phoA plasmid produced no phosphatase activity (q=1.000). However, the strain containing the p452-pgsA-phoA plasmid did produce statistically significant PhoA activity (q<0.001). However, the highest level of phosphatase activity was observed in

strains expressing the OprF-PhoA fusion protein, both at high ($q<0.001$) and moderate ($q<0.001$) expression levels with promoters p264 and p452, respectively. Product yield was approximately six-fold higher in the p264-oprF-phoA system compared to the p452-oprF-phoA system, which reflects the difference between the high- and moderate-strength promotors in *E. coli*. Additionally, the p452-oprF-phoA system produced more product than the p452-pgsA-phoA system ($q<0.001$).

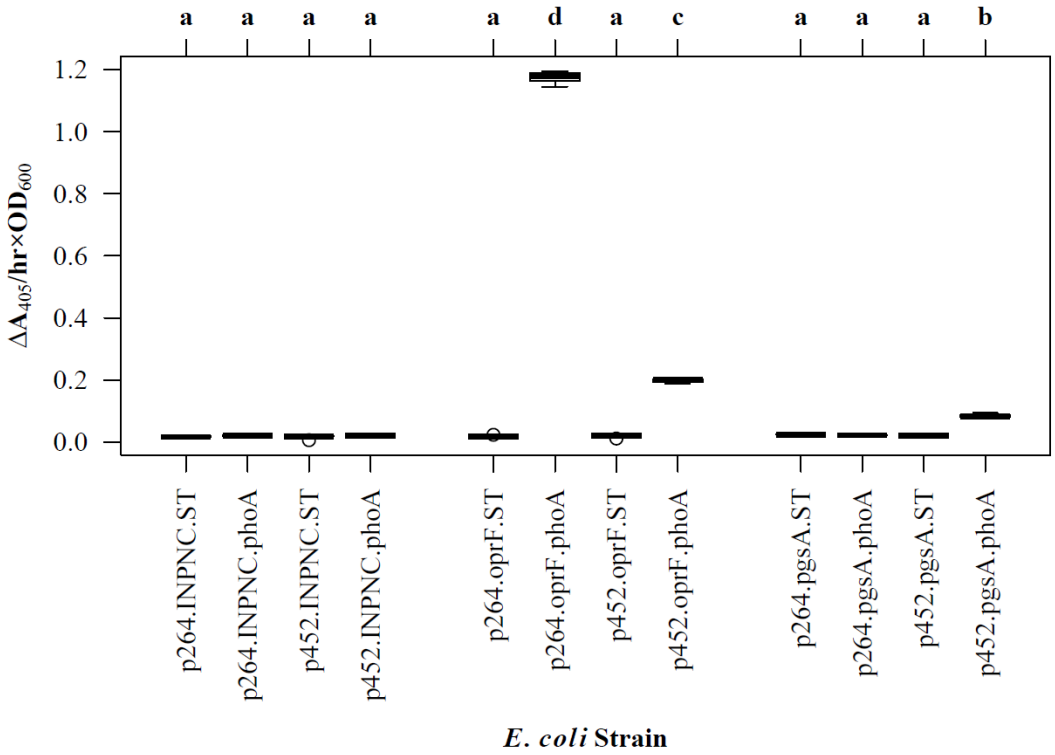


Figure 8. Phosphatase Assay of the Anchor Library in *E. coli*. As a preliminary test, the anchor library was expressed in *E. coli* and the activity of the reporter enzyme, PhoA, was measured in a whole-cell assay. PhoA was translationally fused to three potential anchor proteins, INPNC, OprF188, and PgsA, and the resulting three fusion proteins were produced via two expression vectors: one containing a high-strength promotor (p264), and the other containing a moderate-strength promotor (p452). Each strain expressing a surface display system is matched with a respective negative-control strain that expresses only the anchor protein fused to a Strep-tag (ST). This box-and-whisker plot summarizes the data for eight technical replicates of each *E. coli* strain (N=8). Phosphatase activity is reported as the change in absorbance at 405 nm (ΔA_{405}) per hour, corrected for optical density of the cultures (OD_{600}). The letters above the plot denote statistical groups, as determined by an ANOVA and *post-hoc* Tukey's HSD test.

Table 3. Statistical Analysis: Phosphatase Assay of the Anchor Library in *E. coli*

	p264-INPNC-ST	p264-INPNC-phoA	p452-INPNC-ST	p452-INPNC-phoA	p264-oprF-ST	p264-oprF-phoA	p452-oprF-ST	p452-oprF-phoA	p264-pgsA-ST	p264-pgsA-phoA	p452-pgsA-ST
p264-INPNC-phoA	0.976	-	-	-	-	-	-	-	-	-	-
p452-INPNC-ST	1.000	0.997	-	-	-	-	-	-	-	-	-
p452-INPNC-phoA	0.998	1.000	1.000	-	-	-	-	-	-	-	-
p264-oprF-ST	0.999	1.000	1.000	1.000	-	-	-	-	-	-	-
p264-oprF-phoA	<0.001	<0.001	<0.001	<0.001	<0.001	-	-	-	-	-	-
p452-oprF-ST	0.995	1.000	1.000	1.000	1.000	<0.001	-	-	-	-	-
p452-oprF-phoA	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	-	-	-	-
p264-pgsA-ST	0.400	0.993	0.612	0.941	0.926	<0.001	0.969	<0.001	-	-	-
p264-pgsA-phoA	0.839	1.000	0.951	1.000	0.999	<0.001	1.000	<0.001	1.000	-	-
p452-pgsA-ST	0.969	1.000	0.996	1.000	1.000	<0.001	1.000	<0.001	0.995	1.000	-
p452-pgsA-phoA	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

q values from an ANOVA and *post-hoc* Tukey's HSD test

Given that OprF188 was a successful anchor protein for expression of active PhoA in *E. coli*, the next step was to express the OprF188 constructs in *G. oxydans* and measure phosphatase activity similarly. The OprF188-PhoA fusion protein was successfully expressed in *G. oxydans* when both the high-strength (p264) and moderate-strength (p452) promoters were used (Figure 9). The p264-oprF-phoA strain produced a mean absorbance change of 0.388/(hr×OD_{600}) and a volume activity of 3.21 mU/(mL×OD_{600}), which was significantly different from that of its negative control ($q<0.001$) (Tables 4 and 5). This level of activity is approximately 3-fold lower than that observed when the same construct was expressed in *E. coli.* The p452-oprF-phoA strain produced a mean absorbance change of 0.199/(hr×OD_{600}) and a volume activity of 1.65 mU/(mL×OD_{600}), which was also significantly higher than that of the negative control

(q<0.001). In contrast to the high-strength promotor, there was no reduction in activity when OprF188-PhoA expression was directed by the moderate-strength p452 promoter in *G. oxydans* rather than *E. coli*. In *G. oxydans*, product yield was approximately two-fold higher with the p264-oprF-phoA system compared to the p452-oprF-phoA system, which reflects the difference between the high- and moderate-strength promotors in *G. oxydans*. For comparison, a *G. oxydans* stain containing plasmid p452-oprF-phoA-ST was also assayed for phosphatase activity. This strain failed to produce any product over the course of the assay (q=0.990), suggesting that the Strep-tag may affect the enzymatic activity of PhoA.

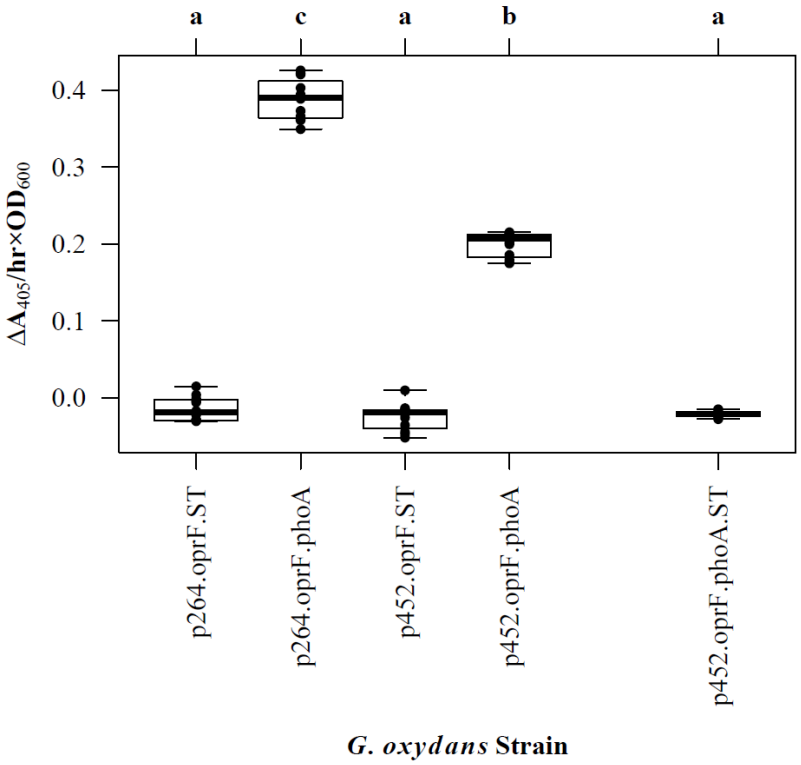


Figure 9. OprF188 for Surface Display in *G. oxydans*. OprF constructs were expressed in *G. oxydans* and phosphatase activity was measured in whole-cell reactions. This box-and-whisker plot summarizes data for 12 replicates of each *G. oxydans* strain (three biological replicates, each with four technical repeats), and the activity of each replicate is shown by the superimposed strip chart. Statistical groups are indicated above the plot and were determined by an ANOVA and *post-hoc* Tukey's HSD test.

Table 4. Statistical Analysis: OprF188 for Surface Display in *G. oxydans*

	p264-oprF-ST	p264-oprF-phoA	p452-oprF-ST	p452-oprF-phoA
p264-oprF-phoA	<0.001	-	-	-
p452-oprF-ST	0.580	<0.001	-	-
p452-oprF-phoA	<0.001	<0.001	<0.001	-
p452-oprF-phoA-ST	0.849	<0.001	0.990	<0.001

q values from an ANOVA and *post-hoc* Tukey's HSD test

Table 5. Phosphatase Activity of *G. oxydans* Strains Expressing the OprF188 Surface Display System

G. oxydans strain	Mean Absorbance Change, $\Delta A_{405}/(hr \times OD_{600})$	Volume Activity, $mU/(mL \times OD_{600})$
p264-oprF-ST	-0.014	-0.12
p264-oprF-phoA	0.388	3.21
p452-oprF-phoA	-0.025	-0.20
p452-oprF-phoA	0.199	1.65
p452-oprF-phoA-ST	-0.021	-0.18

Validation of OprF188 as a Surface Display Anchor in *G. oxydans*

To verify that OprF188 is a suitable anchor for surface display in *G. oxydans*, the OprF188-PhoA fusion protein was modified for the addition of a cleavable linker (CL). Briefly, a 5'-extended version of the gene for alkaline phosphatase, *CL-phoA*, was amplified using a forward primer encoding the amino acid sequence, Ile-Asp-Gly-Arg

(Figure 10A). This sequence is recognized by Factor Xa protease, which cleaves the polypeptide backbone adjacent to the arginine residue (Nagai and Thøgerson 1984; Terpe 2003). Next, plasmid p452-oprF-phoA was digested with Eco105I and HindIII to remove *phoA* (Figure 10B) so that similarly-cut *CL-phoA* amplicon could be inserted. The resulting construct, p452-oprF-CL-phoA, was transformed into *E. coli*, screened via colony PCR (Figure 10C), and confirmed with sequencing.

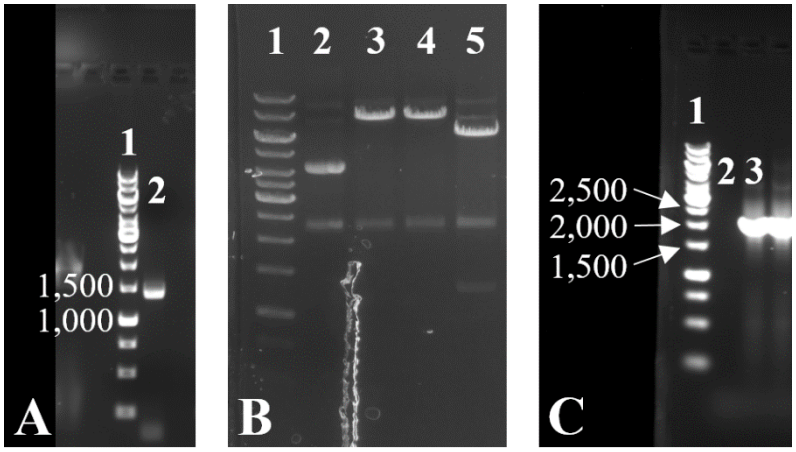


Figure 10. Construction of a Cleavable Surface Display System. (A) To create an OprF188-PhoA fusion protein containing a cleavable linker, a 5'-extended version of the gene encoding PhoA, *CL-phoA*, was amplified from plasmid DNA using primers CL-phoA_F and phoA_R2, containing extended Eco105I and HindIII restriction sites, respectively. Primer CL-phoA_F also encodes the amino acid sequence, Ile-Asp-Gly-Arg, which is cleaved by Factor Xa protease. Lane 1: GeneRuler 1 kb DNA Ladder. Lane 2: *CL-phoA* PCR product; 1,376 bp **(B)** Vector p452-oprF-phoA was cut with Eco105I and HindIII to remove the 1,342 bp *phoA* sequence prior to ligation with the *CL-phoA* amplicon to produce the plasmid, p452-oprF-CL-phoA. Lane 1: GeneRuler 1 kb DNA Ladder; 10,000, 8,000, 6,000, 5,000, 4,000, 3,500, 3,000, 2,500, 2,000, 1,500, 1,000, 750, 500, and 250 bp fragments. Lane 2: uncut p452-oprF-phoA; 7,489 bp circular. Lane 3: p452-oprF-phoA cut with Eco105I; 7,489 bp linear. Lane 4: p452-oprF-phoA cut with HindIII; 7,489 bp linear. Lane 5: p452-oprF-phoA cut with Eco105I and HindIII; 6,147 and 1,342 bp linear. **(C)** The *CL-phoA* amplicon was cut with EcoRI and HindIII and ligated into similarly-cut vector to produce the plasmid, p452-oprF-CL-phoA. This construct was transformed into *E. coli* and putative recombinant colonies were screened via colony PCR using primers oprF_F and pASK_R. Lane 1: GeneRuler 1 kb DNA Ladder. Lane 2: negative-control reaction (no template). Lane 3: p452-oprF-CL-phoA; 2,101 bp.

After confirmation, the cleavable linker construct was transformed into *G. oxydans* and a cleavable linker assay was conducted. Briefly, one sample of cells was treated with Factor Xa, while the other sample was left untreated; both samples were incubated similarly. After incubation, each sample was split into a fraction containing whole cells and a fraction containing supernatant, and both fractions were tested for phosphatase activity. There was no significant difference between the mean phosphatase activity of treated cells compared to untreated control cells (Figure 11) (ANOVA and *post-hoc* Tukey's HSD test, q=0.666). Also, the level of phosphatase activity produced by these cells did not differ from the strain expressing the plasmid, p452-oprF-phoA (without a linker), suggesting that the presence of the cleavable linker itself did not impact activity. The mean level of phosphatase activity observed in the supernatant from treated samples was significantly higher than that from untreated samples (q=0.003).

The Effects of Linkers on Biocatalysis at the Cell Surface

To optimize surface display in AAB, a technique described by Li et al. (2016) was used to integrate a library of linkers into the OprF188 surface display systems. To build a library of linker sequences, specialized primers and overhang PCR was used to add building blocks to *oprF188* and *phoA* (Figure 12A). These building blocks encoded for either flexible or rigid linkers and contained either BamHI or NotI restriction sites, respectively. Essentially, the modified *oprF188* inserts contained abridged linker sequences at their 3' ends and the modified *phoA* inserts contained abridged linker sequences at their 5' ends. Therefore, complete linker sequences resulted from ligation of compatible insert dyads (Figure 13). Prior to ligation, vectors p264-oprF-ST and p452-

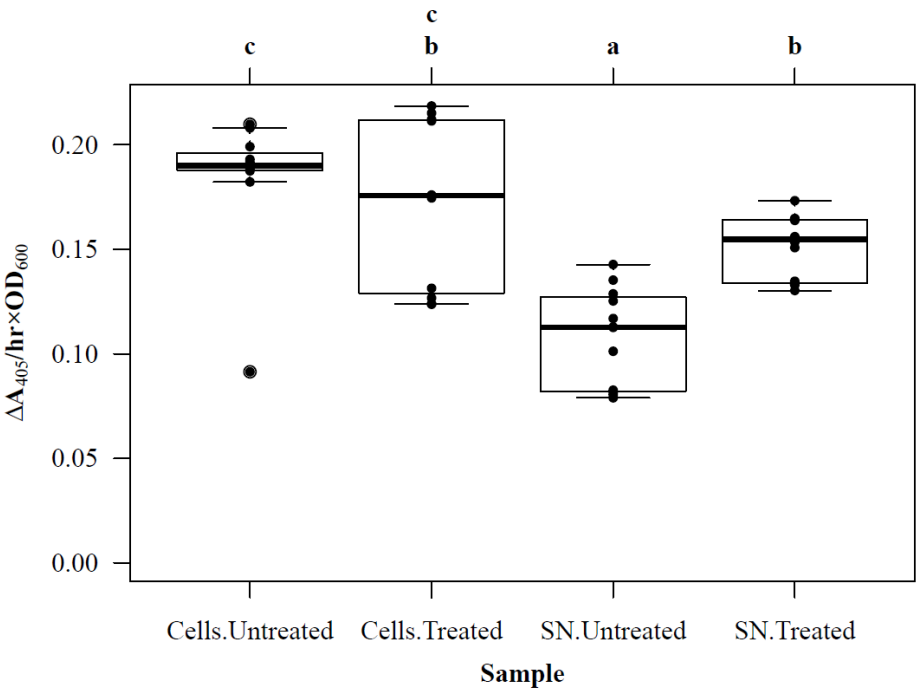


Figure 11. Cleavable Linker Assay. A phosphatase assay was conducted to test the proportion of active PhoA in four sample types prepared from cultures of *G. oxydans* expressing an OprF188-PhoA fusion protein containing a cleavable linker. The sample types were: untreated cells, treated cells, untreated supernatant (SN), and treated supernatant. N=12 for each sample type (three biological replicates, each with four technical repeats), and the phosphatase activity for each replicate is indicated by the superimposed strip chart. The letters above the plot denote statistical groups, as determined by an ANOVA and *post-hoc* Tukey's HSD test.

oprF-ST were cut with EcoRI and HindIII to remove the *oprF-ST* sequence (Figure 12B). The building blocks were then assembled into both expression vectors to create DNA sequences encoding six fusion proteins: three containing flexible linkers composed of the $(GGGGS)_{1–3}$ motif and three containing rigid linkers composed of the $(EAAAK)_{1–3}$ motif. Because the components of the six fusion protein sequences were inserted into both the p264-series and the p452-series expression vectors, the resulting library consisted of 12 constructs varying in promotor strength, linker composition, and linker length. The entire linker library was transformed into *E. coli*, screened via colony PCR (Figure 12C), and confirmed by sequencing.

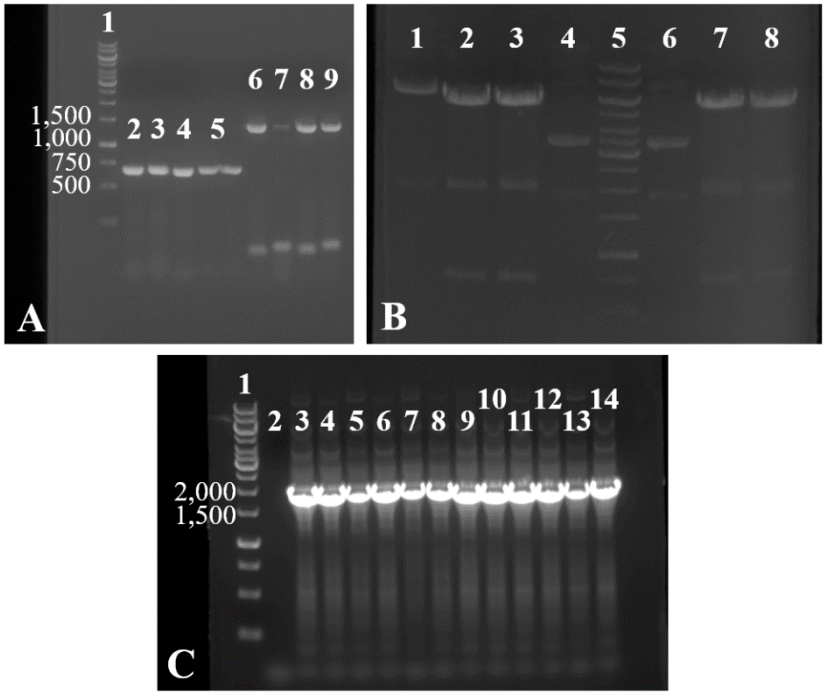


Figure 12. Construction of the Linker Library. (A) Amplification of Linker Inserts. To produce a library of six gene fusions containing variable linker sequences, eight component inserts were amplified via overhang PCR. Four 3'-extended derivatives of the gene encoding OprF188 were amplified using forward primer oprF_F and specialized reverse primers oprF-FL1, oprF-FL2, oprF-RL1, and oprF-RL2_R. In parallel, four 5'-extended derivatives of the gene encoding PhoA were amplified using reverse primer phoA_R2 and specialized forward primers FL1-phoA, FL2-phoA, RL1-phoA, and RL2-phoA_F. Lane 1: GeneRuler 1 kb DNA Ladder. 2: *oprF-FL1*; 677 bp. 3: *oprF-FL2*; 692 bp. 4: *oprF-RL1*; 674 bp. 5: *oprF-RL2*; 689 bp. 6: *FL1-phoA*; 1,363 bp. 7: *FL2-phoA*; 1,378 bp. 8: *RL1-phoA*; 1,363 bp. 9: *RL2-phoA*; 1,378 bp. **(B) Restriction Digestion of Vectors.** Plasmids p264-oprF-ST and 452-oprF-ST were cut with EcoRI and HindIII to remove the 720 bp *oprF-ST* sequence prior to ligation with various linker inserts to create a library of plasmids encoding fusion proteins with variable linkers. Lane 1: p264-oprF-ST cut with HindIII; 6,280 bp linear. 2: p264-oprF-ST cut with EcoRI and HindIII; 5,560 and 720 bp linear. 3: p264-oprF-ST cut with EcoRI and HindIII; 5,560 and 720 bp linear. 4: p264-oprF-ST uncut; 6,280 bp circular. 5: GeneRuler 1 kb DNA Ladder; 10,000, 8,000, 6,000, 5,000, 4,000, 3,500, 3,000, 2,500, 2,000, 1,500, 1,000, 750, 500, and 250 bp fragments. 6: p452-oprF-ST uncut; 6,203 bp circular. 7: p452-oprF-ST cut with EcoRI and HindIII; 5,483 and 720 bp linear. 8: p452-oprF-ST cut with EcoRI and HindIII; 5,483 and 720 bp linear. **(C) Colony PCR Screen of the Linker Library.** The 12 plasmids comprising the linker library were transformed into *E. coli* and putative recombinant colonies were screened via colony PCR using primers oprF_F and pASK_R. Lane 1: GeneRuler 1 kb DNA Ladder. 2: negative-control reaction (no template). 3: p264-oprF-FL1-phoA; 2,086 bp. 4: p452-oprF-FL1-phoA; 2,086 bp. 5: p264-oprF-FL2-phoA; 2,101 bp. 6: p452-oprF-FL2-phoA; 2,101 bp. 7: p264-oprF-FL3-phoA; 2,116 bp. 8: p452-oprF-FL3-phoA; 2,116 bp. 9: p264-oprF-RL1-phoA; 2,086 bp. 10: p452-oprF-RL1-phoA; 2,086 bp. 11: p264-oprF-RL2-phoA; 2,101 bp. 12: p452-oprF-RL2-phoA; 2,101 bp. 13: p264-oprF-RL3-phoA; 2,116 bp. 14: p452-oprF-RL3-phoA; 2,116 bp.

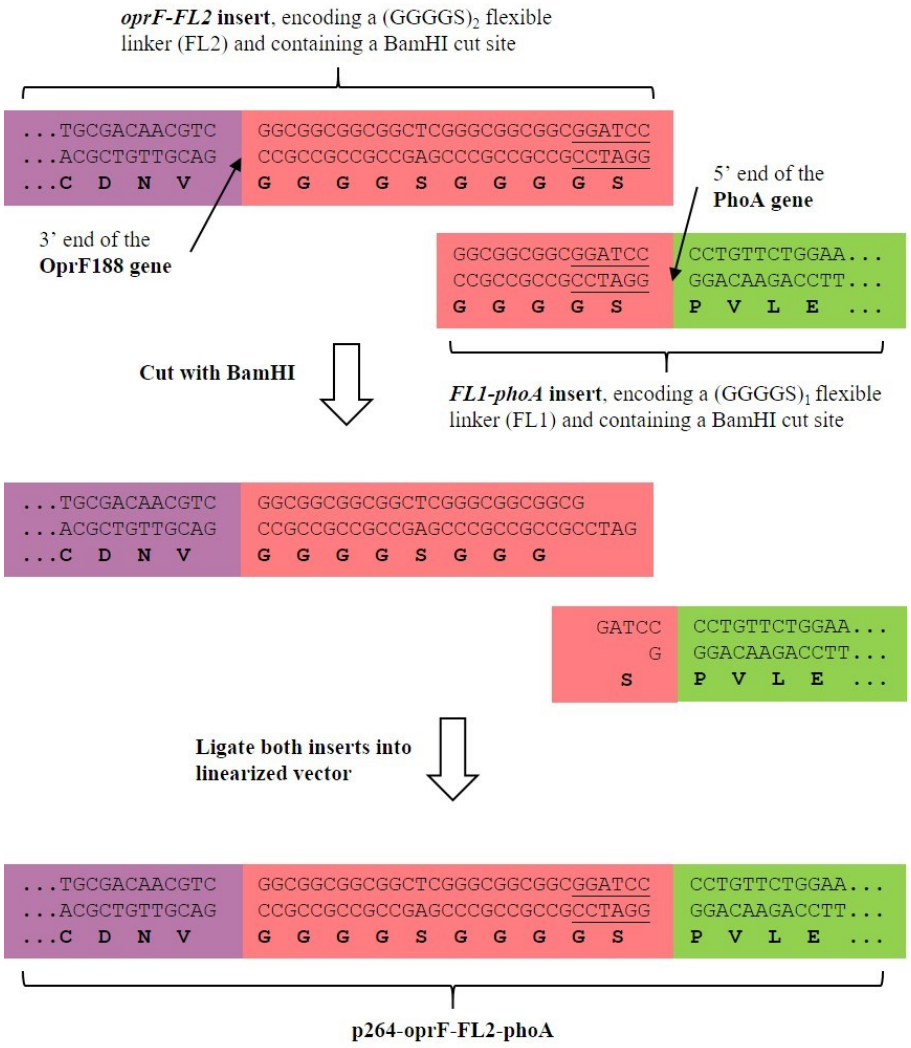


Figure 13. Construction of Plasmid p264-oprF-FL2-phoA. To construct a plasmid for high expression of an OprF188-PhoA fusion protein containing a $(GGGGS)_2$ flexible linker, the first step was to amplify the two inserts, *oprF-FL2* and *FL1-phoA*. Next, *oprF-FL2* was cut with EcoRI and BamHI and *FL1-phoA* was cut with BamHI and HindIII. In parallel, plasmid p264-oprF-ST was cut with EcoRI and HindIII to remove *oprF-ST*. The two inserts were ligated into the linearized vector to produce plasmid p264-oprF-FL2-phoA. The one-letter codes for encoded amino acids are shown (in bold) below the corresponding nucleotide sequence. The BamHI site is underlined.

The linker library was first expressed in *E. coli* and PhoA activity was quantified. The flexible linkers led to a slight decrease in product yield when using the high-expression p264 promoter system (Figure 14, Table 6). While there was no difference between the control (the OprF-PhoA fusion protein without a linker) and the fusion

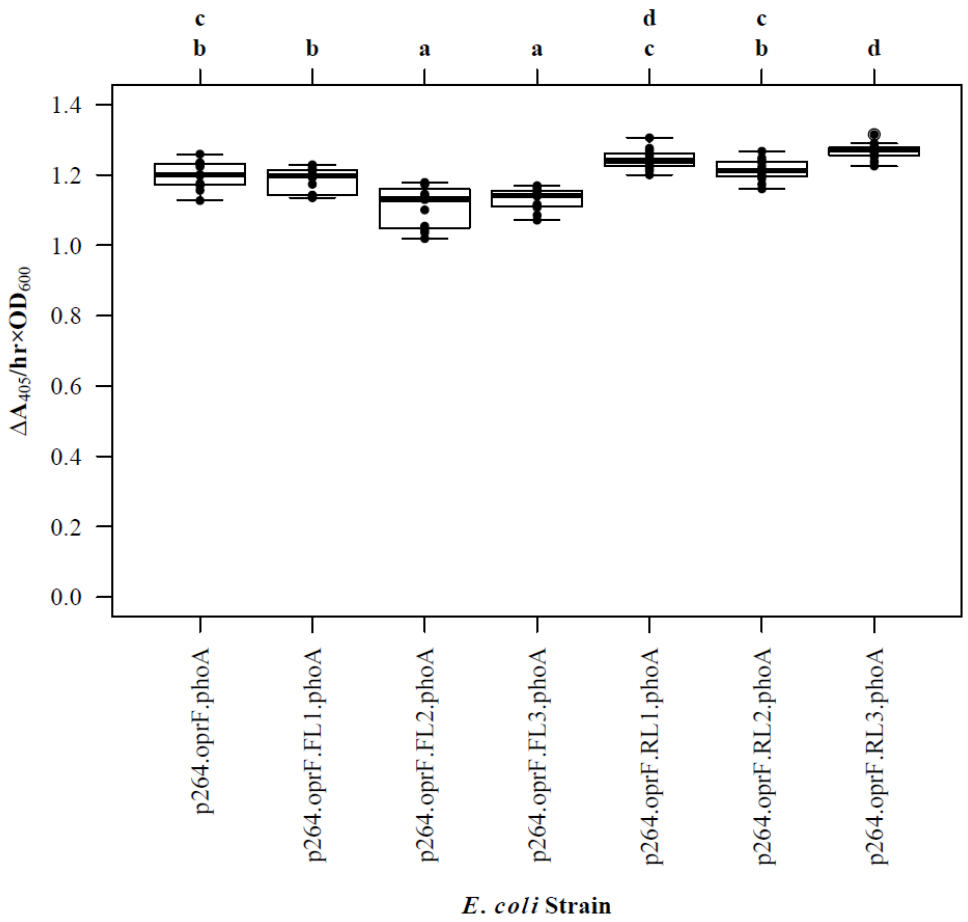


Figure 14. High Expression of the Linker Library in *E. coli*. A library of six OprF188-PhoA fusion proteins with varying linkers was produced via high-expression plasmids in *E. coli*, and activity of the reporter enzyme was quantified in whole-cell reactions. N=12 for each strain, and data points for each replicate are shown by the superimposed strip chart. The letters above the plot denote statistical groups, as determined by an ANOVA and *post-hoc* Tukey's HSD test.

protein containing a small flexible linker (FL1) (q=0.939), the medium (FL2) and large (FL3) flexible linkers led to statistically significant decreases in activity (q<0.001 for each). Meanwhile, the small (RL1) and medium (RL2) rigid linkers did not have any effect on product yield (q=0.071 and 0.969, respectively). The large rigid linker (RL3) led to a slight but statistically significant increase in activity when compared to the control (q<0.001), but there was no difference in activities between the RL1 strain and the RL3 strain (q=0.692). Taken together, these results suggest that, at high expression,

Table 6. Statistical Analysis: High Expression of the Linker Library in *E. coli*

	p264-oprF-phoA	p264-oprF-FL1-phoA	p264-oprF-FL2-phoA	p264-oprF-FL3-phoA	p264-oprF-RL1-phoA	p264-oprF-RL2-phoA
p264-oprF-FL1-phoA	0.939	-	-	-	-	-
p264-oprF-FL2-phoA	<0.001	<0.001	-	-	-	-
p264-oprF-FL3-phoA	<0.001	0.012	0.835	-	-	-
p264-oprF-RL1-phoA	0.071	0.003	<0.001	<0.001	-	-
p264-oprF-RL2-phoA	0.969	0.443	<0.001	<0.001	0.437	-
p264-oprF-RL3-phoA	<0.001	<0.001	<0.001	<0.001	0.692	0.011

q values from an ANOVA and *post-hoc* Tukey's HSD test

rigid linkers had no impact on biocatalysis at the surface of *E. coli*.

In contrast to the high-expression systems, the addition of rigid or flexible linkers increased enzymatic activity when expression was driven by the moderate-strength p452 promoter (Figure 15, Table 7). All the flexible linkers as well as the medium (RL2) and large (RL3) rigid linkers led to equivalent increases in product yield relative to the no-linker control ($q<0.001$ for all). The small rigid linker (RL1) had no effect on product yield ($q=0.525$).

After the effects of linkers on surface display were assessed in *E. coli*, the next step was to evaluate the linker systems in *G. oxydans.* However, attempts to transform p264-oprF-FL2-phoA and p264-oprF-FL3-phoA into *G. oxydans* were unsuccessful,

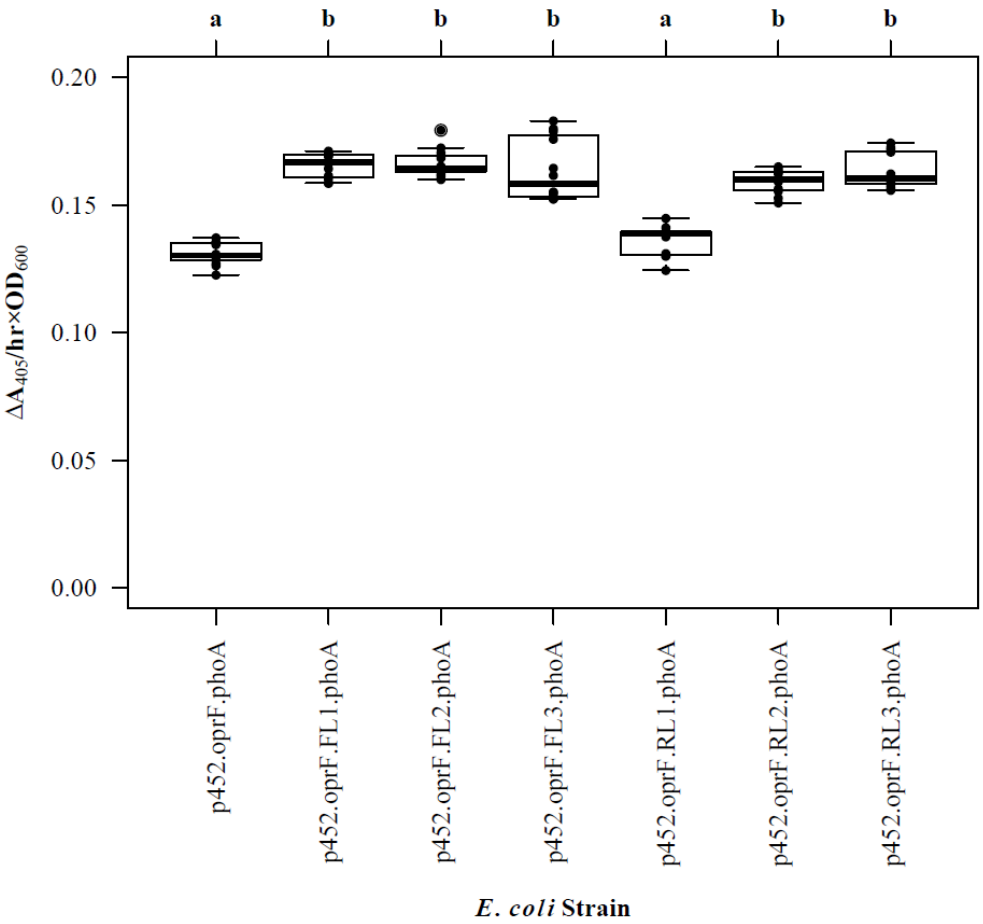


Figure 15. Moderate Expression of the Linker Library in *E. coli*. A library of six OprF188-PhoA fusion proteins with varying linkers was produced via moderate-expression plasmids in *E. coli*, and activity of the reporter enzyme was quantified in whole-cell reactions. N=12 for each strain, and data points for each replicate are shown by the superimposed strip chart. The letters above the plot denote statistical groups, as determined by an ANOVA and *post-hoc* Tukey's HSD test.

suggesting that the medium (FL2) and large (FL3) flexible linkers are toxic to this microorganism at high expression levels. Nonetheless, the remaining four high-expression constructs were stably expressed in *G. oxydans* and phosphatase activity was quantified (Figure 16, Tables 8 and 9). At high expression using the p264 promoter, both the small flexible linker (FL1) and the medium rigid linker (RL2) had no effect on product yield when compared to the no-linker system ($q=0.967$ and 0.765, respectively), while the large rigid linker (RL3) led to a statistically-significant decrease ($q<0.001$).

Table 7. Statistical Analysis: Moderate Expression of the Linker Library in *E. coli*

	p452-oprF-phoA	p452-oprF-FL1-phoA	p452-oprF-FL2-phoA	p452-oprF-FL3-phoA	p452-oprF-RL1-phoA	p452-oprF-RL2-phoA
p452-oprF-FL1-phoA	<0.001	-	-	-	-	-
p452-oprF-FL2-phoA	<0.001	1.000	-	-	-	-
p452-oprF-FL3-phoA	<0.001	0.996	0.967	-	-	-
p452-oprF-RL1-phoA	0.525	<0.001	<0.001	<0.001	-	-
p452-oprF-RL2-phoA	<0.001	0.290	0.162	0.681	<0.001	-
p452-oprF-RL3-phoA	<0.001	0.989	0.942	1.000	<0.001	0.750

q values from an ANOVA and *post-hoc* Tukey's HSD test

Interestingly, the small rigid linker (RL1) not only led to a statistically-significant increase in product yield ($q<0.001$), but it improved biocatalysis by 69%.

At moderate expression with the p452 promoter, the small (FL1) and medium (FL2) flexible linkers led to increased product yields compared to the control ($q=0.003$ and 0.001, respectively) (Figure 17, Tables 9 and 10). In fact, FL2 increased biocatalysis by 64% compared to the control, and by 34% compared to FL1 ($q<0.001$). Conversely, the large flexible linker (FL3) and all the rigid linkers led to significant decreases in product yield compared to the control strain lacking a linker system ($q<0.001$ for all). FL3 decreased activity by 54%, whereas all the rigid linkers decreased biocatalysis by about 70%.

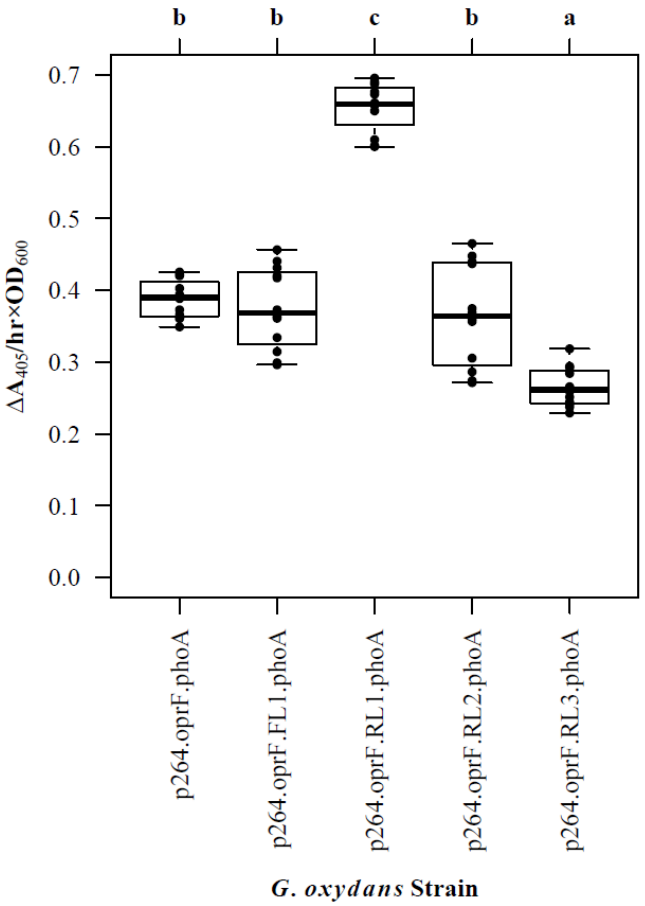


Figure 16. High Expression of the Linker Library in *G. oxydans*. A library of four OprF188-PhoA fusion proteins with varying linkers was produced via high-expression plasmids in *G. oxydans*, and activity of the reporter enzyme was quantified in whole-cell reactions. N=12 for each strain, and data points for each replicate are shown by the superimposed strip chart. The letters above the plot denote statistical groups, as determined by an ANOVA and *post-hoc* Tukey's HSD test.

The Effects of Surface Display on the Growth of *G. oxydans*

To determine the effects of protein production and surface engineering on the health of *G. oxydans* cells, an original method was developed to follow the growth of recombinant *G. oxydans* strains using standard 24-well tissue culture plates. All strains containing OprF188 surface display systems were compared to wild-type *G. oxydans* 621H growth (Figures 18 and 19). At high expression, production of the OprF188 anchor protein fused to a Strep-tag led to a statistically significant increase in mean doubling

Table 8. Statistical Analysis: High Expression of the Linker Library in *G. oxydans*

	p264-oprF-phoA	p264-oprF-FL1-phoA	p264-oprF-RL1-phoA	p264-oprF-RL2-phoA
p264-oprF-FL1-phoA	0.967	-	-	-
p264-oprF-RL1-phoA	<0.001	<0.001	-	-
p264-oprF-RL2-phoA	0.765	0.984	<0.001	-
p264-oprF-RL3-phoA	<0.001	<0.001	<0.001	<0.001

q values from an ANOVA and *post-hoc* Tukey's HSD test

Table 9. Phosphatase Activity of *G. oxydans* Strains Expressing the Linker Library

G. oxydans strain	Mean Absorbance Change, $\Delta A_{405}/(hr \times OD_{600})$	Volume Activity/OD_{600}, mU/ (mL×OD_{600})
p264-oprF-phoA	0.388	3.21
p264-oprF-FL1-phoA	0.376	3.11
p264-oprF-RL1-phoA	0.655	5.42
p264-oprF-RL2-phoA	0.366	3.02
p264-oprF-RL3-phoA	0.265	2.19
p452-oprF-phoA	0.199	1.65
p452-oprF-FL1-phoA	0.244	2.01
p452-oprF-FL2-phoA	0.327	2.70
p452-oprF-FL3-phoA	0.092	0.76
p452-oprF-RL1-phoA	0.053	0.44
p452-oprF-RL2-phoA	0.057	0.47
p452-oprF-RL3-phoA	0.060	0.49

time (97 min) compared to wild-type cells (56 min) (Figure 20) (ANOVA and *post-hoc* Tukey's HSD test, $q<0.001$). In fact, this strain had the second-highest doubling time

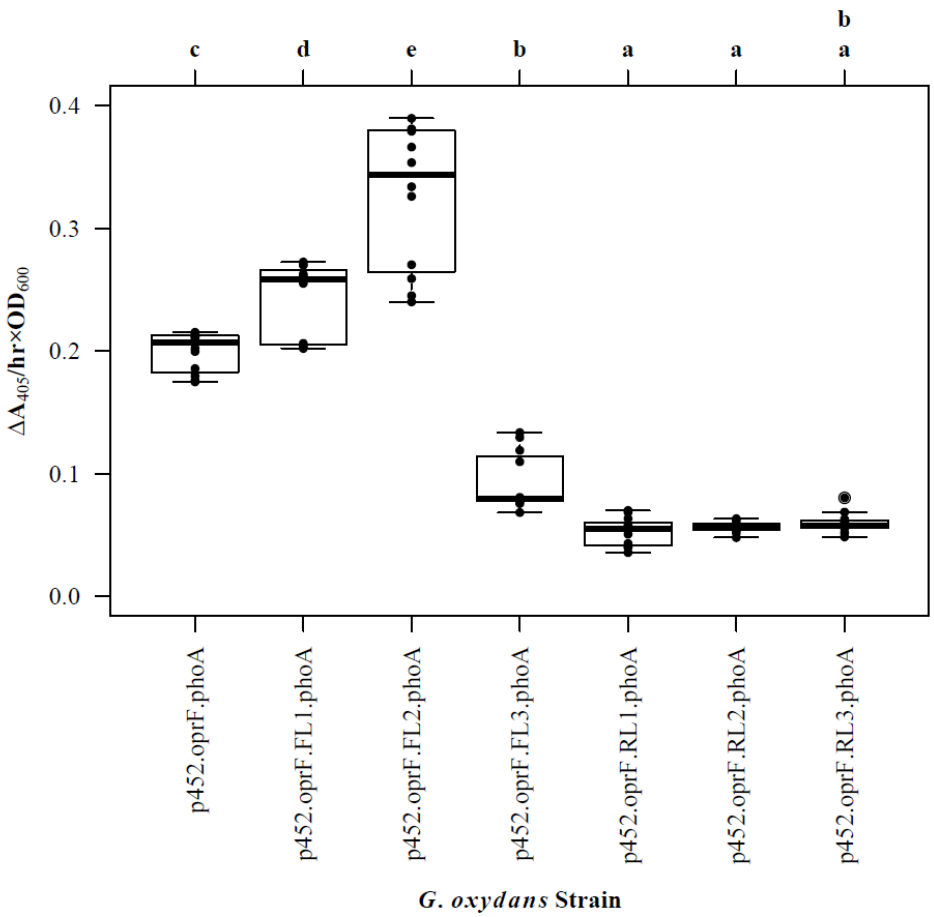


Figure 17. Moderate Expression of the Linker Library in *G. oxydans*. A library of six OprF188-PhoA fusion proteins with varying linkers was produced via moderate-expression plasmids in *G. oxydans*, and activity of the reporter enzyme was quantified in whole-cell reactions. N=12 for each strain, and data points for each replicate are shown by the superimposed strip chart. The letters above the plot denote statistical groups, as determined by an ANOVA and *post-hoc* Tukey's HSD test.

among the strains containing high-expression constructs. However, when PhoA was expressed as a fusion with OprF188 in lieu of the Strep-tag, the observed doubling time was not significantly different than that of non-recombinant cells ($q=0.603$). The presence of the small flexible linker (FL1) increased the doubling time (82 min) compared to the strain expressing the OprF188-PhoA fusion without a linker ($q<0.001$). Conversely, the presence of the small rigid linker (RL1) not only decreased doubling time (27 min) compared to no-linker control strain ($q<0.001$), but its growth rate was higher

Table 10. Statistical Analysis: Low Expression of the Linker Library in *G. oxydans*

	p452-oprF-phoA	p452-oprF-FL1-phoA	p452-oprF-FL2-phoA	p452-oprF-FL3-phoA	p452-oprF-RL1-phoA	p452-oprF-RL2-phoA
p452-oprF-FL1-phoA	0.003	-	-	-	-	-
p452-oprF-FL2-phoA	<0.001	<0.001	-	-	-	-
p452-oprF-FL3-phoA	<0.001	<0.001	<0.001	-	-	-
p452-oprF-RL1-phoA	<0.001	<0.001	<0.001	0.012	-	-
p452-oprF-RL2-phoA	<0.001	<0.001	<0.001	0.031	1.000	-
p452-oprF-RL3-phoA	<0.001	<0.001	<0.001	0.067	0.996	1.000

q values from an ANOVA and *post-hoc* Tukey's HSD test

than that of the wild-type strain (q<0.001). However, the lag time for this strain was 5–7 hours longer than that of wild-type cells (Figure 18). The medium rigid linker (RL2) did not affect growth compared to the no-linker control strain (q=0.166) or wild-type cells (q=0.952). Lastly, the large rigid linker (RL3) led to the highest doubling time among the strains containing high-expression constructs, at 121 minutes. Generally, the lag time was greater for recombinant strains and the final optical density observed for recombinant strains was lower than that of wild-type *G. oxydans*, except for the p264-oprF-RL1-phoA containing strain (Figure 18).

At moderate-expression, production of the OprF188-ST fusion also led to a significant increase in mean doubling time (116 min) compared to wild-type cells (Figure 21) (q<0.001). This time, replacement of the Strep-tag with PhoA did not improve the

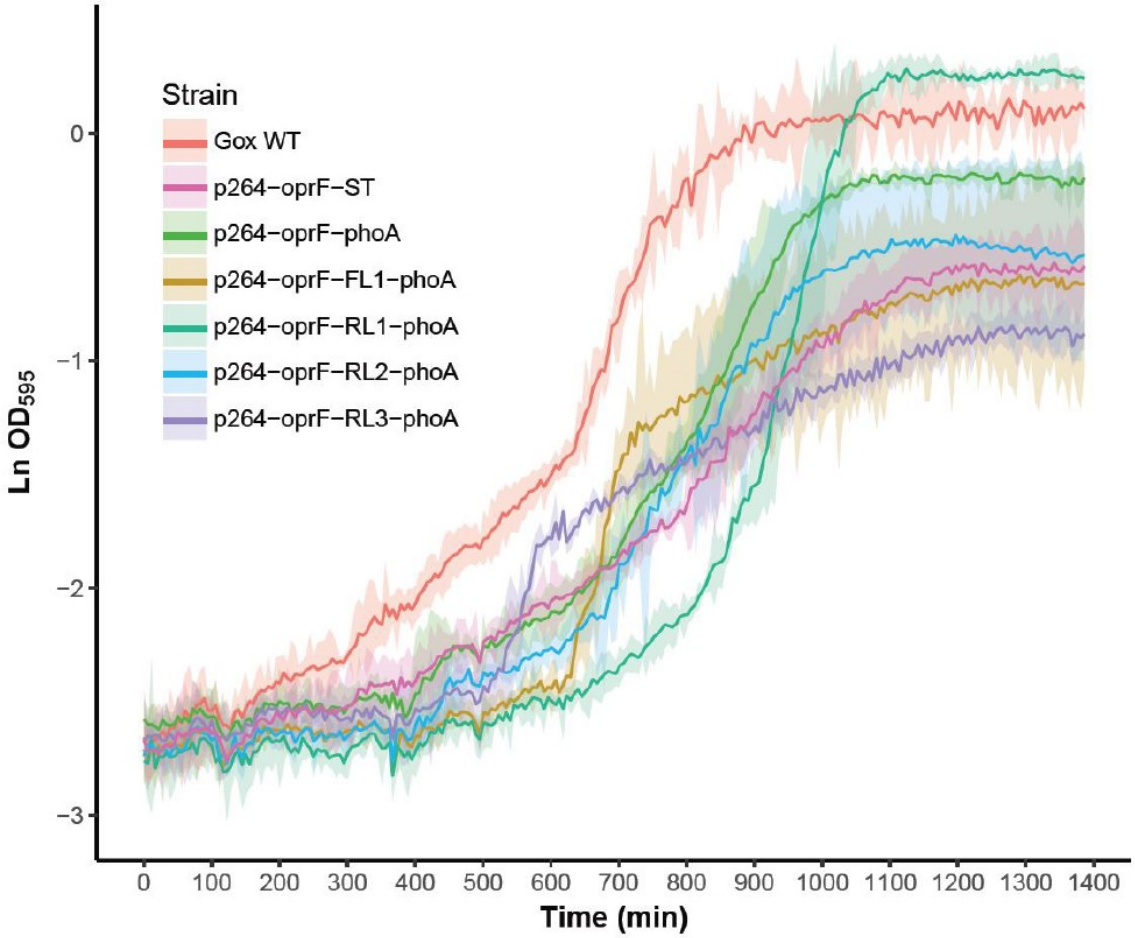


Figure 18. Growth Behavior of *G. oxydans* Strains Containing High-Expression Surface Display Plasmids. *G. oxydans* strains containing high-expression surface display constructs were grown for approximately 23 hours and growth was followed by measuring absorbance at 595 nm. Wild-type *G. oxydans* 621H (*Gox* WT) was included for comparison. The growth curves in this plot represent the mean optical density (natural-log-transformed) for three biological replicates of each *G. oxydans* strain (N=3). The ribbons surrounding each growth curve indicate the 95% confidence interval of the mean.

growth rate—the doubling time for the strain containing plasmid p452-oprF-phoA was no different than that of the strain containing p452-oprF-ST (q=0.075). The presence of the small (FL1) and medium (FL2) flexible linkers led to decreased doubling times (114 min and 106 min, respectively) compared to the no-linker control strain (130 min) (q=0.033 and 0.001, respectively), while the large flexible linker (FL3) had no effect (q=0.999).

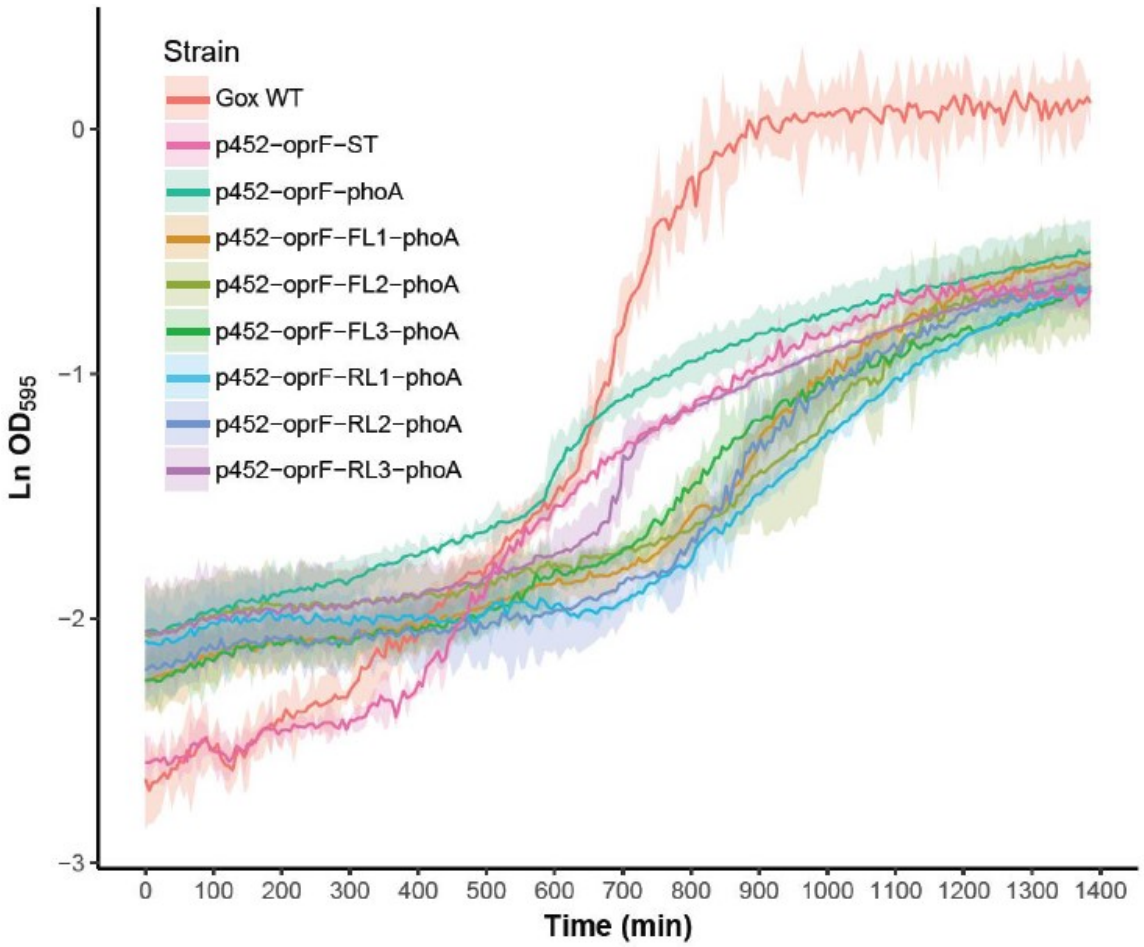


Figure 19. Growth Behavior of *G. oxydans* Strains Containing Moderate-Expression Surface Display Plasmids. *G. oxydans* strains containing moderate-expression surface display constructs were grown for approximately 23 hours and growth was followed by measuring absorbance at 595 nm. Wild-type *G. oxydans* 621H (*Gox* WT) was included for comparison. The growth curves in this plot represent the mean optical density (natural-log-transformed) for three biological replicates of each *G. oxydans* strain (N=3). The ribbons surrounding each growth curve indicate the 95% confidence interval of the mean.

Similarly, the small (RL1) and medium (RL2) rigid linkers decreased doubling time (106 min and 95 min, respectively) compared to the no-linker control strain ($q=0.001$ and $q<0.001$, respectively), while the large rigid linker (RL3) had no effect ($q=0.806$). Generally, lag times were longer, growth rates were lower, and final optical densities were lower for recombinant strains compared to the wild-type strain (Figures 19 and 21).

Surprisingly, the moderate-expression plasmids led to greater negative effects on growth rate and final optical density compared to their high-expression counterparts.

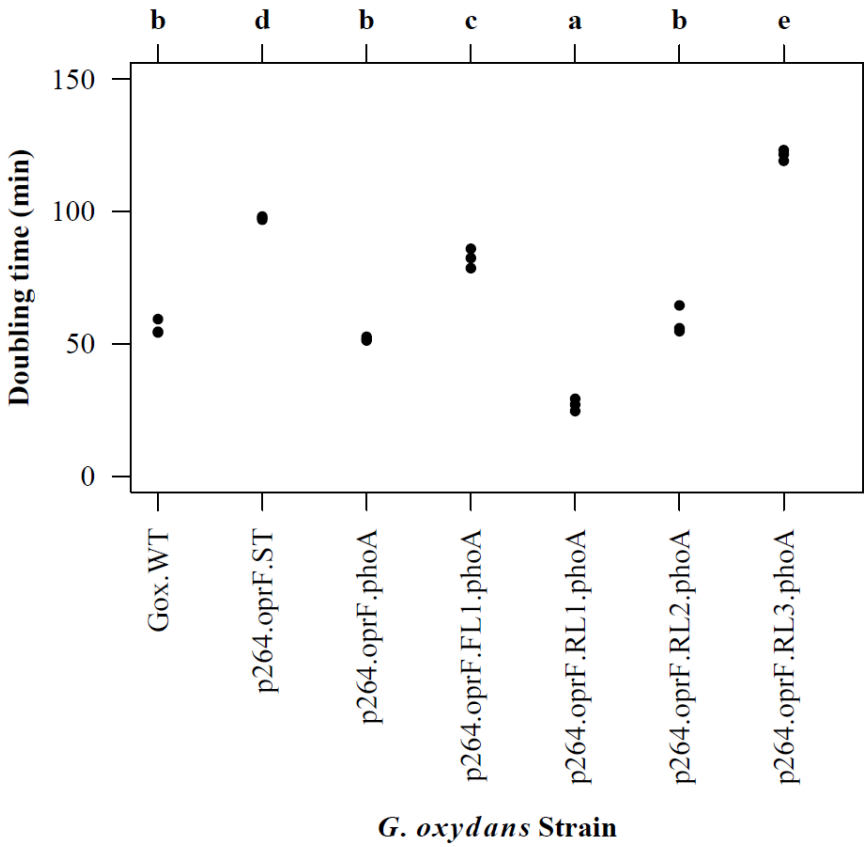


Figure 20. Growth Rates of *G. oxydans* Strains Containing High-Expression Surface Display Constructs. *G. oxydans* strains containing high-expression surface display constructs were grown for approximately 23 hours and growth was followed by measuring absorbance at 595 nm. Wild-type *G. oxydans* 621H (*Gox* WT) was included for comparison. From this data, the doubling time for each biological replicate was calculated (N=3). The letters above the plot denote statistical groups, as determined by an ANOVA and *post-hoc* Tukey's HSD test (q=0.05).

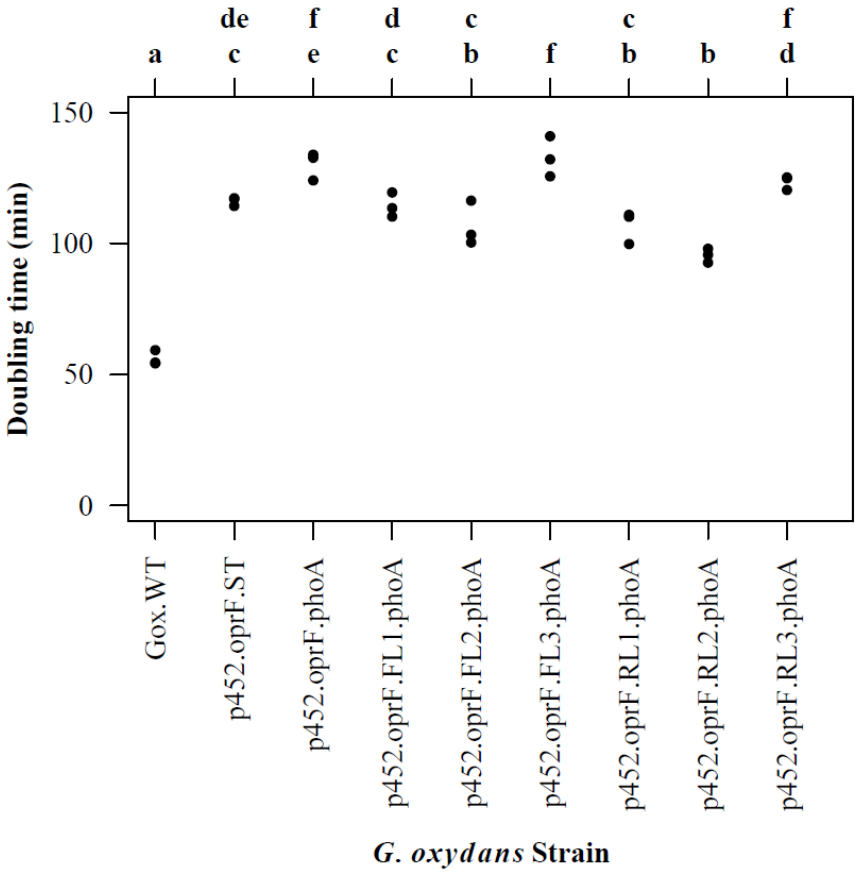


Figure 21. Growth Rates of *G. oxydans* Strains Containing Moderate-Expression Surface Display Constructs. *G. oxydans* strains containing moderate-expression surface display constructs were grown for approximately 23 hours and growth was followed by measuring absorbance at 595 nm. Wild-type *G. oxydans* 621H (*Gox* WT) was included for comparison. From this data, the doubling time for each biological replicate was calculated (N=3). The letters above the plot denote statistical groups, as determined by an ANOVA and *post-hoc* Tukey's HSD test (q=0.05).

DISCUSSION

Proof of Concept: Surface Display in Acetic Acid Bacteria

In this study, the gene encoding PhoA was fused to corresponding genes for three anchor proteins, and the resulting anchor library was expressed in *E. coli*. The only *E. coli* strains that produced appreciable enzymatic activity were those containing the OprF188 surface display systems. To test the ability of OprF188 to transport recombinant enzymes to the cell surface of AAB, these constructs were then expressed in the model acetic acid bacterium, *G. oxydans*, and biocatalysis was quantified. Based on enzymatic activity, this anchor protein consistently localized active PhoA to the cell envelope of this bacterium, regardless of expression level. Nascent PhoA has no activity in the cytoplasm, as it does not fold properly unless secreted to the periplasm, where disulfide-forming enzymes form its tertiary structure (De Geyter et al. 2016; Ehrmann et al. 1990; Hoffman and Wright 1985; Manoil and Beckwith 1985; Michaelis et al. 1983). Furthermore, the OprF188 surface display system developed in this study produced significantly higher yields than the preliminary system developed by Pearson (2014). In that study, a Strep-tag was fused to the C-terminus of the OprF188-PhoA fusion protein so it could be detected by immunoblotting against the tag. The presence of a C-terminal Strep-tag apparently interfered with the activity of PhoA, either by preventing dimer formation or by occluding the active site of the enzyme.

The OprF188 surface display system was expressed via two vectors designed for protein production in *G. oxydans*, one containing a high-strength promotor, and the other a moderate-strength promotor. There was a two-fold difference in phosphatase activity

produced by the respective *G. oxydans* strains. Others have observed up to a three-fold difference in enzymatic activity when using these expression vectors in *G. oxydans* (Kallnik et al. 2010). To determine the effects of protein production on the growth of *G. oxydans*, a novel method was developed to generate growth curves for each *G. oxydans* strain. The strain containing plasmid p264-oprF-ST (used as a negative-control for phosphatase activity) had significantly slower growth compared to wild-type *G. oxydans* 621H. It is possible that this growth defect was caused by the metabolic burden of protein production, especially since *G. oxydans* has a limited ability to generate ATP. Alternatively, the recombinant protein may have overwhelmed secretory machinery, preventing necessary proteins from being processed. Interestingly, the p264-oprF-phoA strain did not exhibit a growth defect, suggesting that PhoA may have enabled recovery from the burden of protein production. It is possible that high expression of PhoA allowed cells to procure more phosphate from the medium. Surprisingly, plasmid p452-oprF-ST, containing the moderate-strength promotor, caused an even greater growth deficit. Normal growth was not recovered when PhoA was expressed in lieu of the Strep-tag, suggesting that moderate expression of PhoA was not enough to scavenge phosphate from the medium.

PhoA is an innately periplasmic enzyme. Thus, its location needed to be verified to provide evidence that OprF188 correctly targeted recombinant enzymes to the outer leaflet of the outer membrane in *G. oxydans*. To verify the location of PhoA, a specialized version of the OprF188-PhoA fusion protein was developed that contained the amino acid sequence, Ile-Asp-Gly-Arg, within the region that links the two proteins together. This sequence is recognized by Factor Xa protease, which cleaves the

polypeptide backbone at the carboxyl end of the arginine residue (Nagai and Thøgerson 1984; Terpe 2003). Factor Xa cleavage sites have been used to validate other surface display systems. For example, Jiang and Boder (2010) used this method to verify localization of recombinant proteins at the cell membrane of yeast. The cleavable linker construct was expressed in *G. oxydans* and cells were subsequently treated with Factor Xa using a method like that described by Jiang and Boder (2010). If the cleavable linker motif and PhoA are exposed at the cell surface via OprF188, then PhoA should be removed from the outer membrane following treatment with the protease. Correspondingly, phosphatase activity will shift from the cells to the surrounding medium. Therefore, phosphatase activity in the supernatant from treated cells should be higher than that in the supernatant from untreated cells, and/or the phosphatase activity of treated cells should be lower than that of untreated cells.

The cleavable linker assay demonstrated that the level of phosphatase activity in the treated supernatant fractions was significantly higher than that in the untreated supernatant fractions, which suggests that PhoA was present at the outer leaflet of the outer membrane, where it was accessible to Factor Xa. That the untreated supernatant fractions contained active PhoA suggests some of the enzyme spontaneously detached from the cell surface, which could only happen if PhoA was properly localizing. Overall, the results of the phosphatase assays and the cleavable linker assay suggest that OprF188 is a suitable anchor protein for surface display in *G. oxydans*.

OprF188 for Surface Display

Outer membrane porin F, or OprF, is an outer membrane protein found naturally in *Pseudomonas aeruginosa* (Sugawara et al. 2012; Schüürmann et al. 2014). Besides acting as a nonspecific, low-permeability porin, it also plays an important structural role in linking the outer membrane to the cell wall in this bacterium (Lee et al. 2005a; Sugawara et al. 2012). The prototypical OMP consists of a single-domain transmembrane β-barrel with exposed loops at either leaflet of the outer membrane (Knowles et al. 2009; Lee et al. 2003; Noinaj et al. 2017; Schüürmann et al. 2014). Therefore, OprF is unusual among OMPs, as it predominately occurs as a multidomain protein. OprF is considered a homologue of OmpA in *E. coli*, and modeling studies of these proteins revealed that they typically are comprised of two domains: an N-terminal domain that forms a small, eight-stranded β-barrel, and a C-terminal domain that associates with the peptidoglycan cell wall in the periplasm (Bodilis and Barray 2006; Brinkman et al. 2000; Sugawara et al. 2012).

In this study, a C-terminal truncated version of OprF was used, OprF188, which contains 188 amino acids of the N-terminus of the original protein. The first 24 amino acids of nascent OprF188 constitute a transmembrane signal peptide (Peterson et al. 2011) which permits secretion across the inner membrane via the Sec translocon. Mature OprF188 consists of eight transmembrane β-strands and ends with a loop structure on the extracellular side of the outer membrane (Figure 22) (Lee et al. 2005a). Thus, OprF188 is essentially the N-terminal β-barrel domain of the untruncated protein (Lee et al. 2005a; Sugawara et al. 2012). OprF and OmpA are unique in that their β-signals are located internally, after their β-barrel domains (Gessmann et al. 2014; Robert et al. 2006;

Sugawara et al. 2012). Thus, truncated OprF188 maintains its β-signal (Lee et al. 2005a; Sugawara et al. 2012). Valine 188 was empirically determined to be the optimal fusion site for passenger proteins (Lee et al. 2005a). In two related surface display studies, this anchor protein was used to localize a 49.9 kDa lipase, which is comparable to the size of the 47 kDa PhoA monomer (Bradshaw et al. 1981), to the outer membrane of *E. coli* (Lee et al. 2005a) and *Pseudomonas putida* (Lee et al. 2005b). Both systems were used to generate chiral products.

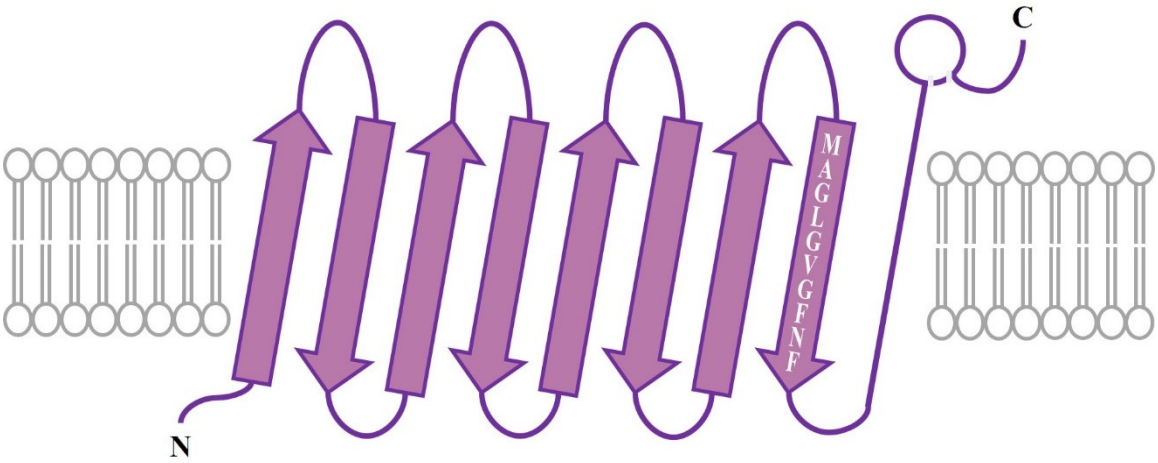


Figure 22. Anchor Protein OprF188. A C-terminal truncated version of outer membrane porin F from *Pseudomonas aeruginosa*, OprF188, was used for surface display of a reporter enzyme in *G. oxydans*. Nascent OprF188 contains a 24-amino acid N-terminal transmembrane signal peptide that directs its excretion across the cell membrane via the SecYEG translocon. This signal is cleaved in the periplasm, meaning that mature OprF188 is 164 residues in length. OprF188 forms a transmembrane β-barrel comprised of eight antiparallel β-strands in the outer membrane. The N-terminus is exposed to the periplasm, while the C-terminus forms a loop structure on the extracellular side of the outer membrane. The final β-strand contains the amino acid sequence, Ala-Gly-Leu-Gly-Val-Gly-Phe-Asn-Phe, which constitutes a β-signal. This signal is recognized by the β-barrel assembly machinery complex, which is responsible for insertion of outer membrane proteins into the outer membrane in gram-negative bacteria. Sources: Lee et al. 2005a; Noinaj et al. 2017; Peterson et al. 2011; Robert et al. 2006. Redrawn from Lee et al. 2005a, with modification.

Interestingly, there are three studies that tried and failed to target PhoA to the cell surface of *E. coli* using *E. coli* OMPs as anchors. When Coulton et al. (1988) fused PhoA to ferrichrome outer membrane transporter, FhuA, they found that PhoA was associated with the outer membrane but not necessarily exposed to the outer surface. Similarly, Murphy and Klebba (1989) fused PhoA to ferrienterobactin outer membrane transporter, FepA, and concluded that PhoA was associated with the outer membrane but only present in the periplasm. Finally, Stathopoulos et al. (1996) fused PhoA to a lipoprotein-OmpA hybrid and, again, the fusion protein was associated with the outer membrane, but PhoA existed exclusively in the periplasm. Close inspection of these studies reveals that the β-signal was unknowingly removed from FhuA, FepA, and OmpA, as C-terminal truncated versions of these proteins were used. Therefore, deletion of the β-signal was likely sufficient to prevent outer membrane insertion by the BAM complex in those studies. Generally, a better understanding of bacterial outer membrane export would enable advances in surface display technology. The use of OprF188 as an anchor in this study and others suggests that it is broadly functional as a surface display anchor protein for gram-negative systems.

INPNC and PgsA

The anchor library also included INPNC and PgsA as potential anchor proteins. Because these proteins produced poor results in *E. coli*, only OprF188 was used for surface display in *G. oxydans*. INPNC is a truncated version of the ice nucleation protein (INP) from *Xanthomonas campestris.* Plant pathogens belonging to the genera *Erwinia*, *Pseudomonas*, and *Xanthomonas* produce INP. In these bacteria, INP functions to

damage plants by forming ice crystals from supercooled water (Lee et al. 2003; Wu et al. 2006). INP and its derivatives have been used for surface display in many studies, and INPNC is capable of localizing passenger proteins over 100 kDa to the cell surface of gram-negative hosts (Lee et al. 2003; Schüürmann et al. 2014; Saleem et al. 2008; van Bloois et al. 2011). INPNC is an INP variant consisting of only the C- and N-terminal domains of the original three-domain protein—the middle domain has been removed (Wu et al. 2006). A few studies have used INPNC for surface display in *E. coli*. For example, Liu et al. (2013) and Yang et al. (2010) used INPNC from *Pseudomonas syringae* to develop systems to degrade toxic organophosphate pesticides. Furthermore, INPNC from *X. campestris* was used to express a 60 kDa transglucosidase at the cell surface of *E. coli* for whole-cell biocatalysis (Wu et al. 2006).

One limitation to using INP for surface display is that it requires a glycosylphosphatidylinositol-like phospholipid anchor for attachment to the outer membrane (Li et. al. 2012; Schüürmann et al. 2014; Wu et al. 2006). Recombinant expression of INP has been reported to increase the level of phosphatidylinositol in *E. coli* (Kozloff et al. 1991). However, no phosphatase activity was observed in the current study when INPNC was used as an anchor protein, suggesting an alternate reason for the lack of expression of active PhoA. Little is known about the mechanism of export for INP (Li et al. 2012), but it is possible that this undescribed mechanism is not compatible with PhoA. Alternatively, the INPNC-PhoA fusion protein may have failed to fold into a stable confirmation.

PgsA is a component of the poly-γ-glutamate synthetase complex, PgsBCA, in *Bacillus subtilis* (Ashiuchi et al. 2001; Narita et al. 2006; Schüürmann et al. 2014).

PgsBCA consists of three proteins and part of the function of PgsA is to anchor the complex to the cell membrane of this gram-positive bacterium (Ashiuchi et al. 2001). Thus, PgsA is usually used for surface display in gram-positive hosts (Schüürmann et al. 2014), but a few studies have demonstrated its functionality in gram-negative systems. For example, Narita et al. (2006) used PgsA to express a 77 kDa amylase and a 34 kDa lipase in *E. coli*. Additionally, Ryu and Karim (2011) displayed three cellulases in a single *E. coli* strain and used the recombinant microbe to produce ethanol from cellulosic biomass. Little is known about how PgsA is naturally exported in *B. subtilis*, and nothing is known about how it is exported in recombinant gram-negative hosts (Schüürmann et al. 2014). Indeed, PgsA contains no discernable signal peptide for either type of bacteria (Peterson et al. 2011). In this study, a low level of phosphatase activity was observed in *E. coli* cells producing the PgsA-PhoA fusion protein via the moderate-expression promotor. However, no activity was observed when the protein was produced at a high level of expression. It is possible that overproduction of the protein via a high-strength promotor led to protein aggregation and inclusion body formation. Taken together, these two results suggest that PgsA was poorly secreted.

The Effects of Fusion Linkers on Biocatalysis via Surface Display

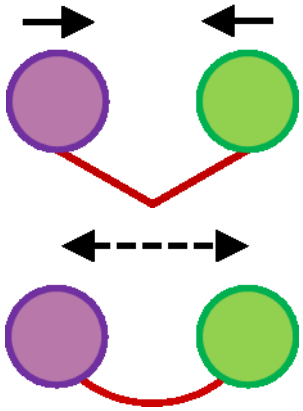
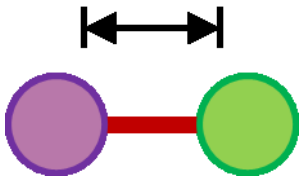
In biotechnology, protein fusion is increasingly being used to improve and expand the capabilities and applications of recombinant proteins. Protein fusion can involve the addition of a small peptide, a component from another protein (moiety), or even a complete second protein to the protein-of-interest (Yang et al. 2016). Synthetic fusion proteins consisting of two separate proteins are essentially no different than naturally-

occurring multi-domain proteins, and both natural and synthetic proteins often require linker sequences to allow proper folding, stability, and function of their discrete domains (Chen et al. 2013). Thus, linker sequences can play a vital role in the design of fusion proteins, as they can affect the properties of the entire biomolecule. Linker sequences are typically composed of repeated and/or patterned motifs of amino acids, and vary in both amino acid composition and motif number which, in turn, impact secondary structure and length of the linker, respectively (Chen et al. 2013). Two types of linkers were used in this study: flexible linkers and rigid linkers (Table 11). Flexible linkers contain small, polar amino acids such as glycine (G) and serine (S). A common flexible linker motif is $(GGGGS)_n$. A common rigid linker sequence is composed of the amino acids, glutamic acid (E), alanine (A), and lysine (K), arranged in a $(EAAAK)_n$ motif (Chen et al. 2013; Li et al. 2016).

Rigid linkers fold to form stable α-helices with internal hydrogen bonding, while flexible linkers do not form ordered conformations and instead occur as random coils, lacking secondary structure (Chen et al. 2013; Li et al. 2016). The $(GGGGS)_n$ flexible linker can be used when interaction between the two components of a fusion protein is desired. This design can provide passive, variable separation of proteins as well. Flexible linkers have been used both to improve stability of fusion proteins and to increase activity (Chen et al. 2013). Linkers containing more rigid units span a greater distance end-to-end and there is less variation in that distance (Li et al. 2016). Per length, rigid linkers provide better separation of the fused components and maintain independence thereof, thereby preventing unwanted interference between the components of the fusion protein. In

general, linker sequences can improve folding of fusion proteins which, in turn, affects expression, and they can increase the activity of enzymes (Chen et al. 2013).

Table 11. Fusion Linkers

Linker type		Advantages	Characteristics	Examples
Flexible		Facilitates interaction between fusion partners **or** Provides passive spatial separation between domains	No secondary structure, composed of small and/or hydrophilic residues	$(GGGGS)_n$ $(G)_n$
Rigid		Maintains distance between domains	α-helical secondary structure or contains many prolines	$(EAAAK)_n$ $(XP)_n$

Recreated from Chen et al. 2013, with modification.

Recombinant surface display systems inherently involve translational fusion of anchor proteins to passenger proteins. What is more, autotransporters naturally include linkers between their anchor domains and passenger domains (Jose 2006; Jose and Meyer 2007). Nonetheless, the effects of linkers on a bacterial surface display system have not previously been investigated. Linkers are sometimes included in the design of surface display systems, but this is rarely made explicit and even fewer studies offer any explanation as to why a particular linker sequence was chosen. In this study, a library of linkers, varying in both composition and length, was generated. Therefore, this is the first

study to test the effects of linkers on biocatalysis at the cell surface of not one, but two bacterial species. The linker library was first expressed in *E. coli* and, generally, the presence of linkers did not dramatically influence product yield. When fusion proteins containing linkers were produced at high expression, flexible linkers decreased product yield slightly, and rigid linkers had no effect on yield (Figure 14). At moderate expression, the presence of linkers increased activity regardless of type or length, except when the short rigid linker (RL1) was used (Figure 15).

In *G. oxydans*, linkers caused more pronounced effects on biocatalysis relative those observed in *E. coli*. Fusion proteins containing the medium (FL2) and large flexible linkers (FL3) —10 and 15 amino acids in length, respectively—were toxic to *G. oxydans* at high expression, as positive transformants were not obtained after multiple attempts. The fusion protein containing the small flexible linker (FL1) was stably produced at high expression and had no effect on biocatalysis, but it did negatively affect the growth rate of the bacterium (Figures 16, 18, and 20). However, when produced via the moderate-strength expression vector, the small and medium flexible linkers (FL1 and FL2) significantly increased product yield (Figure 17). In fact, the $(GGGGS)_2$ flexible linker led to a level of enzymatic activity that was equivalent to some of the high-expression surface display systems. The main purpose in developing the linker library was to optimize biocatalysis at the cell surface of *G. oxydans*. This was achieved through the addition of a small rigid linker (RL1) consisting of a single EAAAK pentapeptide to the OprF188 surface display system. This linker improved phosphatase activity by nearly 70%. Additionally, this linker improved the growth rate (but increased lag time) of the respective *G. oxydans* strain, possibly because the enzyme provided a nutritional

advantage by cleaving inorganic phosphate from substrates within the medium. Interestingly, the opposite effect was observed when the fusion protein containing the small rigid linker was produced via the moderate-strength promotor, decreasing enzymatic activity by about 73%.

While fusion protein linkers appeared to have minor effects on biocatalysis in *E. coli*, they caused more dramatic changes in *G. oxydans*. Therefore, it seems that the effects of linkers on surface display may be species-specific. *E. coli* and *G. oxydans* are phylogenetically distant, belonging to the *Gamma-* and *Alphaproteobacteria*, respectively, and demonstrate very different lifestyles in nature. Thus, there are likely many differences in the composition of their outer membranes and thus the environment at the cell surface of these microorganisms. In *G. oxydans*, few consistent trends appear in the data gathered for the linker library in this microorganism, and the effects of linkers on surface display seemed to be expression-dependent. It is possible that linkers could have affected stability and folding of the fusion proteins, the formation of active PhoA homodimers, or even the catalytic rate of the enzyme. Additionally, some of the observations made in this study are likely specific to alkaline phosphatase, while other reporters may give different results. Overall, these results suggest that linker optimization is an important consideration for each surface display system used and for each host microorganism used.

Surface Display Towards Metabolic Engineering of Acetic Acid Bacteria

Surface display is potentially a powerful tool to enable metabolic engineering of *G. oxydans* and other important AAB. The ability to express enzymes at the outer

membrane could be used to improve current bioprocesses by broadening the substrate range of this bacterium. For example, coexpression of an α-amylase and a glucoamylase would enable the microorganism to hydrolyze starch to procure glucose. Such dual-expression systems have been achieved in yeast surface display studies (Shigetchi et al. 2004). The resulting glucose would be usable by cells for production of gluconate and gluconate derivatives. Furthermore, starch is considered a renewable and economical feedstock compared to using glucose as a carbon source. Similarly, this system could be used for surface expression of lipase enzymes in *G. oxydans*. Additionally, surface display could also be used to immobilize *G. oxydans* cells to create stable bioreactors. For example, immobilized *G. oxydans* cells have improved production of dihydroxyacetone from both pure and crude glycerol (Dikshit and Moholkar 2016).

There are a few ways in which this new surface display system could be improved. First, inducible promotors would permit more control over this system. Second, the sequence for OprF188 could be replaced with one encoding an OMP from *G. oxydans*. It is conceivable that this would lead to greater export efficiency. As for the linker system, PhoA could be replaced with another reporter enzyme to determine whether the results in this study were specific to PhoA or applicable to any surface-displayed enzyme in *G. oxydans*. Recently, new techniques have been proposed to assess surface displays systems. For example, Wendel et al. (2016) developed a way to quantify surface display efficiency using an anti-GFP nanobody. This nanobody can be inserted between the passenger and anchor or to the end of the protein-of-interest. Next, purified GFP is added to medium and fluorescence can be measured. The quantity of GFP that

binds to cells is a relative measure of the amount of protein that is expressed at the cell surface (Wendel et al. 2016).

In conclusion, this study produced a novel molecular tool for strain improvement of AAB, and the OprF188 surface display system described herein is a significant first step towards outer membrane engineering of *G. oxydans*. Such molecular tools will enable metabolic engineering of this unique bacterium to improve and expand its ability to produce value-added products, especially rare sugars and sugar derivatives that serve as important precursors to pharmaceuticals, industrial chemicals, and consumer products.

REFERENCES

Alcade M, Ferrer M, Plou FJ, Ballesteros A (2006) Environmental biocatalysis: from remediation with enzymes to novel green processes. Trends Biotechnol 24(6):281-287

Ashiuchi M, Nawa C, Kamei T, Song JJ, Hong SP, Sung MH, Soda K, Yagi T, Misono H (2001) Physiological and biochemical characteristics of poly γ-glutamate synthetase complex of *Bacillus subtilis*. Eur J Biochem 268:5321-5328

Bradshaw RA, Cancedda F, Ericsson LH, Neumann PA, Piccoli SP, Schlesinger MJ, Shriefer K, Walsh KA (1981) Amino acid sequence of *Escherichia coli* alkaline phosphatase. Proc Natl Acad Sci USA 78(6):3473-3477

Bodilis J, Barray S (2006) Molecular evolution of the major outer-membrane protein gene (*oprF*) of *Pseudomonas*. Microbiology 152(4):1075-1088

Brinkman FSL, Bains M, Hancock REW (2000) The amino terminus of *Pseudomonas aeruginosa* outer membrane protein OprF forms channels in lipid bilayer membranes: correlation with a three-dimensional model. J Bacteriol 182(18):5251-5255

Chen X, Zaro JL, Shen WC (2013) Fusion protein linkers: property, design, and functionality. Adv Drug Deliv Rev 65(10):1357-1369

Coulton JW, Reid GK, Campana A (1988) Export of hybrid proteins FhuA'-'LacZ and FhuA'-'PhoA to the cell envelope of *Escherichia coli* K-12. J Bacteriol 170(5):2267-2275

Crotti E, Rizzi A, Chouaia B, Ricci I, Favia G, Alma A, Sacchi L, Bourtzis K, Mandrioli M, Cherif A, Bandi C, Daffonchio D (2010) Acetic acid bacteria, newly emerging symbionts of insects. Appl Environ Microbiol 76(21):6963-6970

Czaja W, Krystynowicz A, Bielecki S, Brown RM (2006) Microbial cellulose—the natural power to heal wounds. Biomaterials 27:145-151

De Geyter J, Tsirigotaki A, Orfanoudaki G, Zorzini V, Economou A, Karamanou S (2016) Protein folding in the cell envelope of *Escherichia coli*. Nat Microbiol 1(8):16107

De Muynck C, Pereira CSS, Naessens M, Parmentier S, Soetaert W, Vandamme EJ (2007) The genus *Gluconobacter oxydans*: comprehensive overview of biochemistry and biotechnological applications. Crit Rev Biotechnol 27:141-171

De Ley J, Gillis M, Swings J (1984) The genus *Gluconobacter*. In: Krieg NR, Holt JG (eds) Bergey's manual of systematic bacteriology, vol 1. Williams and Wilkins, Baltimore, pp 267-278

Deppenmeier U, Hoffmeister M, Prust C (2002) Biochemistry and biotechnological applications of *Gluconobacter* strains. Appl Microbiol Biotechnol 60(3):233-242

Deppenmeier U, Ehrenreich A (2009) Physiology of acetic acid bacteria in light of the genome sequence of *Gluconobacter oxydans*. J Mol Microbiol Biotechnol 16(1-2):69-80

Dikshit PK, Moholkar VS (2016) Kinetic analysis of dihydroxyacetone production from crude glycerol by immobilized cells of *Gluconobacter oxydans* MTCC 904. Bioresour Technol 216:948-957

Ehrmann M, Boyd D, Beckwith J (1990) Genetic analysis of membrane protein topology by a sandwich gene fusion approach. Proc Natl Acad Sci USA 87:7574-7578

Ferrer M, Beloqui A, Timmis KN, Golyshin PN (2009) Metagenomics for mining new genetic resources of microbial communities. J Mol Microbiol Biotechnol 16(1-2):109-123

Gessmann D, Chung YH, Danoff EJ, Plummer AM, Sandlin CW, Zaccai NR, Fleming KG (2014) Outer membrane β-barrel protein folding is physically controlled by periplasmic lipid head groups and BamA. Proc Natl Acad Sci USA 111(16):5878-5883

Gu Y, Li H, Dong H, Zeng Y, Zhang Z, Paterson NG, Stansfeld PJ, Wang Z, Zhang Y, Wang W, Dong C (2016) Structural basis of outer membrane protein insertion by the BAM complex. Nature 531(7592):64-69

Gupta A, Singh VK, Qazi GN, Kumar A (2001) *Gluconobacter oxydans*: its biotechnological applications. J Mol Microbiol Biotechnol 3(3):445-456

Gustavsson M, Do TH, Lüthje P, Tran NT, Brauner A, Samuelson P, Truong NH, Larsson G (2015) Improved cell surface display of *Salmonella enterica* serovar Enteritidis antigens in *Escherichia coli*. Microb Cell Fact 14:47

Han L, Zheng J, Wang Y, Yang X, Liu Y, Sun C, Cao B, Zhou H, Ni D, Lou J, Zhao Y, Huang Y (2016) Structure of BAM complex and its implications for biogenesis of outer-membrane proteins. Nat Struct Mol Biol 23(3):192-196

Hoffman CS, Wright A (1985) Fusions of secreted proteins to alkaline phosphatase: an approach for studying protein secretion. Proc Natl Acad Sci USA 82:5107-5111

Jiang W, Boder ET (2010) High-throughput engineering and analysis of peptide binding to class II MHC. Proc Natl Acad Sci USA 107(30)13258-13263

Jose J (2006) Autodisplay: efficient bacterial surface display of recombinant proteins. Appl Microbiol Biotechnol 69(6):607-614

Jose J, Meyer TF (2007) The autodisplay story, from discovery to biotechnological and biomedical applications. Microbiol Mol Biol Rev 71(4):600-619

Kallnik V, Meyer M, Deppenmeier U, Schweiger P (2010) Construction of expression vectors for protein production in *Gluconobacter oxydans*. J Biotechnol 150(4):460-465

Keshk SM (2014) Bacterial cellulose production and its industrial applications. J Bioprocess Biotech 4(2):150

Knowles TJ, Scott-Tucker A, Overduin M, Henderson IR (2009) Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. Nature Rev Microbiol 7(3):206-214

Kosciow K, Zahid N, Schweiger P, Deppenmeier U (2014) Production of a periplasmic trehalase in *Gluconobacter oxydans* and growth on trehalose. J Biotechnol 189:27-35

Kosciow K, Domin C, Schweiger P, Deppenmeier U (2016) Extracellular targeting of an active endoxylanase by a TolB negative mutant of Gluconobacter oxydans. J Ind Microbiol Biotechnol 43(7):989-999

Kovach ME, Elzer PH, Hill DS, Robertson GT, Farris MA, Roop RM, Peterson KM (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene 166(1):175-176

Kozloff LM, Turner MA, Arellano F, Lute M (1991) Phosphatidylinositol, a phospholipid of ice-nucleating bacteria. J Bacteriol 173(6):2053-2060

Lee SH, Choi J, Han MJ, Choi JH, Lee SY (2005a) Display of lipase on the cell surface of *Escherichia coli* using OprF as an anchor and its application to enantioselective resolution in organic solvent. Biotechnol Bioeng 90:223-230

Lee SH, Lee SY, Park BC (2005b) Cell surface display of lipase in *Pseudomonas putida* KT2442 using Oprf as an anchoring motif and its biocatalytic applications. Appl Environ Microbiol 71(12):8581-8586

Lee YL, Choi JH, Xu Z (2003) Microbial cell-surface display. Trends Biotechnol 21(1):45-52

Li G, Huang Z, Zhang C, Dong BJ, Guo RH, Yue HW, Yan LT, Xing XH (2016) Construction of a linker library with widely controllable flexibility for fusion protein design. Appl Microbiol Biotechnol 100(1):215-225

Li Q, Yan Q, Chen J, He Y, Wang J, Zhang H, Yu Z, Li L (2012) Molecular characterization of an ice nucleation protein variant (InaQ) from *Pseudomonas syringae* and the analysis of its transmembrane transport activity in *Escherichia coli*. Int J Biol Sci 8(8):1097-1108

Liu R, Yang C, Xu Y, Xu P, Jiang H, Qiao C (2013) Development of a whole-cell biocatalyst/biosensor by display of multiple heterologous proteins on the *Escherichia coli* cell surface for detoxification and detection of organophosphates. J Agric Food Chem 61:7810-7816

Mahoney TF, Ricci DP, Silhavy TJ (2016) Classifying β-barrel assembly substrates by manipulating essential BAM complex members. J Bacteriol 198(14):1984-1992

Macauley S, McNeil B, Harvey LM (2001) The genus *Gluconobacter* and its applications in biotechnology. Crit Rev Biotechnol 21(1):1-25

Manoil C, Beckwith J (1985) Tn*phoA*: A transposable probe for protein export signals. Proc Natl Acad Sci USA 82:8129-8133

Martineau P, Charbit A, Leclerc C, Werts C, O'Callaghan D, Hofnung M (1991) A genetic system to elicit and monitor anti-peptide antibodies without peptide synthesis. Nat Biotechnol 9(2):170-172

Matsushita K, Toyama H, Adachi O (1994) Respiratory chains and bioenergetics of acetic acid bacteria. Adv Microb Physiol 36:247-301

Matsushita K, Toyama H, Adachi O (2004) Respiratory chains in acetic acid bacteria: membrane-bound periplasmic sugar and alcohol respirations. In: Zannoni D (ed)

Respiration in bacteria and archaea, vol 2: diversity of prokaryotic respiratory systems, vol 2. Springer, Dordrecht (Netherlands), pp 81-99

Meyer M, Schweiger P, Deppenmeier U (2013) Effects of membrane-bound glucose dehydrogenase overproduction on the respiratory chain of *Gluconobacter oxydans*. Appl Microbiol Biotechnol 97(8):3457-3466

Michaelis S, Inouye H, Oliver D, Beckwith J (1983) Mutations that alter the signal sequence of alkaline phosphatase in *Escherichia coli.* J Bacteriol 154(1):366-374

Michon C, Langella P, Eijsink VGH, Mathiesen G, Chatel JM (2016) Display of recombinant proteins at the surface of lactic acid bacteria: strategies and applications. Microb Cell Fact 15:70

Murphy CK, Klebba PE (1989) Export of FepA::PhoA fusion proteins to the outer membrane of *Escherichia coli* K-12. J Bacteriol 171(11):5894-5900

Nagai K, Thøgersen HC (1984) Generation of *β*-globulin by sequence-specific proteolysis of a hybrid protein produced in *Escherichia coli*. Nature 309(5971):810-812

Narita J, Okano K, Tateno T, Tanino T, Sewaki T, Sung MH, Fukuda H, Kondo A (2006) Display of active enzymes on the cell surface of *Escherichia coli* using PgsA anchor protein and their application to bioconversion. Appl Microbiol Biotechnol. 70(5):564-572

Nhan NT, de Valdivia EG, Gustavsson M, Hai TN, Larsson G (2011) Surface display of Salmonella epitopes in *Escherichia coli* and *Staphylococcus carnosus*. Microb Cell Fact 10:22

Noinaj N, Gumbart JC, Buchanan SK (2017) The β-barrel assembly machinery in motion. Nat Rev Microbiol 15(4):197-204

Olijve W, Kok JJ (1979) Analysis of growth of *Gluconobacter oxydans* in glucose containing media. Arch Microbiol. 121:283-290.

Pappenberger G, Hohmann HP (2014) Industrial production of L-ascorbic acid (vitamin C) and D-isoascorbic acid. Adv Biochem Eng Biotechnol 143:143-188

Paramasivam N, Habeck M, Linke D (2012) Is the C-terminal insertional signal in Gram-negative bacterial outer membrane proteins species-specific of not? BMC Genomics 13:510

Pearson KS (2014) The development of genetic tools for protein characterization and genetic engineering in the acetic acid bacterium *Gluconobacter oxydans*. Masters thesis, Missouri State University

Peterson TN, Brunak S, von Heijne G, Nielson H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8(10):785-786

Prust C, Hoffmeister M, Liesegang H, Wiezer A, Frick WF, Ehrenreich A, Gottschalk G, Deppenmeier U (2005) Complete genome sequence of the acetic acid bacterium *Gluconobacter oxydans*. Nat Biotechnol 23:195-200

R Core Team (2017) R: a language and environment for statistical computing. R for Statistical Computing, Vienna (Austria), https://www.R-project.org/

Raspor P, Goranovič D (2008) Biotechnological applications of acetic acid bacteria. Crit Rev Biotechnol 28(2):101-124

Rauch B, Pahlke J, Schweiger P, Deppenmeier U (2010) Characterization of enzymes involved in the central metabolism of *Gluconobacter oxydans*. Appl Microbiol Biotechnol 88(3):711-718

Richhardt J, Bringer S, Bott M (2012) Mutational analysis of the pentose phosphate and Entner-Doudoroff pathways in *Gluconobacter oxydans* reveals improved growth of a Δ*edd* and Δ*eda* mutant on mannitol. Appl Environ Microbiol 78(19):6975-6986

Robert V, Volokhina EB, Senf F, Bos MP, Gelder PV, Tommassen J (2006) Assembly factor Omp185 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. PLoS Biol 4(11):e377

Römling U, Galperin MY (2015) Bacterial cellulose biosynthesis: diversity of operons, subunits, products, and functions. Trends Microbiol 23(9):545-557

Ryu S, Karim MN (2011) A whole cell biocatalyst for cellulosic ethanol production from dilute acid-pretreated corn stover hyrdolyzates. Appl Microbiol Biotechnol 91(3):529-542

Saleem M, Brim H, Hussain S, Arshad M, Leigh MB, Zia-ul-hassan (2008) Perspectives on microbial cell surface display in bioremediation. Biotechnol Adv 26:151-161

Schedel M (2000) Regioselective oxidation of aminosorbitol with *Gluconobacter oxydans*, key reaction in the industrial 1-deoxynojirimycin synthesis. In: Rehm HJ, Reed G (eds) Biotechnology, vol 8b. Wiley-VHC, Weinheim (Germany), pp 296-307

Schüürmann J, Quehl P, Gunter F, Jose J (2014) Bacterial whole-cell biocatalysis by surface display of enzymes: toward industrial application. Appl Microbiol Biotechnol 98(19):8031-8046

Shigechi H, Koh J, Fujita Y, Matsumoto T, Bito Y, Ueda M, Satoh E, Fukada H, Kondo A (2004) Direct production of ethanol from raw corn starch via fermentation by use of a novel surface-engineered yeast strain codisplaying glucoamylase and α-amylase. Appl Environ Microbiol 70(8):5037-5040

Steele HL, Jaeger KE, Daniel R, Streit WR (2009) Advances in recovery of novel biocatalysts from metagenomics. J Mol Microbiol Biotechnol 16:25-37

Strathopoulos C, Georgiou G, Earhart CF (1996) Characterization of *Escherichia coli* expressing an Lpp'OmpA(46-159)-PhoA fusion protein localized in the outer membrane. Appl Microbiol Biotechnol 45(1-2):112-119

Sugawara E, Nagaro K, Nikaido H (2012) Alternative folding pathways of the major porin OprF of *Pseudomonas aeruginosa*. FEBS J 279:910-918.

Terpe K (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol 60(5):523-533

Tran AT, Struck DK, Young R (2005) Periplasmic domains define holing-antiholin interactions in T4 lysis inhibition. J Bacteriol 187(19):6631-6640

Tsirigotaki A, Geyter JD, Šoštarić N, Economou A, Karamanou S (2017) Protein export through the bacterial Sec pathway. Nature Rev Microbiol 15(1):21-36

Ufarté L, Potocki-Veronese G, Laville É (2015) Discovery of new protein families and functions: new challenges in functional metagenomics for biotechnologies and microbial ecology. Front Microbiol 6:563

van Bloois E, Winter RT, Kolmar H, Fraaije MW (2011) Decorating microbes: surface display of proteins on *Escherichia coli*. Trends Biotechnol 29(2):79-86

Wei W, Liu X, Sun P, Wang X, Zhu H, Hong M, Mao ZW, Zhao J (2014) Simple whole-cell biodetection and bioremediation of heavy metals based on an engineered lead-specific operon. Environ Sci Technol 48(6):3363-3371

Wendel S, Fischer EC, Martínez V, Seppälä S, Nørholm MHH (2016) A nanobody:GFP bacterial platform that enables functional enzyme display and easy quantification of display capacity. Microb Cell Fact 15:71

Woodley JM (2008) New opportunities for biocatalysis: making pharmaceutical processes greener. Trends Biotechnol 26(6):321-327

Wu PH, Giridhar R, Wu WT. 2006. Surface display of transglucosidase on *Escherichia coli* by using the ice nucleation proteins of *Xanthomonas campestris* and its application in glucosylation of hydroquinone. Biotechnol Bioeng 95(6):1138-1147

Yang C, Freudl R, Qiao C, Mulchandani A (2010) Cotranslocation of methyl parathion hydrolase to the periplasm and of organophosphorus hydrolase to the cell surface of Escherichia coli by the Tat pathway and the ice nucleation protein display system. Appl Environ Microbiol 76(2):434-440

Yang H, Liu L, Xu F (2016) The promises and challenges of fusion constructs in protein biochemistry and enzymology. Appl Microbiol Biotechnol 100(19):8273-8281

Yang W, Xu H (2016) Industrial Fermentation of Vitamin C. In: Revuelta JL, Vandamme EJ (eds) Industrial biotechnology of vitamins, biopigments, and antioxidants. Wiley-VCH Verlag GmbH & Co KGaA, Weinheim (Germany), pp 161-192

Yin K, Lv M, Wang Q, Wu Y, Liao C, Zhang W, Chen L (2016) Simultaneous bioremediation and biodetection of mercury ion through surface display of carboxylesterase E2 from *Pseudomonas aeruginosa* PA1. Water Res 103:383-390

Zeiser J, Mühlenbeck LH, Schweiger P, Deppenmeier U (2014) Characterization of a periplasmic quinoprotein from *Sphingomonas wittichii* that functions as aldehyde dehydrogenase. Appl Microbiol Biotechnol 98(5):2067-2079

APPENDIX. R CODE

```
#Download required packages:
install.packages("dplyr")
install.packages("ggplot2")
install.packages("growthcurver")
install.packages("plyr")
install.packages("multcomp")
install.packages("reshape2")

#Load required packages:
library("dplyr")
library("ggplot2")
library("growthcurver")
library("plyr")
library("multcomp")
library("reshape2")

#Figure 8:
AnchorData=read.csv('AnchorData.csv', header=TRUE)
AnchorDatamelt <- melt(AnchorData)
ANOVAAnchorData<-aov(value~variable, data=AnchorDatamelt)
summary(ANOVAAnchorData)
TukeyHSD(ANOVAAnchorData, conf.level=0.95)
tuk<-glht(ANOVAAnchorData, linfct=mcp(variable="Tukey"))
cld(tuk)
par(family='serif')
boxplot(AnchorData, las=2, at=c(1,2,3,4,6,7,8,9,11,12,13,14),
ylab=expression(bold(paste(Delta,'A'[405]*'/hr×OD'[600]))),  par(mar=c(10,4,3,0.5),
par(mgp=c(2.5,1,0))))
stripchart(value ~ variable, vertical=TRUE, data=AnchorDatamelt,
at=c(1,2,3,4,6,7,8,9,11,12,13,14), method='overplot', add=TRUE, pch=20, col='black')
mtext(expression(bold(bolditalic(E.~coli)~Strain)),side=1,line=8)
groups<-c('a','a','a','a','a','d','a','c','a','a','a','b')
mtext(groups, at=c(1,2,3,4,6,7,8,9,11,12,13,14), side=3, line=0.75, font=2)
axis(side=3, label=NA, at=c(1,2,3,4,6,7,8,9,11,12,13,14))

#Figure 9:
OprFDataGox=read.csv('OprFDataGox.csv', header=TRUE)
OprFDataGoxmelt<-melt(OprFDataGox)
ANOVAOprFDataGox<-aov(value~variable, data=OprFDataGoxmelt)
summary(ANOVAOprFDataGox)
TukeyHSD(ANOVAOprFDataGox, conf.level=0.95)
tuk<-glht(ANOVAOprFDataGox, linfct=mcp(variable="Tukey"))
cld(tuk)
```

```
par(family='serif')
boxplot(OprFDataGox, las=2, at=c(1,2,3,4,6),
ylab=expression(bold(paste(Delta,'A'[405]*'/hr×OD'[600]))), par(mar=c(10,4,3,0.5),
par(mgp=c(2.5,1,0))))
stripchart(value ~ variable, vertical=TRUE, data=OprFDataGoxmelt, method='overplot',
add=TRUE, pch=20, col='black', at=c(1,2,3,4,6))
mtext(expression(bold(bolditalic(G.~oxydans)~Strain)), side=1, line=8.5)
groups<-c('a','c','a','b','a')
mtext(groups, at=c(1,2,3,4,6), side=3, line=0.75, font=2)
axis(side=3, label=NA, at=c(1,2,3,4,6))
ddply(OprFDataGoxmelt, c("variable"), summarize, Mean=mean(value), SD=sd(value),
SEM=sd(value)/sqrt(length(value)))

#Figure 11:
Cleavable<-read.csv('Cleavable2.csv', header=TRUE)
Cleavablemelt<-melt(Cleavable)
CleavableANOVA<-aov(value~variable,data=Cleavablemelt)
TukeyHSD(CleavableANOVA, conf.level=0.95)
tuk<-glht(CleavableANOVA, linfct=mcp(variable="Tukey"))
cld(tuk)
par(family='serif')
boxplot(Cleavable, las=2, ylim=c(0,0.22), outline=TRUE,
ylab=expression(bold(paste(Delta,'A'[405]*'/hr×OD'[600]))), par(mar=c(8.5,4.5,3,0.5),
par(mgp=c(3,1,0))))
stripchart(value ~ variable, vertical =TRUE, outline=FALSE, data=Cleavablemelt,
method='overplot', add=TRUE, pch=20, col='black', outline=FALSE)
mtext(expression(bold(Sample)), side=1, line=6.5)
groups<-c('c','b','a','b')
mtext('c', at=c(2), line=1.75, font=2)
mtext(groups, at=c(1,2,3,4), side=3, line=0.75, font=2)
axis(side=3, label=NA, at=c(1,2,3,4))

#Figure 14:
Ecoli264Linker<-read.csv("Ecoli-264-Linker.csv", header=TRUE)
Ecoli264Linkermelt<-melt(Ecoli264Linker)
Ecoli264LinkerANOVA<-aov(value~variable, data=Ecoli264Linkermelt)
TukeyHSD(Ecoli264LinkerANOVA, conf.level=0.95)
tuk<-glht(Ecoli264LinkerANOVA, linfct = mcp(variable="Tukey"))
cld(tuk)
par(family='serif')
boxplot(Ecoli264Linker, las=2, ylim=c(0.00,1.40),
ylab=expression(bold(paste(Delta,'A'[405]*'/hr×OD'[600]))), par(mar=c(11,4,4,0.5),
par(mgp=c(2.5,1,0))))
stripchart(value ~ variable, vertical=TRUE, data=Ecoli264Linkermelt, method='overplot',
add=TRUE, pch=20, col='black')
mtext(expression(bold(bolditalic(E.~coli)~Strain)), side=1, line=9)
```

```
groups<-c('b','b','a','a','c','b','d')
groups2<-c('c','d','c')
mtext(groups, at=c(1,2,3,4,5,6,7), side=3, line=0.75, font=2)
mtext(groups2, at=c(1,5,6), side=3, line=1.75, font=2)
axis(side=3, label=NA, at=c(1,2,3,4,5,6,7))

#Figure 15:
Ecoli452Linker<-read.csv("Ecoli-452-Linker.csv", header=TRUE)
Ecoli452Linkermelt<-melt(Ecoli452Linker)
Ecoli452LinkerANOVA<-aov(value~variable, data=Ecoli452Linkermelt)
TukeyHSD(Ecoli452LinkerANOVA, conf.level=0.95)
tuk<-glht(Ecoli452LinkerANOVA, linfct=mcp(variable="Tukey"))
cld(tuk)
par(family='serif')
boxplot(Ecoli452Linker, las=2, ylim=c(0.00,1.40),
ylab=expression(bold(paste(Delta,'A'[405]*'/hr×OD'[600]))), par(mar=c(11,4,4,0.5),
par(mgp=c(2.5,1,0))))
stripchart(value ~ variable, vertical=TRUE, data=Ecoli452Linkermelt, method='overplot',
add=TRUE, pch=20, col='black')
mtext(expression(bold(bolditalic(E.~coli)~Strain)), side=1, line=9)
groups<-c('a','b','b','b','a','b','b')
mtext(groups, at=c(1,2,3,4,5,6,7), side=3, line=0.75, font=2)
axis(side=3, label=NA, at=c(1,2,3,4,5,6,7))

#Figure 16:
Gox264Linker<-read.csv("Gox-264-Linker.csv", header=TRUE)
Gox264Linkermelt<-melt(Gox264Linker)
Gox264LinkerANOVA<-aov(value~variable, data=Gox264Linkermelt)
TukeyHSD(Gox264LinkerANOVA, conf.level=0.95)
tuk<-glht(Gox264LinkerANOVA, linfct=mcp(variable="Tukey"))
cld(tuk)
par(family='serif')
boxplot(Gox264Linker, las=2, ylim=c(0,0.7),
ylab=expression(bold(paste(Delta,'A'[405]*'/hr×OD'[600]))), par(mar=c(11,5,4,0.5),
par(mgp=c(2.5,1,0))))
stripchart(value ~ variable, vertical=TRUE, data=Gox264Linkermelt, method='overplot',
add=TRUE, pch=20, col='black')
mtext(expression(bold(bolditalic(G.~oxydans)~Strain)), side=1, line=9)
groups<-c('b','b','c','b','a')
mtext(groups, at=c(1,2,3,4,5), side=3, line=0.75, font=2)
axis(side=3, label=NA, at=c(1,2,3,4,5))

#Figure 17:
Gox452Linker<-read.csv("Gox-452-Linker.csv", header=TRUE)
Gox452Linkermelt<-melt(Gox452Linker)
Gox452LinkerANOVA<-aov(value~variable, data=Gox452Linkermelt)
```

```
TukeyHSD(Gox452LinkerANOVA,conf.level=0.95)
tuk<-glht(Gox452LinkerANOVA, linfct=mcp(variable="Tukey"))
cld(tuk)
par(family='serif')
boxplot(Gox452Linker, las=2, ylim=c(0,0.7),
ylab=expression(bold(paste(Delta,'A'[405]*'/hr×OD'[600]))), par(mar=c(11,5,4,0.5),
par(mgp=c(2.5,1,0))))
stripchart(value ~ variable, vertical=TRUE, data=Gox452Linkermelt, method='overplot',
add=TRUE, pch=20, col='black')
mtext(expression(bold(bolditalic(G.~oxydans)~Strain)), side=1, line=9)
groups<-c('c','d','e','b','a','a','a')
groups2<-c('b')
mtext(groups, at=c(1,2,3,4,5,6,7), side=3, line=0.75, font=2)
mtext(groups2, at=c(7), side=3, line=1.75, font=2)
axis(side=3, label=NA, at=c(1,2,3,4,5,6,7))

#Figure 18:
data264<-read.csv("264 GC.csv")
platemap452<-read.csv("452 Platemap.csv")
shape264<-melt(data264, id=c("Time", "Temperature"), variable.name="Well",
value.name="OD595")
annotated264<-inner_join(shape264, platemap264, by="Well")
conf_int95<-function(data){n<-length(data)
error<-qt(0.975, df=n-1) * sd(data)/sqrt(n)
return(error)}
stats264<-annotated264 %>% group_by(Strain, Time) %>%
summarize(N=length(OD595), Average=mean(OD595), CI95=conf_int95(OD595))
%>%filter(!is.na(Strain))

#Figure 19:
data452<-read.csv("452 GC.csv")
platemap264<-read.csv("264 Platemap.csv")
shape452<-melt(data452, id=c("Time", "Temperature"), variable.name="Well",
value.name="OD595")
annotated452<-inner_join(shape452, platemap452, by="Well")
conf_int95<-function(data){n<-length(data)
error<-qt(0.975, df=n-1) * sd(data)/sqrt(n)
return(error)}
stats452 <- annotated452 %>% group_by(Strain, Time) %>%
summarize(N=length(OD595), Average=mean(OD595), CI95=conf_int95(OD595))
%>%filter(!is.na(Strain))

#To calculate doubling times for Figures 20 and 21:
d<-read.csv("GC.csv")
Plate<-SummarizeGrowthByPlate(d)
write.csv(Plate,"DT.csv")
```

```
#Figure 20:
DT264=read.csv('264 DT.csv',header=TRUE)
DT264melt<-melt(DT264)
ANOVADT264<-aov(value~variable,data=DT264melt)
TukeyHSD(ANOVADT264, conf.level=0.95)
tuk<-glht(ANOVADT264, linfct=mcp(variable="Tukey"))
cld(tuk)
par(family='serif')
par(mar=c(10,4,3,0.5), mgp=c(2.5,1,0))
stripchart(value ~ variable, vertical=TRUE, data=DT264melt,method='stack', pch=20,
col='black', las=2, ylab=expression(bold('Doubling time (min)')), ylim=c(0,150))
mtext(expression(bold(bolditalic(G.~oxydans)~Strain)), side=1, line=9)
groups<-c('b','d','b','c','a','b','e')
mtext(groups, at=c(1,2,3,4,5,6,7),side=3, line=0.75, font=2)
axis(side=3,label=NA,at=c(1,2,3,4,5,6,7))
ddply(DT264melt, c("variable"), summarise, Mean=mean(value), SD=sd(value),
SEM=sd(value)/sqrt(length(value)))

#Figure 21:
DT452=read.csv('452 DT.csv',header=TRUE)
DT452melt<-melt(DT452)
ANOVADT452<-aov(value~variable,data=DT452melt)
TukeyHSD(ANOVADT452, conf.level=0.95
tuk<-glht(ANOVADT452, linfct=mcp(variable="Tukey"))
cld(tuk)
par(family='serif')
par(mar=c(10,4,3,0.5), mgp=c(2.5,1,0)))
stripchart(value ~ variable, vertical=TRUE, data=DT452melt, method='stack', pch=20,
col='black', las=2, ylab=expression(bold('Doubling time (min)')), ylim=c(0,150))
mtext(expression(bold(bolditalic(G.~oxydans)~Strain)), side=1, line=9)
groups<-c('a','c','e','c','b','f','b','b','d')
groups2<-c('de','f','d','c','c','f')
mtext(groups, at=c(1,2,3,4,5,6,7,8,9), side=3, line=0.75, font=2)
mtext(groups2, at=c(2,3,4,5,7,9), side=3, line=1.75, font=2)
axis(side=3, label=NA, at=c(1,2,3,4,5,6,7,8,9))
ddply(DT452melt, c("variable"), summarise, Mean=mean(value), SD=sd(value),
SEM=sd(value)/sqrt(length(value)))
```