Engineering *Lactococcus lactis* for the scaffold protein-mediated surface display of recombinant enzymes

Andrew Wieczorek

A Thesis

In

The Department

Of

Biology

Presented in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy at

Concordia University

October 2012

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: **Andrew Wieczorek**

Entitled: **Engineering *Lactococcus lactis* for the scafford-mediated surface display of recombinant enzymes**

and submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY (Biology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Chair
Dr. G. Dafni

_____ External Examiner
Dr. S-L. Wong

_____ External to Program
Dr. P. Darlington

_____ Examiner
Dr. R. Storms

_____ Examiner
Dr. J. Turnbull

_____ Thesis Supervisor
Dr. V. Martin

Approved by _____
Dr. S. Dayanandan, Graduate Program Director

December 18, 2012 _____
Dr. B. Lewis, Dean, Faculty of Arts and Science

Abstract

Engineering *Lactococcus lactis* for the scaffold-mediated surface display of recombinant enzymes

Andrew Wieczorek

Concordia University, 2012. DOCTOR OF PHILOSOPHY (Biology).

Multi-enzyme complexes are responsible for the synthesis of a number of biochemical compounds and the degradation of complex polymers. An example of the latter is the degradation of cellulose by enzyme complexes termed “cellulosomes” which are produced by several bacteria of the class Clostridia. The basic structure of a cellulosome comprises a central scaffold protein, which associates with a multitude of cellulases via cohesin-dockerin interactions. The advent of cellulose utilization as feedstock for producing biofuels has garnered much interest towards designing custom-tailored recombinant cellulosomes and expressing them in microbes of interest. The metabolic diversity among bacteria also make this approach an appealing strategy for bestowing cellulolytic capabilities upon organisms which produce non-biofuel commodity chemicals such as lactic acid, succinic acid, acetone, amino acids, food additives and carotenoids. In addition, the display of recombinant multi-enzyme complexes in bacteria can yield novel insights into the mechanisms and parameters affecting their secretion, assembly and function. The industrially relevant lactic acid bacterium, *Lactococcus lactis*, is a model organism for the secretion and display of recombinant proteins, and the numerous biological techniques available for its manipulation make this organism particularly appealing for such a task. In this thesis, I present my work describing the incremental

steps taken towards the surface display of custom-tailored multi-enzyme complexes on the surface of *L. lactis*. Chapter 1 describes the proof of concept for this project, including the choice of promoters, secretion signal peptide, and reporter enzymes. It also discusses the major bottlenecks observed based on the organism's physiology. Chapter 2 describes the engineering of scaffold chimeras with cohesins of different specificity and the display of two enzymes on such scaffolds. I also investigated the catalytic profiles of the resulting complexes when enzymes were simultaneously or sequentially bound to the displayed scaffold. Finally, chapter 3 describes the optimization of the type 2 dockerin-cohesin interaction by the inclusion of the CipA "X" module, as well as the engineering of enzyme complexes with novel architectures by use of secondary "adapter" scaffolds and the subsequent assembly of multi-scaffold complexes. Also investigated is the potential of using a dual-plasmid system for the full *in vivo* assembly of such complexes without the exogenous addition of components.

Acknowledgments

At a crossroads I came to as an undergraduate student, I was seeking an environment, which would nurture both my scientific and creative inclinations. My final semester at Concordia would provide me with the opportunity to embark on a journey which would accomplish just that, thanks to my meeting Dr. Vincent Martin. He has allowed me to apply unconventional and creative ideas towards my scientific career. As the first graduate student in the Martin lab, coming from a background, which was not rich in knowledge of molecular biology, he taught me the value of being both an independent learner and creative thinker. I am thankful to have had a supervisor whose office was always open for me to just pop in and suggest somewhat eccentric avenues to explore in my research. Some were rejected and some are included in this thesis, but my mentor always took me seriously. Thanks Dr. M.

My parents have sacrificied everything and anything they could to ensure that I had every opportunity in my life to do what I am passionate about. I owe my life to them. My labmates are no longer labmates, but lifelong friends. They have made these the best years of my life. Thanks Nick, Dom, Euan, and Coco. Thank you Paul for always supporting me, and even coming to my first conference. Thank you Mara for your constant patience, care, love, support, and patience. You've saved my life more than once. Finally, and certainly not least, I will forever be eternally grateful to the amazing faculty at Concordia. Thank you Dr. Storms and Dr. Turnbull for your guidance and advice. Jo, you're quite the dancer. Dr. Zerges, Dr. Gulick, and the rest of the Biology faculty and staff, I have sat in your classes as an undergraduate student, assisted you as a TA in your labs, attended and

gave departmental seminars which you gave me tremendous feedback on, and now am privileged to even have the chance to teach and train undergraduate students in our department just as you did for me. I am eternally grateful for the opportunity and experiences you have given me. Thank you.

Dedication

Being a graduate student is both filled with times of rejoice as well as times of despair, and my parents have been along my side every step of the way. It is cliché to perhaps say that I would not be here if not for my parents, but I understand this in ways many others could not possibly understand. I dedicate the sacrifice and hard work I put into this endeavor to you, since you have worked so hard and sacrificed so much for me. I also dedicate this to Paul for always encouraging me to do anything and everything I want. You were at my concerts, conferences, scotch nights and all the rest. I dedicate this to you as well. Mara, you are my rock in every sense of the word. The beginning of my graduate studies coincided with us falling in love all over again, and I honestly cannot imagine having done this without you.

Family and friends, I dedicate this to you. Babcia, Mark, and Uncle Morris, in life you gave me love and support, and in death you give me a new perspective on life. Thank you.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CBP ..... Consolidated Bioprocessing

EGL ..... Endoglucanase

CBH ..... Cellobiohydrolase

BGL ..... (Beta)-glucosidase

CBD ..... Cellulose-binding domain

SLH ..... S-layer homologous (domain)

PASC ..... Phosphoric acid-swollen cellulose

XBD ..... Xylan-binding Module

CEM ..... Cellulose-enzyme-microbe (Complex)

SSF ..... Simultaneous saccharification and fermentation

LAB ..... Lactic acid bacteria

GRAS ..... Generally regarded as safe

sp_{Usp45} ..... Secretion signal peptide from Usp45 protein

NICE ..... Nisin-controlled gene-expression (system)

PnisA ..... Inducible promoter of *nisA* gene

TBD ..... Toluidine-blue (agar)

cwa_{M6} ..... Cell-wall anchor of M6 protein

M17 ..... Growth medium for *L. lactis* (lacking carbon source)

$CipA_{frags}$ ..... Fragments of scaffold protein CipA

P59 ..... Constitutive lactococcal promoter

rbs_{Usp45} ..... Ribosome-binding site of *usp45* gene

rbs_{NisA}	Ribosome-binding site of *nisA* gene
lacZ-α	Gene encoding α fragment of *E. Coli* β-galactosidase LacZ
coh1 (coh1C1)	First type 1 cohesin of CipA
coh2 (coh1C2)	Second type 1 cohesin of CipA
coh3 (coh1C3)	Third type 1 cohesin of CipA
coh9 (coh1C9)	Ninth type 1 cohesin of CipA
CBM3a (CBD)	Cellulose-binding module (Cellulose-binding domain)
GM17	M17 media supplemented with glucose
PCR	Polymerase chain reaction
GUS	Buffer for β-glucuronidase assays
BSA	Bovine serum albumin
dock1	Type 1 dockerin domain of CelS
$OlpB_{frags}$	Fragments of anchor protein OlpB
$SdbA_{frag}$	Fragment of anchor protein SdbA
Coh2O2	Second type 2 cohesin of OlpB
link	Inter-cohesin linker region
Coh2S1	Type 2 cohesin of SdbA
dock2	Type 2 dockerin domain from CipA
Z Buffer	Buffer for β-galactosidase assay
PNP	*p*-nitrophenyl β-D-glucuronide
HisX6	N-terminal His-His-His-His-His-His for affinity chromatography
ORF	Open-Reading Frame

PRS ..... DNA encoding *PnisA-rbs*$_{NisA}$-*sp*$_{Usp45}$

RS ..... DNA encoding *rbs*$_{NisA}$-*sp*$_{Usp45}$

X-dock2 ..... CipA$_{frag}$ containing X module and dock2 domain

SR ..... Streptococcal regeneration (media)

Goal and Objectives

Goal: Recombinant multi-enzyme complexes can potentially contribute to bioconversion processes that require several reactions being carried out simultaneously or sequentially, (Bayer, et al., 2004; Conrado, et al., 2008). Studying the assembly parameters of such recombinant complexes may also shed light on how natural multi-enzyme complexes assemble *in vivo.* The model organism *L. lactis* is an attractive host for the production of recombinant protein complexes due to its ability to secrete and surface display several recombinant proteins (Le Loir, et al., 2005; Leenhouts, et al., 1999), and due to the availability of a controlled gene expression system which may potentially prevent cellular toxicity (Mierau & Kleerebezem, 2005). My goal was therefore to display recombinant multi-enzyme complexes with precise enzyme compositions and architectures on the surface of *L. lactis in vivo*.

Objectives: A first objective to achieve this goal was the surface display of a simple scaffold on the surface of *L. lactis*, capable of binding a fusion enzyme via interactions between domains located on each component. As a proof of concept experiment, this is included in chapter 1. The use of non-enzyme reporters was also attempted, and preliminary results are thus included in appendices.

As a second objective, chimeric scaffolds consisting of cohesins of different specificities were displayed in order to precisely dictate the location of two dockerin-fused enzymes on the resulting scaffolds. I also sought to investigate whether the simultaneous or sequential addition of enzymes resulted in differences in binding at the

respective cohesin sites. Chapter 2 therefore describes the strategy which I adopted towards the generation of such bi-enzymatic complexes and the study of their assembly on the surface of *L. lactis*. A library of non-enzyme reporter proteins were also engineered for the future potential simultaneous targeting of specific protein pairs onto the displayed scaffolds. A description of this approach is therefore provided in the appendices, along with the results of a preliminary investigation of their ability to bind wild type *C. thermocellum* cells.

My third objective was to generate two-scaffold complexes as inspired by the cellulosomes of *A. cellulolyticus* and *R. flavefaciens* as well as the production of all components *in vivo* by use of a dual-plasmid system. Chapter 3 therefore describes the successful assembly of two-level enzyme complexes, as well as a preliminary analysis of the ability to generate both enzyme and scaffold components *in vivo*.

1 Introduction

1.1 Multi-enzyme complexes

Macromolecular enzyme complexes catalyze an array of biochemical and metabolic processes such as the degradation of proteins (Lowell, Ballou, et al., 1988; Lowell, Smith, et al., 1988) or recalcitrant polymers (Bayer, et al., 2004) as well as the synthesis of metabolic products via substrate channeling (Conrado, et al., 2008). From a biotechnological perspective, mimicking such process by incorporating catalytic modules or enzymes of interest within synthetic protein complexes can significantly enhance the efficiency of such bioprocesses via substrate channeling (Dueber, et al., 2009) and increased enzyme synergy (Bayer, et al., 2004). Examples include the spatial organization of enzymes through compartmentalization in organelles, co-localization on membranes, or assembly in complexes using protein scaffolds or fusions, which all play an important role in controlling the flow of metabolites in a cell (Conrado, et al., 2008; H. Lee, et al., 2011). The spatial organization of multi-enzyme pathways can serve many functions such as substrate channeling to reduce the loss of intermediates to competing side reactions. Channeling can also be used to prevent the accumulation of toxic or unstable metabolites (Sampson & Bobik, 2008). Higher localized concentration of proteins and metabolites, dubbed molecular crowding, also decreases product/reactant diffusion and increases yields and rates of metabolite production (Conrado, et al., 2008).

Spatial organization is used to control the stoichiometry of the proteins that make up the complex and to protect proteins from degradation (H. Lee, et al., 2011). Synergism between enzymes in a complex can also result in an activity that is higher than the sum of

its parts, as demonstrated for cellulosomes, which will be discussed further in following sections (Bayer, et al., 2004). Inspired by nature and driven by the need to achieve high production yields in industrial microbes, metabolic engineers have started tinkering with the spatial organization of enzymes in cells using synthetic protein scaffolds and organelles (Delebecque, et al., 2011; Dueber, et al., 2009; Farhi, et al., 2011).

1.2 Cellulose: A model substrate for multi-enzyme complexes

Interest in organizing enzymes on synthetic protein scaffolds stems from the desire to produce intracellularly, secrete, and/or surface-display enzyme complexes to control flux of metabolic pathways (Dueber, et al., 2009; H. Lee, et al., 2011), or degrade complex polymers (Anderson, et al., 2011; Tsai, et al., 2009). The organism of choice, in large part, depends on the desired metabolic end product, and much attention in past decades has been on the production alternative fuels such as bioethanol. Petroleum is the major source of energy for satisfying transportation needs, and it is a major feedstock for the production of commodity chemicals (Shanmugam & Ingram, 2008). In most industrialized countries, energy security is being threatened by depleting petroleum reserves (Farrell, et al., 2006), and this is being addressed from a biotechnological perspective (S. K. Lee, et al., 2008; Ragauskas, et al., 2006).

Lignocellulosic biomass is a sustainable alternative to fossil resources, and has the added advantage of not competing with human and animal nutrition. Cellulose is the main polymer component of lignocellulosic feedstock, and since it is composed of glucose monomers, it is potentially a good feedstock for fermentation by microorganisms

that have been engineered to produce high yields of fuels and commodity chemicals such as ethanol, butanol, lactic acid, succinic acid, acetic acid, and many others (Shanmugam & Ingram, 2008). Still, the net bioconversion of cellulose into valuable chemicals requires multiple steps, which include hydrolysis of cellulose into fermentable sugars, followed by subsequent fermentation into metabolic end-products. To reduce costs, a strategy known as consolidated bioprocessing (CBP) has been proposed, which entails the *in situ* production of cellulases by the fermenting organism (L. R. Lynd, et al., 2005). This strategy consolidates enzyme production, hydrolysis and fermentation into a single step. However, CBP requires an organism efficient at both degrading cellulose and fermenting glucose to a single product at high titers. Such an organism does not exist in nature (L. Lynd, 1996). In addition, the complete breakdown of cellulose to glucose requires the cooperation of three different types of cellulases. Endoglucanases (EGLs) cleave amorphous cellulose randomly at endo sites to release cellodextrins of various lengths (L. R. Lynd, et al., 2002). Cellobiohydrolases (CBHs), on the other hand, are required for the hydrolysis of crystalline cellulose, and release cellobiose by acting at the reducing and non-reducing ends of cellulose strands (L. R. Lynd, et al., 2002). Finally, β-glucosidases (BGLs) produce glucose from the hydrolysis of the cellobiose and cello-oligomers produced by EGLs and CBHs. The three types of enzymes are believed to act synergistically (L. R. Lynd, et al., 2002). EGLs cleave at random inside strands, creating termini for CBHs, which in turn contribute to loosening of cellulose crystallinity, making further material available to EGLs (L. R. Lynd, et al., 2002). Some cellulases, as well as other proteins involved in cellulose degradation, carry a cellulose-binding domain (CBD)

that is used to tether them to the polymeric substrate, and allows for the processive degradation of cellulose by crawling along its strands (Watanabe & Tokuda, 2010). Certain organisms have therefore evolved to assemble their cellulases into multi-enzyme complexes termed cellulosomes, notably to enhance synergy between enzymes and promote substrate channeling (Bayer, et al., 2004).

1.3 Cellulosomes

Cellulolytic fungi are capable of degrading crystalline cellulose by secreting cocktails of hydrolytic enzymes that act in synergy to degrade the polymer (Y. H. Zhang & Lynd, 2004). An alternative strategy adopted by certain fungi as well as anaerobic bacteria of the class Clostridia entails the hydrolysis of cellulose by cellulosomes (Bayer, et al., 2004). Cellulosomes are protein complexes comprised of a multitude of hydrolytic enzymes with varying catalytic properties that associate with a central scaffold protein to enhance synergy when degrading cellulose (Gerngross, et al., 1993). The assembly of the cellulosome complex is mediated via non-covalent interactions between non-catalytic dockerin and cohesin domains. These domains serve as the "glue" which hold the complex together and dictate its architecture. Two characteristics of a dockerin and cohesin pair determine the specificity of the interaction: the species from which they are derived (Bayer, et al., 2004), as well as the type of cohesin and dockerin partner involved (Leibovitz & Beguin, 1996). In other words, the cohesin-dockerin interaction is species-specific, and type 1 and type 2 cohesins from a single specie do not interact with

dockerins of the opposite type (e.g. type 1 cohesins do not interact with type 2 dockerins, and vice-versa).

1.3.1 The cellulosome of *C. thermocellum*

C. thermocellum is of great interest as a candidate for CBP since it has an extraordinary cellulolytic capacity, it is thermophilic (cellulose hydrolysis is accelerated at high temperatures), and can be co-cultured with pentose-fermenting thermophilic organisms such as *Thermoanaerobacter saccharolyticum* (Bayer, et al., 2004; Schwarz, 2001). It has also been demonstrated that electrotransformation of this organism with foreign DNA is possible (Tyurin, et al., 2004). For these reasons, the cellulosome of *C. thermocellum* has been extensively studied both biochemically (Bayer, et al., 1985; Choi & Ljungdahl, 1996; Mayer, et al., 1987), and via proteomic analyses (Gold & Martin, 2007; Raman, et al., 2009; Zverlov, et al., 2005). The main component, CipA, is a large 200 kDa scaffold protein that brings enzymes into close proximity with one another (Kruus, et al., 1995) (Fig. 1). CipA contains a CBD that binds the cellulosome to the cellulose fibers, bringing the complex in close proximity to the substrate upon which the different cellulases bound to CipA can act. CipA also contains nine type I cohesin domains, which interact with C-terminal type I dockerin domains found on cellulosomal enzymes (Fig. 1). CipA also contains a C-terminal type 2 dockerin domain, which binds non-covalently to type 2 cohesin domains found on cell wall anchor proteins OlpB and SdbA (Leibovitz & Beguin, 1996; Lemaire, et al., 1995). These anchor proteins remain associated with the cell surface by means of S-layer homologous (SLH) domains (Adams, et al., 2006).

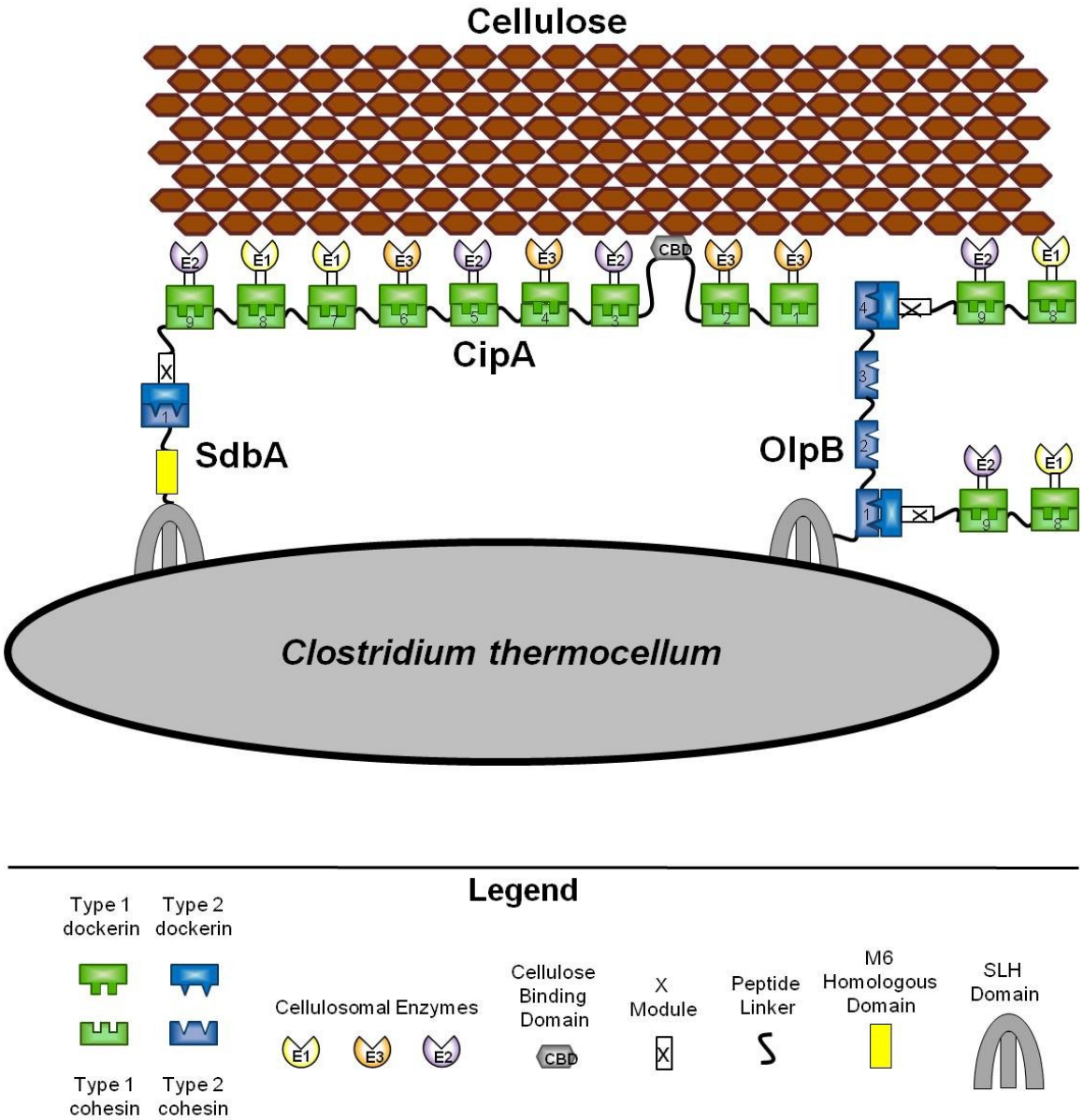

Figure 1. The cellulosome of *C. thermocellum*. Scaffold protein CipA contains nine type 1 cohesin domains which interact with type 1 dockerin domains located at the C-terminus of cellulosomal enzymes. CipA also includes a CBD, which binds with cellulose fibers, bringing the cells in close contact with the crystalline substrate. A C-terminal X module and type 2 dockerin domain are responsible for CipA binding with type 2 cohesins on surface anchor proteins OlpB and SdbA, which in turn carry SLH domains capable of interacting with the cell surface. Peptide linkers separate cohesin domains, and SdbA contains a domain showing high degrees of homology with Streptococcal M proteins located just downstream of its unique cohesin domain.

With little to no difference between the binding specificity of a type I dockerin with any of the type I cohesins found on CipA (Lytle, et al., 1996), it is still somewhat unclear as to how the cellulosome may or may not ensure optimal positioning of the enzymes within the complex.

1.3.2 Anchored cellulosomes of other anaerobic bacteria

The cellulosome of *C. thermocellum* has been well characterized, however insights into the cellulosome architectures of other clostridia reveal that this design is a relatively simple one. For example, *Ruminococcus flavefaciens* produces a cellulosome comprising three divergent cohesin-dockerin pairs mediating the association of scaffolds ScaA, ScaB, ScaX, and multiple enzymes (Rincon, et al., 2003). The cellulosome of *Acetivibrio cellulolyticus* is also more complex than those of the Clostridium genera (Fig. 2). ScaC plays a role similar to OlpB in *C. thermocellum* and serves as an anchor protein that mediates attachment of the cellulosome with the cell surface by means of SLH domains (Q. Xu, et al., 2003). A specific dockerin-cohesin interaction mediates the attachment of adapter scaffold ScaB with ScaC, while a second dockerin-cohesin pair mediates attachment of ScaA with ScaB. ScaA contains a CBD as well as 7 cohesins with a different specificity than the ScaB and ScaC counterparts, capable of binding the dockerin domains located on the multiple hydrolytic enzymes. Interestingly, this architecture also reveals that ScaA itself bears catalytic activity characteristic of a family 9A glycohydrolase (Q. Xu, et al., 2003) (Fig. 2). Therefore, the architectural variability among natural cellulosomes, combined with an ultimate goal of developing a CBP- capable organism,

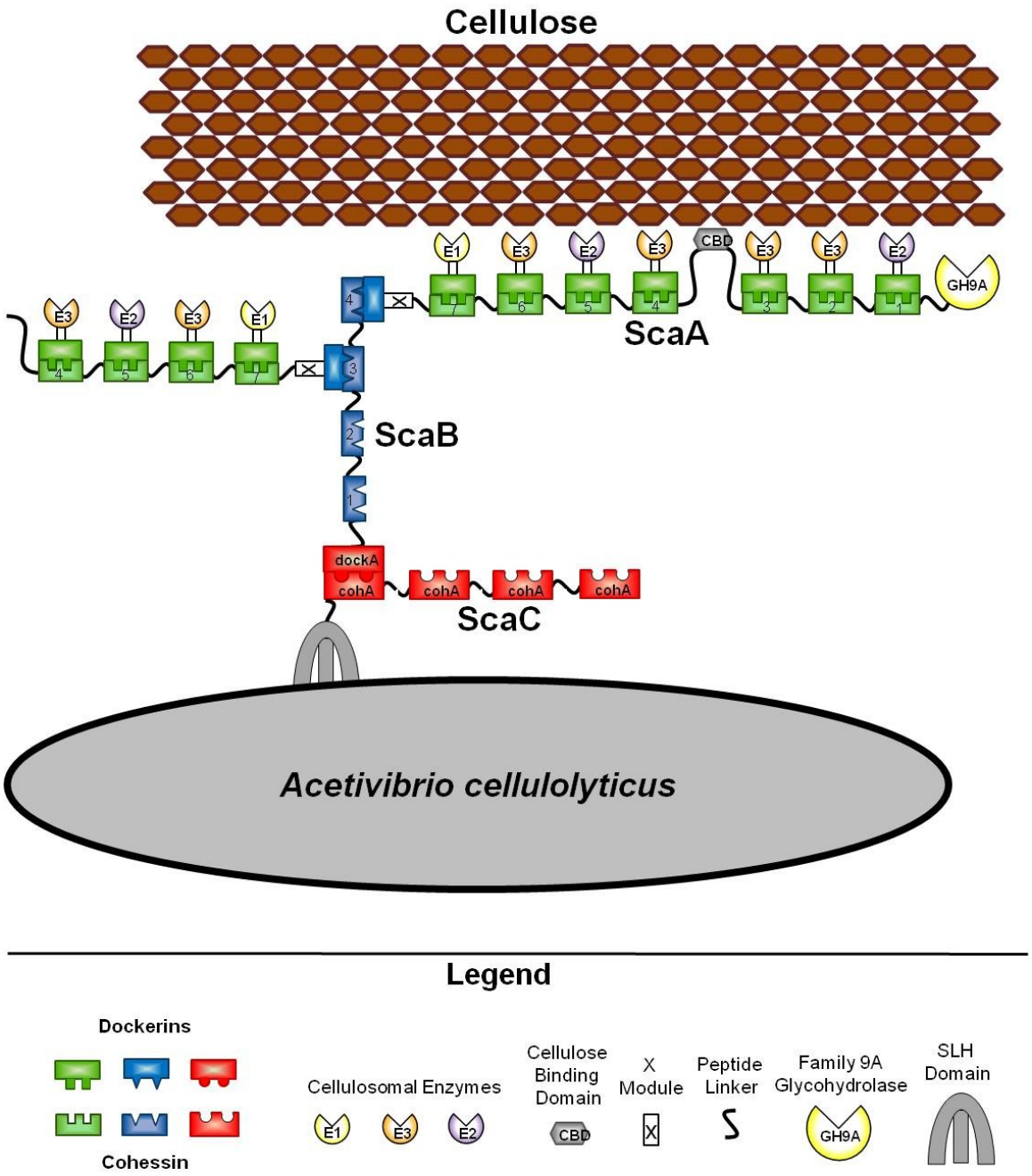

Figure 2. The cellulosome of *A. cellulolyticus*. Scaffold protein ScaA contains seven type 1 cohesin domains which interact with type 1 dockerin domains located at the C-terminus of cellulases. ScaA also carries a CBD that binds with cellulose fibers, bringing the cells in close contact with the crystalline substrate, as well as a family 9A glycohydrolase at its N-terminus. A C-terminal X module and dockerin domain are responsible for ScaA binding with one of four cohesins located on ScaB. ScaB contains a C-terminal dockerin domain, which interacts with one of four cohesin domains located on anchor protein ScaC, which also carries SLH domains capable of interacting with the cell surface.

has yielded extensive research into designing and expressing recombinant cellulosomes in the last decade (Fierobe, et al., 2002; Fierobe, et al., 2001; Fierobe, et al., 2005). Such variability in the architectural design of cellulosomes from different organisms has been a significant source of inspiration for the engineering of protein scaffolds and multi-enzyme complexes (Fierobe, et al., 2005; Mingardon, et al., 2007; Perret, et al., 2004; Wen, et al., 2010).

1.4 Recombinant cellulosomes

Combining enzymes of a biochemical pathway in a larger multi-enzyme complex can yield several benefits such as substrate channeling (Conrado, et al., 2008) as well as synergy among neighboring enzymes (Bayer, et al., 2004). In the case of cellulose hydrolysis, substrate channeling is exemplified by longer chain polysaccharides produced by non-processive cellulases becoming the substrate for processive cellulases, which can produce short chain cellodextrins and cellobiose as primary products (Schwarz, 2001). Therefore, synergy results when such enzymes are localized in close proximity to one another (Bayer, et al., 2004). From a biotechnological perspective, optimizing the spatial organization of pathway enzymes through co-localization on protein scaffolds or fusions has the potential to greatly enhance the channeling of hydrolysis intermediates to enzymes which will use them as substrates in further reactions (Conrado, et al., 2008). Research groups have therefore sought to design recombinant cellulosomes and investigate the effects of enzyme composition and spatial organization on the hydrolytic activity of the resulting complexes. Cohesins and dockerins with different specificities

originating from different species have been used as building blocks to engineer custom-designed recombinant cellulosomes or cellulosome-inspired complexes with defined architectures and enzyme compositions (Cho, et al., 2004; Fierobe, et al., 2005; Mingardon, et al., 2007; Wen, et al., 2010). The strategies adopted to produce recombinant cellulosomes can be divided into three categories: (i) the production of enzymes and scaffolds in engineered host strains followed by their subsequent purification and assembly *in vitro* (Figure 3A), (ii) the production of all components in a single strain resulting in the *in vivo* assembly of resulting complexes in the culture supernatant (Figure 3B), and (iii) the surface-tethering of scaffolds for the *in vivo* assembly of artificial cellulosomes on the cell surface of the host organism (Figure 3C). In Table 1, successfully generated recombinant cellulosome components are listed according to host organism and assembly strategy.

1.4.1 *In vitro* assembly of recombinant cellulosomes

The initial efforts in assembling recombinant cellulosomes involved the production and purification of individual proteins in *E. coli*, followed by their assembly *in vitro* and enzymatic characterization of the resulting complex. Desirable characteristics for a bacteria designed to overexpress individual components include ease of manipulation of the organism, and low endogenous proteolytic activity. The *in vitro* assembly of recombinant cellulosomes involves the production of individual components and subsequent assembly of the complex being carried out in separate steps. Therefore, optimization of culture conditions, production, and purification of each component, can

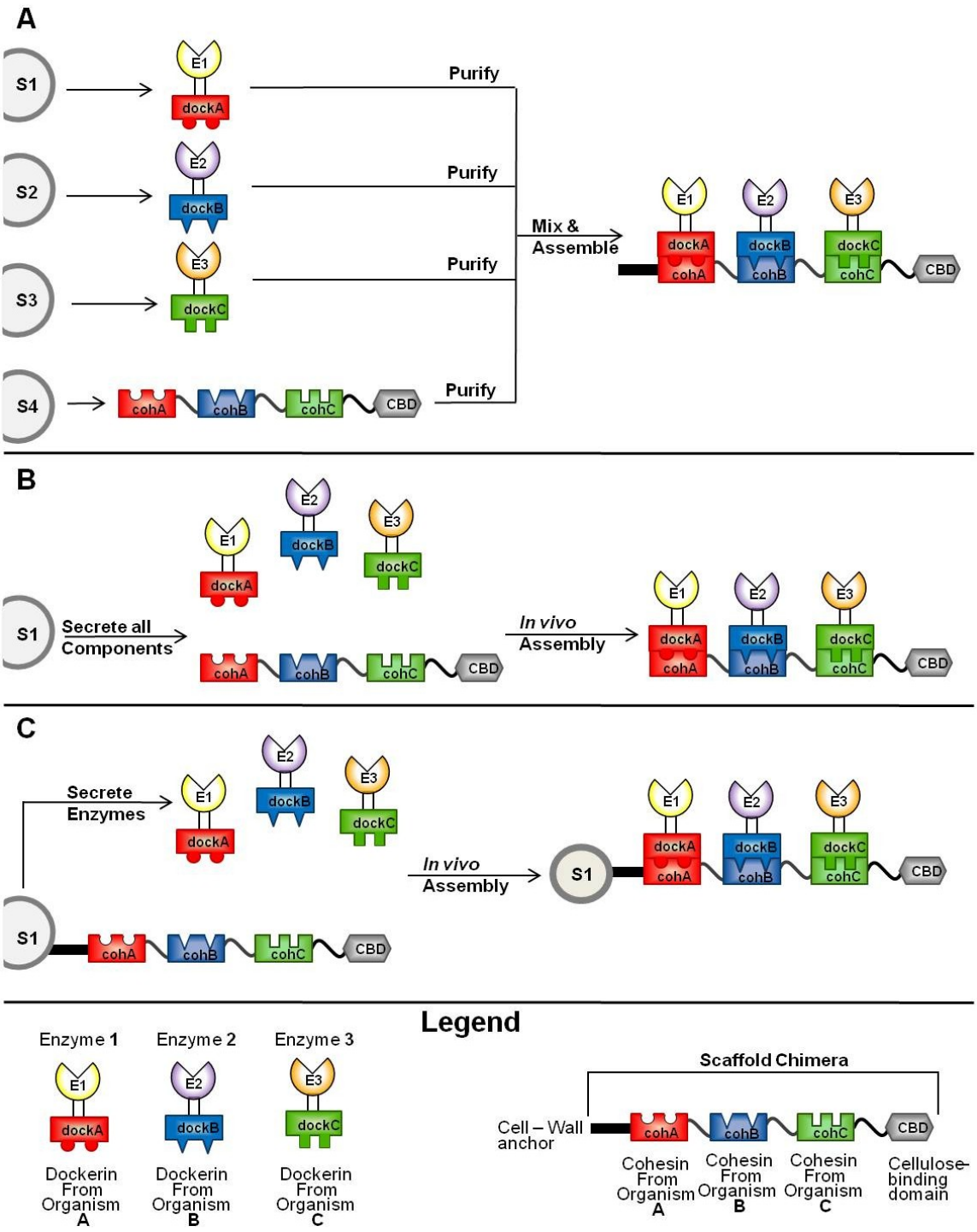

Figure 3. Strategies for the assembly of artificial cellulosome complexes. (A) Enzymes-dockerin fusions and scaffold chimeras are produced by different strains of a host organism (e.g. S1, S2, S3, S4), purified, and subsequently assembled *in vitro*. (B) Enzymes and scaffold subunits are secreted by a single host organism into the culture supernatant where they self-assemble into cellulosomes *in vivo*. (C) A host organism tethers a scaffold protein to its surface while secreting recombinant enzyme-dockerin fusions, resulting in the *in vivo* assembly of the cellulosome complex on the cell surface.

TABLE - 1 Summary of organisms engineered to assembly recombinant cellulosomes, methods of production, and lists of scaffolds and enzymes assembled into multi-enzyme complexes.

Strategy for cellulosome production	Organism (s)	# Divergent Cohesins on scaffold*	# Enzymes simultaneously In complex*	Scaffolds** (cohesin specificty)	Enzymes or Proteins Targeted to Scaffold(s)	Reference
In vitro assembly	*E. coli*	1	2	CipA (coh_{th})	CelD	(Kataeva, et al., 1997)
		1	2	CipA (coh_{th})	CelE	(Ciruela, et al., 1998)
		1	2	Mini-CbpA (coh_{cv})	CelE / CelH / CelS	(Murashima, Kosugi, et al., 2002)
		2	2	Scaf1-4 (coh_{th} / coh_{cl})	CelA / CelF	(Fierobe, et al., 2001)
		2	2	Scaf1-5 (coh_{th} / coh_{cl})	CelA / CelC / CelE CelF / CelG	(Fierobe, et al., 2002)
		3	3	Scaf6 (coh_{th} / coh_{cl} / coh_{rf})	CelA / CelC / CelE CelF / CelG	(Fierobe, et al., 2005)
		3	3	Scaf3, Scaf6 (coh_{th} / coh_{cl} / coh_{rf})	CelF / CelG	(Mingardon, et al., 2007)
		3	3	ScafATF (coh_{ac} / coh_{th} / coh_{rf})	Cel5 / Xyn10 / Xyn11	(Morais, et al., 2010b)
		4	4	Scaf-BTFA (coh_{ac} / coh_{th} / coh_{rf} / coh_{bc})	Cel5 / Cel45 / Xyn10 / Xyn11	(Morais, et al., 2010a)
	B. subtilis / E. coli	1	1	Mini-CbpA (coh_{cv})	EngB	(Murashima, Chen, et al., 2002)
In vivo assembly (secreted)	*B. subtilis*	1	1	Mini-CbpA (coh_{cv})	EngB	(Cho, et al., 2004)
		1	1	Mini-CbpA (coh_{cv})	XynB / EngB	(Arai, et al., 2007)
	C. Aceto-butylicum	1	1	Mini-CipA (coh_{ca})	Cel48A	(Sabathe & Soucaille, 2003)
		1	1	Scaf3 (coh_{cl} / coh_{th})	Cel48F / Cel9E	(Perret, et al., 2004)
In vivo assembly (anchored)	*S. cerev-isiae*	1	3	CipA3 (coh_{th})	EGII / CBHII / BGLI	(Tsai, et al., 2009)
		2	1	Scaf3p (coh_{cl} / coh_{th})	GFP / Cel5A	(Lilly, et al., 2009)
		2	4	ZZ-cohcoh (Z domain / coh_{cv})	EGII / BGLI	(Ito, et al., 2009)
		3	3	Scaf-ctf (coh_{th} / coh_{cl} / coh_{rf})	CelE / CelA / CelG	(Tsai, et al., 2009)
		3	3	Scaf-ctf (coh_{th} / coh_{cl} / coh_{rf})	CelE / CelA / CBHII / Bgl1	(Tsai, et al., 2010)
	L. lactis	1	2	$CipA_{frags}$ (coh_{th})	UidA	(Wieczorek & Martin, 2010)
	B. subtilis	3	3	Scaf (coh_{th} / coh_{cl} / coh_{rf})	Cel8A / Cel9E / Cel9G	(Anderson, et al., 2011) (You, et al., 2012)
		1	3	Mini-CipA (coh_{th})	Cel5 / Cel9 / Cel48	

*Correspond to complexes engineered to contain largest number of divergent cohesins and integrated enzymes
**Scaffolds listed are those containing the largest number of cohesin domains from that study. Names in parenthesis correspond to types of cohesins included in the most complex scaffolds. Coh: cohesin domain. Subscript indicates organism of origin: th (*C. thermocellum*), cv (*C. cellulovorans*), cl (*C. cellulolyticum*), rf (*R. flavefaciens*), ac (*A. cellulolyticum*), bc (*B. cellulosolvens*). Z domain: *S. aureus* Protein A binding domain.

theoretically be carried out in different organisms of choice (Fig. 3A).

1.4.1.1 Expression of cellulosomal proteins in *E. coli*

Early work on the assembly of cellulosomes *in vitro* focused on how a scaffold-bound enzyme had increased activity on cellulose compared with the isolated enzyme. In a study by Kataeva and coworkers, EGL CelD was shown to bind stoichiometrically with fragments of the CipA proteins, and CelD-CipA complexes demonstrated increased activity on cellulose compared with enzyme alone (Kataeva, et al., 1997). A major observation was that the activity of the complex was dependent on the presence of a CBD, not necessarily the amount of CelD present. The authors hypothesized that the CBD located on scaffold fragments was either indirectly contributing to the hydrolysis process by optimally positioning CelD to act on the crystalline substrate, or that it was playing a more direct role, participating in the partial decomposition of the substrate and ultimately, allowing access to CelD (Kataeva, et al., 1997). A subsequent study by Ciruela and colleagues revealed that the binding of another EGL, CelE, with full length CipA, resulted in artificial cellulosomes with increased activity on crystalline cellulose compared to free enzymes (Ciruela, et al., 1998). Interestingly, although the CBD of CipA was capable of binding both crystalline and amorphous cellulose, the increase in activity observed when CelE was complexed with CipA was only observed on the former, suggesting the pivotal role of the scaffold-enzyme complex in degrading the crystalline substrate. Both studies conducted by Kataeva and Ciruela involved the incorporation of a single enzyme into artificial cellulosomes. A study by Murashima and coworkers involved

the production of a truncated version of the *Clostridium cellulovorans* scaffold protein CbpA (Mini-CbpA) as well as three enzymes, EngE, EngH, and EngS, and their subsequent assembly *in vitro* into artificial cellulosomes containing combinations of two enzymes (Murashima, Kosugi, et al., 2002). Synergistic activity of the enzymes on crystalline cellulose was affected by both the type and stoichiometric ratios of enzyme used. Optimal combinations of enzymes were determined based on increased activity on crystalline cellulose. In this case, however, the effects of enzyme positioning within the complex could not be deduced due to the non-specific binding of each enzyme to any of the two cohesins present on the scaffold. The multiple enzyme activities required to degrade crystalline cellulose and the possibility to optimize the stoichiometry and the relative positioning of the enzymes within the complex prompted the construction of recombinant cellulosomes with precise enzymatic compositions.

The construction of artificial scaffold proteins containing cohesins of different specificities was used to precisely dictate the enzyme composition of designer cellulosomes (Fierobe, et al., 2002; Fierobe, et al., 2001; Fierobe, et al., 2005; Mingardon, et al., 2007). Initial work describing the construction and utility of scaffold chimeras was carried out by Fierobe and coworkers, where the fusion of cohesins derived from the cellulosomes of *C. thermocellum* and *Clostridium cellulolyticum* allowed complexes with two enzymes to be generated (Fierobe, et al., 2001). The authors engineered a total of four scaffolds, which contained two divergent cohesins positioned at various locations relative to the CBD, which was absent in one construct. Two *C. cellulolyticum* cellulases, CelA and CelF, were engineered to contain either native or *C. thermocellum* dockerins. All

components were over-produced in *E. coli,* purified and assembled *in vitro* into tri-modular cellulosomes. The authors once again demonstrated the necessity of the CBD for increased hydrolysis of the cellulose substrate, and observed that the sequential or simultaneous assembly of each component yielded similar activities. Increased activity was observed when enzymes were assembled onto the chimeric scaffold as opposed to free enzyme mixtures, suggesting synergistic effects, which may include substrate channeling. In a subsequent study, Fierobe and colleagues successfully generated a library of 75 different chimeric cellulosomes (Fierobe, et al., 2002) and tested their activities on both crystalline and amorphous substrates. The enzymes incorporated into the complexes consisted of pairwise combination of *C. cellulolyticum* cellulases CelA, CelC, CelE, CelF, or CelG. Synergy due to enzyme assembly on the chimeric scaffolds was only observed when acting on the more recalcitrant crytalline substrates Avicel and bacterial microcrystalline cellulose (BMC), with less or no synergy observed when acting on the less crystalline substrates such as bacterial cellulose and phosphoric acid-swollen cellulose (PASC). Interestingly, a complex containing a chimeric scaffold that contained two CBDs showed less activity on the highly crystalline substrates than a complex containing a scaffold with a single CBD, probably due to a more restricted conformation of the scaffold and enzymes within the complex. The factors deemed responsible for the enhanced activity of bifunctional complexes included targeting of the complex to the cellulose substrate as well as the complementing activities of both enzymes. Therefore, to further amplify the synergistic and overall activities of such artificial cellulosomes, Fierobe and coworkers generated trifunctional cellulosomes, which contained three

different cellulases as well as a CBD (Fierobe, et al., 2005). In order to precisely position the desired enzymes within the complexes, a third dockerin-cohesin pair derived from *R. flavefaciens* was used in which the interaction is characterized by both high affinity and lack of cross-reactivity with the other cohesin-dockerin pairs of clostridial origin. Upon incorporation of three cellulases, the complexes demonstrated significantly higher activity than their bifunctional counterparts. The synergy among enzymes was also demonstrated when assembled onto the chimeric scaffold versus free enzyme mixtures.

The work described above reported on the construction of artificial cellulosomes that mimic natural systems where all enzymes assemble onto a single scaffold protein, which in turn mediate the attachment of the complex to the cellulose substrate by means of a CBD. To generate artificial cellulosomes with novel geometries and potentially higher overall activities on cellulose, Mingardon and coworkers constructed chimeric scaffolds and cellulases designed to self-assemble in a precise spatial arrangement (Mingardon, et al., 2007). The strategy investigated several factors that may have an effect on the resulting enzymes or multi-enzyme complexes. First, chimeric GH9- and GH48- family enzymes were constructed by fusing the catalytic modules with a CBD. The activities of the resulting enzymes on crystalline substrates were increased by fusion with CBD, but were lower on amorphous substrates. PASC contains more reactive sites than crystalline cellulose and it is hypothesized that incorporation of the CBD reduced the overall mobility of the enzyme on that substrate, whereas the necessity of the CBD for Avicel degradation was demonstrated by increased activity of the chimeric enzymes compared with parental counterparts. A “hybrid cellulosome” was engineered to contain two

enzymes on a central chimeric scaffold, while a “covalent cellulosome” was generated by covalently linking them in a single polypeptide chain (Mingardon, et al., 2007). Although the covalent cellulosome did demonstrate increased activity on Avicel compared with free enzymes, probably due to enzyme proximity and synergy, it demonstrated lower activity than the hybrid cellulosome, which more closely resembles the architecture of natural clostridial cellulosomes. The restricted mobility of the individual enzymatic units within the covalent cellulosome likely accounted for its inferior catalytic activity. Inclusion of a third cohesin and dockerin pair derived from *R. flavefaciens* resulted in the generation of novel geometries. “Cyclic”, “polymeric”, and “mixed” cellulosomes were generated by fusing enzymes and scaffolds together, resulting in geometries dictated by interactions of corresponding dockerin and cohesin domains, however all of these conformations resulted in reduced activity on Avicel when compared with the hybrid cellulosome, which closely mimicks the structure of natural clostridial cellulosomes. For detailed description of each of these geometries, the reader is encouraged to refer to the original publication (Mingardon, et al., 2007). Some other notable observations were that the least effective cellulosome contained the most CBDs and that in certain architectures, cohesin-dockerin pairs could dissociate, most probably due to conformational strain.

Cellulosic biomass is not naturally found in the form of pure cellulose, but consists of lignocellulosic material which includes other complex polymers such as xylans (Thomson, 1993). In an effort to bestow xylanase activity upon engineered cellulosomes, Morais and colleagues investigated the possibility of targeting two xylanases as well as a xylose binding domain (XBD) to a scaffold containing three divergent cohesins from *A.*

cellulolyticus, *C. thermocellum*, and *R. flavefaciens* (Morais, et al., 2010b). This was part of a larger effort to convert multiple free cellulases and hemicellulases from the cellulolytic fungi *Thermobifida fusca* into cellulosome-integrated enzymes. The assembled complexes were characterized by a 1.5 fold increase in activity on the complex substrate hatched wheat straw when compared with free enzyme mixtures, and the authors attributed this to substrate targeting by the XBD as well as to the proximity of the enzymes within the complex (Morais, et al., 2010b). These results were further improved upon in a subsequent study whereby an artificial cellulosome was generated to accommodate a total of four enzymes (two EGLs and two xylanases) (Morais, et al., 2010a). This was achieved by the addition of another dockerin-cohesin pair derived from *Bacteriodes cellulosolvens* resulting in synthetic scaffold ScafBTFA. An overall 2.4-fold increase in activity on hatched wheat straw was observed compared with the free enzyme mixtures due to the proximity of the enzymes on a single scaffold.

1.4.1.2 Expression of cellulosome components in *B. subtilis*

While *E. coli* remains an attractive host for the overexpression, production, and purification of enzymes and scaffolds due to its ease of manipulation, the presence of endogenous proteases can contribute to the degradation of the recombinant proteins (Murashima, Chen, et al., 2002). Another attractive host for the production of recombinant cellulosomes is *B. subtilis*, since it can be easily genetically manipulated, is characterized by fast growth, and is an efficient protein secretor (Wong, 1995). A strain of *B. subtilis* deficient in eight major extracellular proteases, *B. subtilis* WB800, was

engineered and used as a host for the production and secretion of *C. cellulovorans* EngE since this enzyme was shown to be partially degraded in *E. coli* (Murashima, Chen, et al., 2002). Murashima and colleagues were successful in using this protease-deficient strain to produce EngE, and subsequent incubation with scaffold Mini-CbpA, which contains a CBD as well as two cohesins, resulted in assembly of an enzyme-scaffold complex capable of binding cellulose (Murashima, Chen, et al., 2002).

1.4.2 *In vivo* secretion and assembly of recombinant cellulosomes

The overexpression and purification of individual scaffolds and enzymes for the assembly of artificial cellulosomes poses extra costs and steps towards cellulose hydrolysis. Rather, the development of a CBP-capable organism would require the production, secretion and *in vivo* assembly of artificial cellulosomes in the extracellular space (Fig. 3B).

1.4.2.1 Secretion of recombinant cellulosomes by *B. subtilis*

As an extension of Murashima and colleagues' work, Cho and colleagues constructed an expression cassette encoding both Mini-CbpA and EngE on a single vector, which was established in *B. subtilis* WB800 (Cho, et al., 2004). This resulted in the secretion and subsequent assembly of both enzyme and scaffold components into an artificial cellulosome complex that localized to the culture supernatant. This study was the first to report on the *in vivo* assembly of artificial cellulosomes by a single organism, whereby the cellulosome demonstrated activity on CMC but not the crystalline substrate

Avicel. Similarly, Arai and colleagues used three strains of *B. subtilis* WB800 that were engineered to secrete either EngB, XynB, or MiniCbpA (Arai, et al., 2007). By co-culturing enzyme and scaffold producing strains, complexes formed in the supernatant and were characterized by the appropriate enzymatic activity. This provided a novel method for assembling protein complexes *in vivo* based on intercellular complementation.

1.4.2.2 Secretion of recombinant cellulosomes by *C. acetobutylicum*

C. acetobutylicum is an organism that is used in the production of acids and solvents including acetone, butanol, and ethanol (Mitchell, 1998). The potential to engineer this organism to degrade cellulose as a cheap and abundant carbon source has garnered significant attention in the past decade. This bacterium is not cellulolytic, however its genome sequence revealed a cellulosomal gene cluster encoding a number of hydrolytic enzymes as well as a scaffold protein CipA (Nolling, et al., 2001; Sabathe, et al., 2002). Sabathe and colleagues were successful in engineering *C. acetobutylicum* to secrete and assemble a functional minicellulosome *in vivo* (Sabathe & Soucaille, 2003). Since this organism does not secrete CipA, the authors replaced the original signal peptide with that of the *C. cellulolyticum* scaffold protein CipC. Overexpression and secretion of a truncated version of CipA containing two cohesin domains and a CBD resulted in its binding with endogenous cellulase Cel48A, and formation of a cellulosome *in vivo* (Sabathe & Soucaille, 2003). In analyzing the activity of the recombinant cellulosome on Avicel, bacterial cellulose, PASC and carboxymethyl cellulose (CMC), no detectable activity was observed when using the crystalline substrates, as is the case for

native *C. acetobutylicum*. Low levels of activity were observed on CMC and PASC, however these levels did not exceed those of the wild-type cellulosome. Perret and colleagues engineered this organism to secrete artificial scaffold chimera Scaf3 (Perret, et al., 2004). The Scaf3 consisted of miniCipC1, which is a truncated form of *C. cellulolyticum* scaffold CipC, and an additional cohesin from *C. thermocellum* scaffold CipA. After visualizing the chimeric scaffold using SDS-PAGE, the protein was blotted on a nitrocellulose membrane and subsequently shown to bind both Cel48 and Cel9 containing a dockerin from *C. cellulolyticum*, as well as Cel9 with a dockerin from *C. thermocellum*.

1.4.3 *In vivo* surface-anchoring of recombinant cellulosomes

Cellulosomes promote synergy of the enzymes within the complex when associated with the substrate (Garcia-Campayo & Beguin, 1997; Kosugi, et al., 2004; Kruus, et al., 1995). In natural and recombinant systems, synergistic effects are further augmented from the cellulosome's association with the surface of the cell, yielding cellulose-enzyme-microbe (CEM) ternary complexes (Bayer, et al., 1983; Lu, et al., 2006; L. R. Lynd, et al., 2005; L. R. Lynd, et al., 2002; Miron, et al., 2001; Ng, et al., 1977; Schwarz, 2001; Zverlov, et al., 2008). CEM ternary complexes are thought to be beneficial in many ways such as limiting the escape of hydrolysis products and enzymes, increasing access to substrate hydrolysis products, minimizing the distance products must diffuse before cellular uptake occurs, concentrating enzymes at the substrate surface, protecting hydrolytic enzymes from proteases and thermal degradation, as well as optimizing the

chemical environment at the substrate-microbe interface (Lu, et al., 2006; L. R. Lynd, et al., 2005; L. R. Lynd, et al., 2002; Miron, et al., 2001; Schwarz, 2001; Zverlov, et al., 2008). In several cellulosome-producing bacteria, including *C. thermocellum*, the cellulosome is anchored to the surface of cells, resulting in one of the most efficient examples of bacterial cellulose hydrolysis (Bayer, et al., 2004; Schwarz, 2001). In an effort to mimic this approach, microbial engineers have adopted this strategy as a next logical step towards the improvement of recombinant cellulosomes.

1.4.3.1 Anchoring recombinant cellulosomes on the cell surface of *S. cerevisiae*

As mentioned previously, much interest towards the development of a CBP-capable organism comes from a desire to generate biofuels such as ethanol from cellulosic substrates. Therefore, significant attention has been directed at endowing cellulolytic capacity to *S. cerevisiae*. Lilly and colleagues were successful in targeting hybrid scaffold Scaf3p to the cell surface of *S. cerevisiae* by fusing it with the glycosyl phosphatidylinositol (GPI) signal peptide of the Cwp2 protein for linking to the β-1,6 glucan of the yeast cell wall (Lilly, et al., 2009). The scaffold contained two divergent cohesins from *C. thermocellum* and *C. cellulolyticum* as well as a CBD. Microsocopy revealed that the CBD was functional in adhering cells to filter paper, and the successful targeting of a Cel5a-dockerin fusion protein to the scaffold confirmed functionality of the cohesin domains. The ability to generate scaffold chimeras using non-cohesin domains was established by Ito and colleagues (Ito, et al., 2009). They generated artificial scaffolds by fusing the Z domain of *Staphylococcus aureus* Protein A with a cohesin from the *C.*

cellulovorans cellulosome and displayed them on the cell surface (Ito, et al., 2009). The scaffold chimeras were engineered to contain two Z domains as well as two cohesins for targeting different enzymes to the cell surface. The authors fused two enzymes, EGII and BGLI, to either a dockerin domain or Fc domain, which successfully bound the enzymes to the cohesin and Z domains, respectively (Ito, et al., 2009). Hydrolysis experiments revealed that co-displaying EGII-FC and BGL-dock fusion protein resulted in cells capable of degrading β-glucan. However, in the absence of a CBD on the engineered scaffold, this strain would most likely be inefficient at hydrolyzing more recalcitrant crystalline cellulosic substrates. A different approach to ethanol production was adopted by Tsai and coworkers, where yeast strains were engineered to display a scaffold containing three divergent cohesins from *C. thermocellum*, *C. cellulolyticum* and *R. flavefaciens* as well as a CBD (Tsai, et al., 2009). Three enzymes, *C. thermocellum* CelA, and *C. cellulolyticum* CelE and CelG were overproduced in *E. coli* and successfully targeted to corresponding cohesin domains on the scaffold by fusion with appropriate dockerin domains, resulting in the surface-display of trifunctional cellulosomes. The anchor used in this study consisted of displaying the Aga2 protein, which interacted with the Aga1 protein fused with the scaffold. Replacing endoglucanase CelG with *C. thermocellum* β-glucosidase BglA resulted in significant increases in glucose liberation from PASC, and the resulting strain was capable of directly producing ethanol from this substrate. Incubating cells in the presence of PASC resulted in ethanol production that corresponds to 95% of the theoretically attainable ethanol yield. The authors also observed no accumulation of glucose in the medium during the fermentation assays, suggesting that the released

glucose was immediately taken up by cells during the simultaneous saccharification and fermentation (SSF) process (Tsai, et al., 2009).

The production of both enzymes and scaffold in a single yeast strain was achieved by Wen and colleagues (Wen, et al., 2010). The scaffold contained three cohesins as well as a CBD and was successfully displayed by use of the α-agglutinin adhesion receptor. *In vivo* secretion of an EGL, CBH, and BGL resulted in the assembly of tetrameric complexes (one scaffold and three enzymes) and the resulting yeast strain was capable of directly converting PASC to ethanol at a yield of 1.8 g/L. Interestingly, the authors also observed that when Bgl1 was positioned within the complex, in close proximity to EGII and CBHII, increased degradation of PASC was achieved, most probably due to removal of the cellobiose at the cell surface which may have been inhibiting EGII and CBHII. In comparison with the work by Tsai and colleagues, this represented the first report of producing and assembling a trifunctional cellulosome on the cell surface by the *in vivo* production of all components. The relatively low levels of EGII and Bgl1 produced by this strain, however, suggested that burdening the secretion machinery of the organism was a potential bottleneck. To address this issue, the Chen group adopted a different approach, which entailed intercellular complementation by a yeast consortium (Tsai, et al., 2010). In this case, one strain produced a scaffold containing three divergent cohesins and a CBD, while each of three other strains produced an exoglucanase, EGL, or BGL, which were targeted to specific sites on the artificial scaffold by fusion with corresponding dockerin domains. The authors also reported that an optimal ratio of each

strain within the consortium resulted in two-fold increase in ethanol production when compared with a consortium containing equal proportions of each strain.

1.4.3.2 Anchoring recombinant cellulosomes on the cell surface of *L. lactis*

While engineering cellulosomes in microbes has focused mostly on ethanol-producing microbes such as *S. cerevisiae*, microorganisms producing other biofuel molecules or commodity chemicals would also benefit from this capacity. Wieczorek and Martin engineered a strain of *L. lactis* to anchor mini-scaffolds on the cell surface (Wieczorek & Martin, 2010). While several bacterial species non-covalently anchor cellulosomes to the cell surface by means of S-layer homologous domains, other organisms such as *R. flavefaciens* display cellulosomes by covalently anchoring them to the cell wall by sortase (Rincon, et al., 2005). The authors in this study fused fragments of *C. thermocellum* CipA scaffold with a C-terminal LPXTG-containing anchor motif from *Streptococcus pyogenes* M6 protein, resulting in their successful surface-display. By fusing the scaffolds with the export-specific reporter, *S. aureus* nuclease NucA, the authors were able to easily detect them in the extracellular medium. Fusion of *E. coli* β-glucuronidase UidA with the dockerin from major *C. thermocellum* cellulosomal enzyme CelS, resulted in its successful targeting to the surface-displayed scaffolds. While the assembled complexes were not cellulolytic, the investigation yielded insights into parameters affecting secretion and anchoring of the recombinant scaffolds, including the observation that scaffold size was not a significant bottleneck in display efficiency. The strain used was deficient in its major extracellular housekeeping protease HtrA, which

was demonstrated to be responsible for the degradation of secreted recombinant proteins (Miyoshi, et al., 2002). In a subsequent study, the authors fused type 1 and type 2 cohesins to generate scaffold chimeras capable of binding UidA and *E. coli* β-galactosidase LacZ fused with type 1 and type 2 dockerins (chapter 3). This yielded novel insights into the assembly of displayed complexes, suggesting that enzyme size and position relative to the cell surface may play a role in determining the overall net enzymatic profile of the displayed complexes.

1.4.3.3 Anchoring recombinant cellulosomes on the cell surface of *B. subtilis*

Cellulosomes have also been engineered in *B. subtilis* (Arai, et al., 2007). The attractiveness of this host is compounded by several characteristics including its ability to metabolize C_5 and C_6 sugars as well as its natural ability to uptake long-chain cellodextrins (You, et al., 2012). Anderson and colleagues used a strategy similar to the Martin group's by employing the sortase-mediated anchoring of proteins on the cell surface (Anderson, et al., 2011). This group initially demonstrated proof of concept by displaying a single enzyme, Cel8A, and subsequently went on to display cohesin domains capable of interacting with an appropriate Cel8A-dockerin fusion. It was observed that proteolytic degradation of the displayed enzymes resulted in an 80% decrease in activity after only 6 hrs, an effect hypothesized to result from the presence of the extracellular housekeeping protease WprA. Inserting this system into a $WprA^-$ strain resulted in a significant reduction in the observed proteolysis of the enzymes. The most complex artificial cellulosome generated by this group included a surface-anchored chimeric scaffold

containing three divergent cohesins and a CBD. Incubation of cells with enzyme-dockerin fusions purified from *E. coli* resulted in the assembly of functional minicellulosomes on the cell surface. Soon afterwards, the Zhang group reported the engineering of a scaffold-displaying *B. subtilis* strain capable of binding three enzymes and the subsequent assembly of an artificial cellulosome on the cell surface (You, et al., 2012). These authors investigated the effect of the CEM ternary complex by comparing a cell-bound artificial cellulosome, a cell-free artificial cellulosome, and a commercial fungal cellulose mixture. When comparing the activity of cell-bound cellulosomes vs. cell-free cellulosomes, a larger significant increase in CEM synergy on Avicel as opposed to amorphous cellulose was observed in the cell-displayed constructs. The authors suggested this effect was due to larger product inhibition at the boundary layer when active on crystalline cellulose. EGLs demonstrate higher hydrolysis activity on amorphous cellulose while CBHs are more sensitive to product inhibition (Liao, et al., 2011), therefore the observed results demontrated the benefits of anchoring cellulosomes on the cell surface.

1.4.4 Future prospectives for cellulosome production

The generation of "custom-designed" cellulosomes with optimized ratios and relative positioning of enzymes has thus resulted in recombinant protein complexes characterized by hydrolytic capacities where the overall activities are higher than the sum of their parts. Still, significant advances are necessary for the cost-effective transformation of cellulose into valuable commodity chemicals such as bioethanol and lactic acid to become an industrial standard. For example, from a microbiological

perspective, engineering an organism with enhanced secretion and surface anchoring capacity for such complexes may be of significant interest. Indeed, the native metabolic diversity of microbes designed to utilize cellulose as an energy source, as well as the advent of synthetic biology through which non-native and novel pathways can be introduced into these organisms, suggest that the bioconversion of cellulosic substrates into valuable chemicals is not so far from reach. Constructing more efficient recombinant cellulases, as well as the assembly of cellulosomes with complex architectures inspired by bacteria such as *R. flavifaciens* and *A. cellulolyticus*, are other possible avenues to explore in this field.

From a more general perspective, biofuel production and cellulose degradation are only a small subset of bioconversion processes that may benefit from the recombinant display of custom-designed multi-enzyme complexes. The wide range of metabolic diversity among bacteria provides opportunities for the development of cellulosome-inspired complexes in organisms capable of producing other industrially relevant products (Shanmugam & Ingram, 2008). In addition, the variability in biochemical composition of complex substrates to be used for such bioconversion processes would entail a different optimized mixture of hydrolytic enzymes. Therefore, developing a system where researchers could "plug" virtually any enzyme of choice into a specific "socket" on a synthetic scaffold could prove of great value in the future. This would also yield a better understanding of the parameters that affect cellulosome assembly, including enzyme size, position, sequence of enzyme loading, and overall

complex architecture. Therefore, the display of synthetic protein scaffolds on surrogate hosts for purposes other than bioethanol production is highly desirable.

1.5 Heterologous scaffold proteins production

Limitations preventing the easy genetic manipulation of cellulolytic clostridia lower their appeal as candidates for the expression of custom-designed cellulosomes (Tyurin, et al., 2004). In addition, the assembly of multi-enzyme complexes on a host organism has the potential to benefit avenues of research extending beyond biofuels production. The ordered assembly of multiple enzymes on a single scaffold could be of use for the optimization of other bioprocesses requiring the simultaneous or sequential activity of several enzymes (Conrado, et al., 2008). Characteristics that may be of value in choosing an organism to display a multi-enzyme complex include the genetic tractability, low levels of endogenous proteases, high secretion efficiency of heterologous proteins, as well as tools for anchoring scaffolds on the cell surface.

E. coli has been used for the expression and purification of proteins and enzymes, however translocating recombinant proteins out of the cell for display *in vivo* requires their passage through both an inner and outer membrane (Saier, 2006). Gram- positive bacteria are of particular appeal since the secretion and display of recombinant proteins in *B. subtilis* (Wong, 1995) and lactic acid bacteria (LAB) has been achieved with substantial success (Le Loir, et al., 2005). The LAB, which include members of the Lactobacillus and Lactococcus genera, have multitude of tools for protein display, either using covalent or non-covalent mechanisms (Leenhouts, et al., 1999).

1.5.1 *Lactococcus lactis*: A model host organism

Lactococcus lactis is a Gram-positive bacterium most closely related to members of the Streptococcus genus (Bolotin, et al., 2001). It is generally regarded as safe (GRAS), and a variety of molecular techniques are well documented for its manipulation and metabolic engineering (Holo & Nes, 1995; Kleerebezem, et al., 2002). Strains of *L. lactis* have been engineered to secrete and/or surface-display a wide variety of proteins ranging from 9.8 to 165 kDa (Avall-Jaaskelainen, et al., 2003; Bermudez-Humaran, et al., 2003; Cortes-Perez, et al., 2005; Dieye, et al., 2003; Enouf, et al., 2001; Leenhouts, et al., 1999; Lindholm, et al., 2004; Narita, et al., 2006; Piard, et al., 1997; Raha, et al., 2005; Ramasamy, et al., 2006; Ribeiro, et al., 2002). Methods for transforming *L. lactis* have been established (Holo & Nes, 1995), a strain deficient in a major extracellular protease has been engineered (Miyoshi, et al., 2002), and secretion signal peptides have been identified (van Asseldonk, et al., 1990) and used to drive the targeting of recombinant proteins to the extracellular space (Le Loir, et al., 2005). In addition, several strategies for anchoring recombinant proteins to the surface of LAB have been established (Leenhouts, et al., 1999).

1.5.2 Secretion in *L. lactis*

Few proteins are naturally secreted in *L. Lactis* (Bolotin, et al., 2001; Poquet, et al., 1998; van Asseldonk, et al., 1990). Only eight genes are implicated in protein secretion in the organism, and it does not contain the *secDF* gene known to improve

secretion efficiency (Bolotin, et al., 2001). Interest in developing an export-specific reporter for this organism resulted in the use of *Staphylococcal aureus* nuclease (Nuc) with its native signal peptide sp_{Nuc} (Le Loir, et al., 1998). Le Loir and colleagues investigated the secretion efficiency of Nuc and determined that only 70% of the produced enzyme was successfully secreted (Le Loir, et al., 1998). Increased secretion efficiency was achieved by insertion of a nine-residue synthetic peptide (LEISSTCDA) between the signal peptide and Nuc coding sequence (Le Loir, et al., 1998; Le Loir, et al., 2001). Site-directed mutagenesis experiments revealed that the resulting increase in secretion efficiency correlated with the insertion of negatively-charged residues at the N-terminus of the mature protein (Le Loir, et al., 2001). Use of this synthetic peptide also increased the net amount of protein generated, suggesting that secretion provides a means to allow heterologous proteins to escape proteolysis by intracellular proteases. The authors also hypothesized that use of a homologous signal peptide may increase the secretion efficiency and subsequently tested this hypothesis. Previous characterization of a native protein of unknown function, Usp45, led to the discovery that this protein is secreted by *L. lactis* (van Asseldonk, et al., 1990), with the 22 amino acids at its N-terminus acting as a signal peptide to direct secretion. Therefore, Le Loir and colleagues tested the effects of changing the native sp_{Nuc} for the homologous signal peptide of Usp45 (sp_{Usp45}) (Le Loir, et al., 2001). The authors were able to significantly increase the secretion efficiency of Nuc by replacing sp_{Nuc} for sp_{Usp45}. This homologous signal peptide has also been used to drive the secretion of a variety of recombinant proteins in *L. lactis* (Le Loir, et al., 2005). The mechanism of secretion signal processing is based on the sec

pathway common to several Gram-positive bacteria whereby chaperones recognize the signal peptide, and once translocation across the cell envelope is complete, cleave between the alanine and aspartate residues found in this sequence (Dieye, et al., 2001). In one report, use of a signal peptide from *Lactobacillus brevis* to drive secretion of the *E. coli* FedF adhesin resulted in higher secretion efficiency compared with sp_{Usp45} (Lindholm, et al., 2004). Nonetheless, the high secretion efficiency achieved when using sp_{Usp45} therefore makes it an attractive signal peptide for the secretion and display of recombinant proteins and/or enzymes in *L. lactis*.

Secretion of recombinant proteins by microbes, as opposed to their intracellular production, has been used as a general strategy in order to increase yields (Le Loir, et al., 2005). In addition, protein secretion has provided a starting point towards engineering *L. lactis* as a live vaccine (Bermudez-Humaran, et al., 2004; Ribeiro, et al., 2002). Ribeiro and colleagues were successful in not only producing the immunogenic *Brucella abortus* ribosomal protein L7/L12, but also observed a significant increase in protein production when it was targeted to the extracellular medium (Ribeiro, et al., 2002). In another study, Bermudez-Humaran and coworkers expressed the human papilommavirus E7 antigen in larger quantities by secretion when compared with its targeting to the cytoplasmic space (Bermudez-Humaran, et al., 2002). While these studies demonstrated that immunogenic proteins could be produced in *L. lactis*, they also revealed that secretion of such proteins is a means to generate larger quantities of recombinant protein in this bacterium.

1.5.3 Decreasing proteolysis

When secreting proteins, proteases can be of great concern as they can significantly decrease the resulting amount of intact protein found in the extracellular medium. A number of native proteases in *E. coli* and *B. subtilis* have been well characterized (Gottesman, 1996; Simonen & Palva, 1993). In contrast to these bacteria, *L. lactis* has a unique extracellular housekeeping protease HtrA, whose function was verified by the construction of a deletion mutant (Poquet, et al., 2000). It was revealed that this protease participates in propeptide processing and in the degradation of recombinant proteins (Poquet, et al., 2000). Strain *htrA*-NZ9000 is deficient in the HtrA protease, which has been deemed responsible for the degradation of several successfully secreted recombinant proteins (Cortes-Perez, et al., 2006; Miyoshi, et al., 2002). This strain demonstrates decreased proteolysis of secreted proteins, as well as an increase in their overall stability. The proteins tested in one study consisted of the E7 and L7/L12 antigens, the bovine rotavirus antigen NSP4, as well as *Staphyloccus hyicus* lipase (Miyoshi, et al., 2002). In all cases, decreased proteolysis was observed. Interestingly, in a later study, HtrA was shown to be responsible for increased secretion efficiency of extracellularly-targeted proteins (Sriraman & Jayaraman, 2008). Sriraman and Jayaraman suggested that use of an HtrA mutant might not be ideal for achieving the highest amount of protein production since such mutants produce less net recombinant protein when compared with wild-type. HtrA deletion mutants also exhibit increased cell aggregation (Foucaud-Scheunemann & Poquet, 2003), a phenomenon which may in turn reduce exposure of the secretion machinery to the extracellular space, potentially

resulting in decreased secretion efficiency (Sriraman & Jayaraman, 2008). It is therefore of importance to understand that when secreting recombinant proteins in *L. lactis htrA*-NZ9000, a certain trade-off exists between decreasing proteolytic degradation and decreasing secretion efficiency at the same time. Nonetheless, in avenues of research where the primary objective is not necessarily to produce the highest amount of protein, but rather to produce stable ones, the overall decrease in proteolysis of an HtrA mutant is a very appealing characteristic.

1.5.4 Constitutive vs. inducible promoters

The industrial scale-up of systems fermentations with microbes designed to produce recombinant proteins using constitutive promoters is appealing since the exogenous addition of inducers is avoided. In developing a gene expression and protein-targeting system for *L. lactis*, Dieye and colleagues engineered expression cassettes to specifically target NucA to either the cytoplasm, supernatant, or cell wall (Dieye, et al., 2001). Expression of all cassettes was under the control of the strong constitutive lactococcal promoter *P59*. In this case, the authors did not observe any toxicity effects resulting from the constitutive overexpression of NucA, yet for the overproduction of other surface-targeted proteins, controlled gene expression is required (Wieczorek & Martin, 2010). Although Dieye and coworkers demonstrated the efficiency of NucA production under the control of the strong constitutive *P59* promoter, controlled expression of other heterologous proteins in *L. lactis* can increase protein yields and in

some cases, reduce cellular toxicity (Bermudez-Humaran, et al., 2004; de Vos, 1999; Narita, et al., 2006; Wieczorek & Martin, 2010).

The most developed system for the controlled expression of heterologous proteins in *L. lactis* is based on the mechanism of production of the 34-amino acid antimicrobial peptide nisin (Mierau & Kleerebezem, 2005). Nisin naturally binds to the cell wall synthesis precursor lipid II, generating pores in the cytoplasmic membrane, leading to cell death. Of particular interest is that the mechanism of nisin production is by auto-induction, where the presence of the peptide induces expression of the gene encoding it, *nisA* (Kuipers, et al., 1993). A major advance in utilizing nisin as an inducer for the production of recombinant proteins involved development of the NICE (Nisin-Controlled Gene Expression) system, which has been employed for over 15 years (Mierau & Kleerebezem, 2005). This system relies on two proteins involved in signal transduction and subsequent expression of genes under control of the *nisA* promoter. NisK is a histidine-protein kinase residing in the cell membrane and acts as a receptor for binding nisin. NisK then phosphorylates response regulator NisR, which in its activated state is a transcriptional activator of the *nisA* promoter *PnisA*. A number of proteins have been successfully produced in *L. lactis* where expression of the respective genes was under control of *PnisA* (Mierau & Kleerebezem, 2005). Strain *htrA*-NZ9000 has been a host organism of choice for the controlled production of recombinant proteins since it is both deficient in major housekeeping protease HtrA, and contains chromosomal integrations of the *nisR* and *nisK* genes necessary for induction of transcription by *PnisA* which is typically located on an expression vector (Miyoshi, et al., 2002). The inducible expression

of recombinant proteins is of particular importance when targeting proteins to the cell wall, since uncontrolled overexpression of recombinant proteins can impair cell wall biosynthesis, resulting in cellular toxicity (Bermudez-Humaran, et al., 2004; de Vos, 1999; Mierau & Kleerebezem, 2005; Narita, et al., 2006).

1.5.5 Export-specific reporter NucA

Initially used to screen for promoters (Poquet, et al., 1998), *Staphylococcus aureus* nuclease (NucA) enables the detection of proteins that are secreted when expressed as C- or N-terminal fusions. Detection of nuclease is achieved by overlaying plates containing grown colonies with Toluidine-Blue DNA (TBD) agar (Poquet, et al., 1998). A metachromatic shift results when the DNA is degraded by nuclease, resulting in pink halo formation around colonies secreting the recombinant NucA protein. In some cases, NucA has also been shown to increase the secretion efficiency and/or overall production when translationaly fused to poorly secreted proteins (Dieye, et al., 2003; Ribeiro, et al., 2002).

1.5.6 Anchoring mechanisms for LAB

Gram positive organisms surface display proteins by means of covalent and non-covalent interactions with the cell wall, S-Layer, or phospholipid bilayer (Leenhouts, et al., 1999) (Fig. 4).

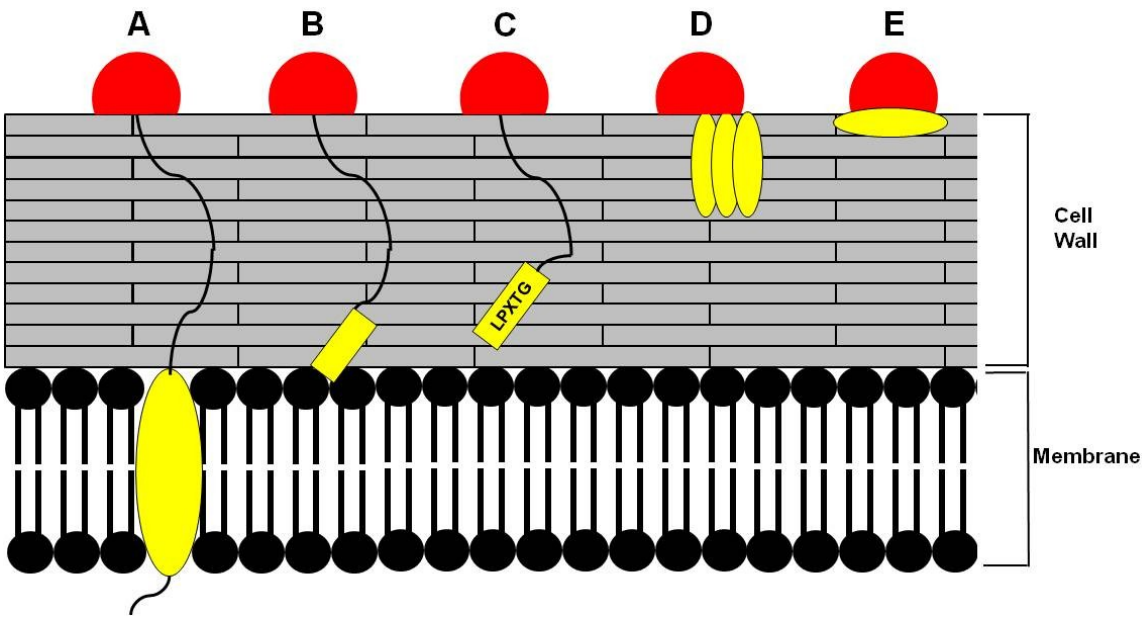

Figure 4. Modes of anchoring proteins to the cell surface of LAB. Yellow regions represent the various types of anchoring domains. Display is mediated by (A) transmembrane anchor region, (B) lipoprotein anchor domain (covalent), (C) LPXTG-containing cell-wall anchoring domain (covalent), (D) AcmA-repeats cell-wall binding domain, (E) SLH domains.

As discussed earlier, *C. thermocellum* localizes cellulosomal anchoring proteins to the S-layer via the non-covalent interactions of SLH domains with the cell wall (Schwarz, 2001). The S-layer consists of proteins that form porous lattices exterior to the cell wall. Subunits of the S-layer are identical, and account for in some cases, approximately 20% of the total cell protein content (Leenhouts, et al., 1999). Some LAB also contain S-layers, notably *L. brevis*, *Lactobacillus acidophilus*, *Lactobacillus crispatus*, and *Lactobacillus helveticus* (Leenhouts, et al., 1999). Fusion of SLH domains with recombinant proteins of interest has been used to achieve surface-display in LAB (Avall-Jaaskelainen, et al., 2002). However, since some LAB such as *L. casei* and *L. lactis* do not contain S-layers, use of this anchoring mechanism is not possible (Callegari, et al., 1998).

Another cellulosome-displaying bacterium, *R. flavefaciens*, anchors its cellulosome to the cell wall by means of a transpeptidase enzyme, sortase (Rincon, et al.,

2005). This enzyme catalyzes a transpepdidase reaction, cleaving within a recognized LPXTG motif, and covalently attaching the protein to lipid II, a cell-wall precursor that is incorporated into the peptidoglycan. *L. lactis* has its own sortase enzyme, and it has been demonstrated that a protein containing the appropriate C-terminal anchor can be successfully targeted to the cell wall (Piard, et al., 1997). The *Streptococcus pyogenes* M6 protein contains a 500 bp C-terminal anchor region, consisting of a well-conserved LPXTG motif, a hydrophobic region, and a positively-charged tail (Piard, et al., 1997). Positioning of the anchor is mediated by interactions between the hydrophobic region and cytoplasmic membrane, as well as the charged tail and negatively-charged phospholipids on the cytoplasmic face. Once positioned, sortase enzyme cleaves between the threonine and glycine of the LPXTG motif, forming the covalent bond between the threonine residue and the pentaglycine backbone of the peptidoglycan. By fusing proteins with the M6 anchor domain (cwa_{M6}), they can be successfully displayed on the surface of *L. lactis* (Dieye, et al., 2003; Dieye, et al., 2001; P. Lee & Faubert, 2006; Ribeiro, et al., 2002; Wieczorek & Martin, 2010). Lee and Faubert successfully displayed *Giardia lamblia* CWP2 on the cell surface of *L. lactis* by use of cwa_{M6}, and resulting strains were immunogenic (P. Lee & Faubert, 2006). Dieye and colleagues successfully displayed both NucA (Dieye, et al., 2001) and antigens VP2 and VP3 from infectious bursal disease virus (Dieye, et al., 2003), while Ribeiro and coworkers successfully displayed L7/L12 antigens on the surface of *L. lactis* as well (Ribeiro, et al., 2002).

1.5.7 Factors affecting protein secretion and anchoring in LAB

The history of recombinant protein production in *L. lactis* clearly suggests that protein size is not a bottleneck for their secretion. While smaller proteins, such as the 9.8 kDa Afp1 protein from *Streptomyces tendae*, have been successfully secreted in *L. Lactis* (Freitas, et al., 2005), so has the much larger 165 kDa dextransucrase from *Leuconostoc mesenteroides* (Neubauer, et al., 2003). It seems, rather, that protein conformation may affect a protein's ability to be efficiently secreted. To develop a LAB-based live vaccine delivery strategy for young cattle to combat bovine rotavirus, Enouf and colleagues attempted to secrete the antigenic protein NSP4 in *L. lactis* (Enouf, et al., 2001). The NSP4 protein was not detected in the supernatant, suggesting poor secretion, even when fused to sp_{Usp45} signal peptide. Other instances of low secretion efficiency of proteins between 9.8 and 165 kDa have also been reported. For example, Azevedo and co-workers observed low secretion efficiency of the *Brucella abortus* GroEL heat shock protein (Miyoshi, et al., 2006). A similar result was observed for a recombinant bovine β-lactoglobulin (BLG), a cow's milk allergen, in *L. lactis*, with the highest secretion efficiency reported at 5% (Chatel, et al., 2001; Nouaille, et al., 2005). Although the secretion efficiencies of NSP4, GroEL and BLG were consistency low, fusion of the heterologous protein to a signal peptide versus intracellular production resulted in increased yields of recombinant protein, most probably due their protection from intracellular degradation by the sec-pathway chaperones.

1.6 Hypothesis and rationale of this study

I hypothesize that by using *L. lactis* as a host organism, the surface display of multi-enzyme complexes can be achieved, and that the specific order and composition of enzymes within such complexes can be dictated by the architecture of engineered recombinant scaffolds. I plan to use fragments of CipA containing only type 1 cohesins as a proof of concept towards this goal, followed by the fusion of such simple scaffolds with type 2 cohesins in order to specifically position test enzymes at desired locations on such scaffolds. Fusion of reporter enzymes UidA and LacZ with appropriate dockerin domains and their subsequent binding to such surface-displayed chimeric scaffolds will hopefully generate insights into factors affecting the resulting enzymatic profiles of such complexes, including the sequential and simultaneous binding of these enzymes to the chimeric scaffolds. I hypothesize that by generating secondary scaffolds capable of binding both surface displayed scaffolds as well as test enzymes, I can demonstrate the potential of using such secondary scaffolds as auxiliary platforms to bind enzymes of interest.

Chapter 2

Engineering the cell surface display of cohesins for assembly of cellulosome-inspired enzyme complexes on *Lactococcus lactis*

Wieczorek, A.S. and V.J. Martin. 2010. *Engineering the cell surface display of cohesins for assembly of cellulosome-inspired enzyme complexes on Lactococcus lactis.* Microb Cell Fact. 9: p. 69.

2.1 Introduction

This chapter describes the incremental steps taken towards the cell surface display of small cellulosome scaffold proteins in *Lactococcus lactis*, a first and necessary step for the eventual engineering of extracellular protein complexes in this and other bacterial hosts. For this purpose, I chose *L. lactis,* a Gram-positive bacterium with established commercial value. *L. lactis* is of specific interest as it is generally regarded as safe (GRAS), has been used to produce valuable commodity chemicals such as lactic acid (Petrov, et al., 2008) and bioactive compounds (Hernandez, et al., 2007), and has been successfully engineered to secrete and/or display on its cell surface, a wide variety of proteins ranging from 9.8 to 165 kDa (Le Loir, et al., 2005). The metabolic engineering tools available in conjunction with the successful controlled expression and production of enzymes and proteins (Le Loir, et al., 2005) make it an ideal candidate for the recombinant expression of cellulosomal components. In addition, molecular tools for secreting and anchoring recombinant proteins by this bacterium have been previously established. I also hypothesized that by using a mutant strain lacking the major

extracellular housekeeping protease HtrA, I would be able to circumvent any undesirable proteolysis that may occur at the cell surface. While constitutive promoters are beneficial in large-scale fermentation processes, I also decided to test the nisin-inducible promoter in order to avoid any toxicity issues that may arise from overproduction of cell wall-targeted scaffolds.

Using *L. lactis* as a host, my experimental rationale was to display various fragments of the scaffold protein CipA ($CipA_{frags}$) on the cell surface. I subsequently attempted to bind a test enzyme, *E. coli* β-glucuronidase fused with the dockerin domain from CelS (dock1), on resulting scaffolds. This chapter of my thesis therefore describes the proof of concept that *L. lactis* can be engineered to display functional cohesin-containing scaffolds on its cell surface.

2.2 Materials and Methods

2.2.1 Bacterial strains and plasmids used in this chapter

The bacterial strains and plasmids reported in this chapter are listed in Table 2. *E. coli* strains were grown in Luria-Bertani medium at 37°C with shaking (220 rpm). *Lactococcus lactis htrA*-NZ9000 was grown in M17 medium (Terzaghi & Sandine, 1975) supplemented with 1% (w/v) glucose (GM17) at 30°C without agitation. *C. thermocellum* was grown in ATCC1191 medium at 55°C with 0.2% (w/v) cellobiose as a carbon source. Where appropriate, antibiotics were added as follows: for *E. coli*, ampicillin (100 μg/mL), erythromycin (150 μg/mL), chloramphenicol (10 μg/mL) and kanamycin (30 μg/mL); for *L. lactis,* erythromycin (5 μg/mL) and chloramphenicol (10 μg/mL). General molecular

Table 2 - Strains and plasmids used in this chapter. [a]Vector pSCNIII was a gift provided by Jos Seegers (unpublished data). pAW100 series of vectors are nisin-inducible and contain an intact rbs_{usp45}. pAW300 series vectors are nisin-inducible and contain an intact rbs_{nisA}. pAW500 series vectors are pAW300 variants lacking an N-terminal NucA fusion. *P59*, constitutive lactococcal promoter; *PT7*, inducible T7 promoter; P_{nisA}, inducible *nisA* promoter; rbs_{usp45}, Usp45 ribosome-binding site; rbs_{nisA}, *nisA* ribosome-binding site; sp_{Usp45}, signal sequence of Usp45; *nucA*, staphylococcal nuclease; cwa_{M6}, anchor motif of M6 protein; *llt2*, transcriptional terminator of *rrnB* operon; t_{trpA}, transcriptional terminator of *trpA.*

Strain	Genotype / Decription	Source
L. lactis htrA-NZ9000	Mutant MG1363 derivative (*nisRK* genes on the chromosome) lacking *htrA*	(Miyoshi, et al., 2002)
E. coli TG1	*supE thi-1* Δ(*lac-proAB*) Δ(*mcrB-hsdSM*)*5* (rK– mK–) [F′ *traD36 proAB lacI*q*ZΔM15*]	ATCC
E. coli DH5α	*fhuA2* Δ*(argF-lacZ)U169 phoA glnV44 Φ80* Δ*(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	Invitrogen
E. coli BL21 (DE3)	F^- *ompT gal dcm lon* $hsdS_B(r_B^- m_B^-)$ *λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])*	Novagen
Plasmid		
pVE5524	Ery^r, Amp^r; pBS::pIL252::t_{trpA}::*P59*::rbs_{usp45}::sp_{Usp45}-*nucA*-cwa_{M6}-*t1t2*	(Dieye, et al., 2001)
pVE5523	Ery^r, Amp^r; pBS::pIL252::t_{trpA}::*P59*::rbs_{usp45}::sp_{Usp45}-*nucA-t1t2*	(Dieye, et al., 2001)
pSIP502	Ery^r; $P_{nisA::}rbs_{nisA::}$*uidA*	(Sorvig, et al., 2003)
pSCNIII	Cm^r	J. Seegers[a]
pUC19	Amp^r	(Yanisch-Perron, et al., 1985)
pET28(b)	Kn^r	Novagen

pSIPsp-nuc	Ery^r; *P_{nisA}*::rbs$_{nisA}$::sp$_{Usp45}$-*nucA*	This Work
pUC104	Amp^r; *t_{trpA}*::*P_{nisA}*::*rbs_{usp45}*::*sp*$_{Usp45}$-*nucA*	This Work
pUC104mod	Amp^r; *t_{trpA}*::*P59*::*rbs_{usp45}*::*sp*$_{Usp45}$-*nucA*	This Work
pUC304	Amp^r; *t_{trpA}*::*P_{nisA}*::*rbs_{nisA}*::*sp*$_{Usp45}$-*nucA*	This Work
pUC504	Amp^r; *t_{trpA}*::*P_{nisA}*::*rbs_{nisA}*::*sp*$_{Usp45}$	This Work
pAW004	Ery^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P59*::*rbs_{usp45}*::*sp*$_{Usp45}$-*nucA-MCS-cwa*$_{M6}$-*t1t2*	This Work
pAW005	Ery^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P59*::*rbs_{usp45}*::*sp*$_{Usp45}$-*nucA-MCS-t1t2*	This Work
pAW004Z	Ery^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P59*::*rbs_{usp45}*::*sp*$_{Usp45}$-*nucA-lacZα-cwa*$_{M6}$-*tlt2*	This Work
pAW005Z	Ery^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P59*::*rbs_{usp45}*::*sp*$_{Usp45}$-*nucA- lacZα-tlt2*	This Work
pAW004ZC	Cm^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P59*::*rbs_{usp45}*::*sp*$_{Usp45}$-*nucA-lacZα-cwa*$_{M6}$-*tlt2*	This Work
pAW005ZC	Cm^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P59*::*rbs_{usp45}*::*sp*$_{Usp45}$-*nucA- lacZα-tlt2*	This Work
pGEMc9	Amp^r; pGEMT::with cloned *coh9* from *cipA*	This Work
pGEMc1	Amp^r; pGEMT::with cloned *coh1* from *cipA*	This Work
pGEMc1-c2	Amp^r; pGEMT::with cloned *coh1-coh2* from *cipA*	This Work
pGEMcbm-c3	Amp^r; pGEMT::with cloned *cbm3a-coh3* from *cipA*	This Work
pGEMcbm	Amp^r; pGEMT::with cloned *cbm3a* from *cipA*	This Work
pAW104	Cm^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P_{nisA}*::*rbs_{usp45}*::*sp*$_{Usp45}$-*nucA-LacZα-cwa*$_{M6}$-*tlt2*	This Work
pAW105	Cm^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P_{nisA}*::*rbs_{usp45}*::*sp*$_{Usp45}$-*nucA-LacZα-tlt2*	This Work
pAW301	Cm^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P_{nisA}*::*rbs_{nisA}*::*sp*$_{Usp45}$-*nucA-cwa*$_{M6}$-*tlt2*	This Work
pAW302	Cm^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P_{nisA}*::*rbs_{nisA}*::*sp*$_{Usp45}$-*nucA-tlt2*	This Work
pAW304	Cm^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P_{nisA}*::*rbs_{nisA}*::*sp*$_{Usp45}$-*nucA-lacZα-cwa*$_{M6}$-*tlt2*	This Work
pAW305	Cm^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P_{nisA}*::*rbs_{nisA}*::*sp*$_{Usp45}$-*nucA-lacZα-tlt2*	This Work
pAW307	Cm^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P_{nisA}*::*rbs_{nisA}*::*sp*$_{Usp45}$-*nucA-coh9-cwa*$_{M6}$-*tlt2*	This Work
pAW308	Cm^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P_{nisA}*::*rbs_{nisA}*::*sp*$_{Usp45}$-*nucA-coh9-tlt2*	This Work
pAW310	Cm^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P_{nisA}*::*rbs_{nisA}*::*sp*$_{Usp45}$-*nucA-coh1-cwa*$_{M6}$-*tlt2*	This Work
pAW311	Cm^r, Amp^r; pBS::pIL252::*t_{trpA}*::*P_{nisA}*::*rbs_{nisA}*::*sp*$_{Usp45}$-*nucA-coh1-tlt2*	This Work

pAW334	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*nucA-coh1-coh2-cwa*$_{M6}$-*tlt2*	This Work
pAW335	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*nucA-coh1-coh2-tlt2*	This Work
pAW328	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*nucA-cbm3a-coh3-cwa*$_{M6}$-*tlt2*	This Work
pAW329	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*nucA-cbm3a-coh3-tlt2*	This Work
pAW331	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*nucA-cbm3a-cwa*$_{M6}$-*tlt2*	This Work
pAW332	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*nucA-cbm3a-tlt2*	This Work
pAW504	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*lacZα-cwa*$_{M6}$-*tlt2*	This Work
pAW505	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*lacZα-tlt2*	This Work
pAW507	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*coh9-cwa*$_{M6}$-*tlt2*	This Work
pAW508	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*coh9-tlt2*	This Work
pAW510	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*coh1-cwa*$_{M6}$-*tlt2*	This Work
pAW511	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*coh1-tlt2*	This Work
pAW534	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*coh1-coh2-cwa*$_{M6}$-*tlt2*	This Work
pAW535	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*coh1-coh2-tlt2*	This Work
pAW528	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*cbm3a-coh3-cwa*$_{M6}$-*tlt2*	This Work
pAW529	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*cbm3a-coh3-tlt2*	This Work
pAW531	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*cbm3a-cwa*$_{M6}$-*tlt2*	This Work
pAW532	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*cbm3a-tlt2*	This Work
pETdock1	Kn^r; pET28(b)::with cloned *dock1* from *celS*	This Work
pETUdock1	Kn^r; pET28(b)::*PT7*::*6xHis-uidA-dock1*	This Work
pETU	Kn^r; pET28(b)::*PT7*::*6xHis-uidA*	This Work

biology techniques for *E. coli* were performed as previously described (Sambrook & Russell, 2001). Genomic DNA was isolated from *C. thermocellum* as previously described (Wang & Wu, 1993). To make competent cells, *L. lactis* was grown in M17 medium

(Terzaghi & Sandine, 1975) supplemented with 1% (w/v) glucose, 25% (w/v) sucrose and 2% (w/v) glycine and cells were transformed as previously described (Holo & Nes, 1989). M17 media was supplied by Oxoid, LB media was supplied by Novagen, all antibiotics, ρ-nitrophenyl-β-D-glucuronide and nisin were provided by Sigma, and X-gal and IPTG were supplied by Fermentas.

2.2.2 Assembly of cassettes for scaffold protein expression and targeting

The *E. coli-L. lactis* shuttle vectors pVE5524 and pVE5523 were used as backbone plasmids for targeting fragments of the CipA scaffold protein to the cell surface or supernatant, respectively (Dieye, et al., 2001). The various $CipA_{frags}$ were produced as fusions with the N-terminal signal peptide from the lactococcal Usp45 (Genbank Accession no. AAA25230.1) secreted protein (sp_{Usp45}) (van Asseldonk, et al., 1990) and for targeting to the cell wall, as a fusion with the C-terminal anchor from the *Streptococcus pyogenes* M6 (Genbank accession no. AAA26920.1) protein (cwa_{M6}) (Piard, et al., 1997) (Fig. 5). Expression cassettes were designed to allow the optional fusion of $CipA_{frags}$ with an N-terminal nuclease reporter (NucA) used for detection of the fusion proteins in the extracellular milieu (Dieye, et al., 2003; Ribeiro, et al., 2002) (Fig. 5). The strong constitutive lactococcal promoter *P*59 (Dieye, et al., 2001) and the P_{nisA} nisin-inducible promoter from the *nisA* gene of *L. lactis* (Sorvig, et al., 2003) were tested for optimal expression of the recombinant scaffolds. Two ribosome-binding sites were also tested, that of the *usp45* gene (rbs_{usp45}) (Dieye, et al., 2001) and that of the *nisA* gene (rbs_{nisA})

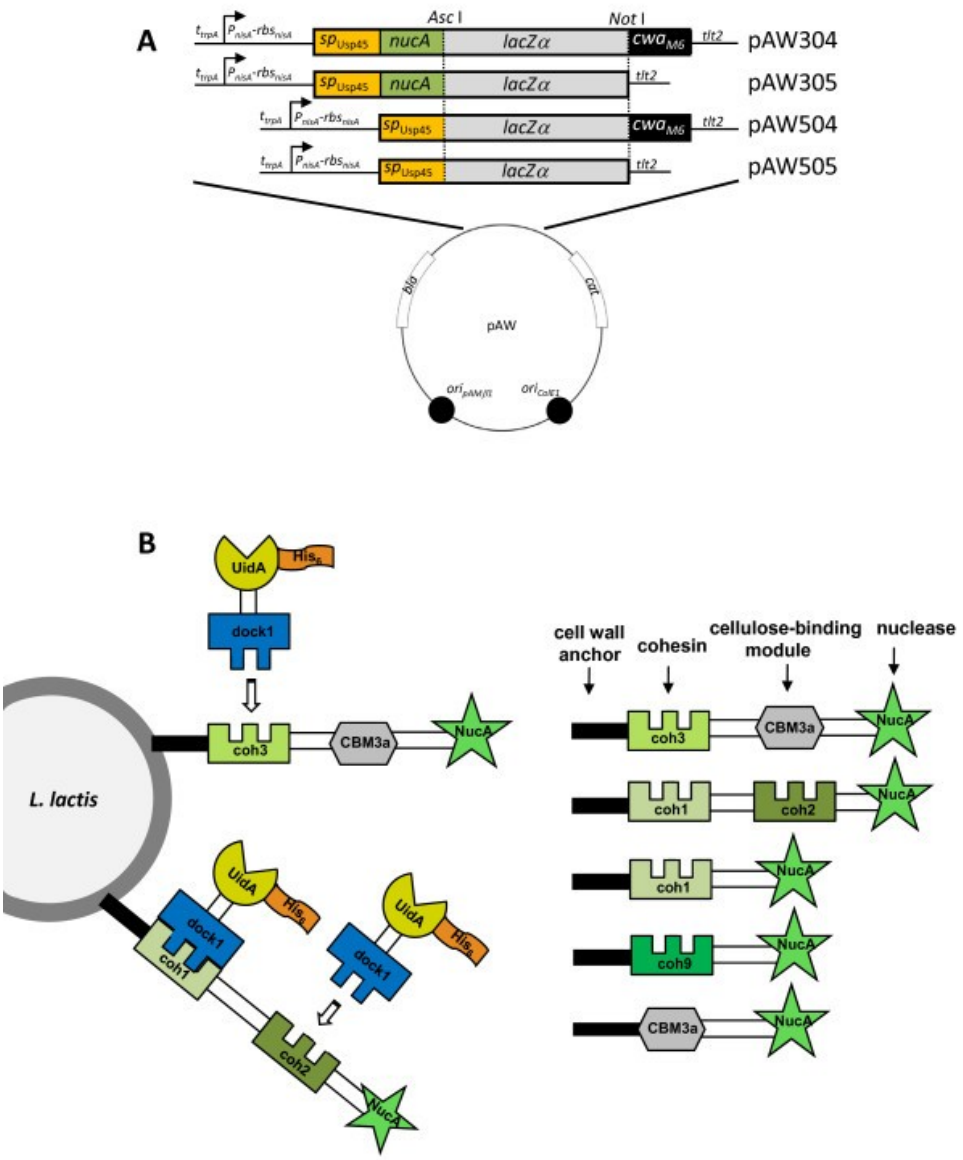

Figure 5. pAW series of *cipA*$_{frag}$ expression vectors and strategy for complex assembly. (A) Vectors were designed for facilitated insertion of fragments of the gene encoding the cellulosomal scaffold protein CipA, into *Asc*I-*Not*I restriction sites. Scaffolds can be optionally expressed with or without an N-terminal nuclease reporter and/or a C-terminal cell wall anchor motif. pAW304 is designed for expression, secretion, and cell wall-targeting of CipA fragments (CipA$_{frags}$) as fusions with the N-terminal NucA reporter. pAW305 is designed for the expression and secretion of CipA$_{frags}$ as a fusion with the N-terminal NucA reporter, but without the C-terminal anchor motif. pAW504 is designed for expression, secretion, and cell wall-targeting of CipA$_{frags}$ without the N-terminal NucA reporter. pAW505 is designed for the expression and secretion of CipA$_{frags}$ with neither the N-terminal NucA reporter nor the C-terminal anchor motif. **(B)** Graphic depiction of the surface-display strategy of engineered scaffolds and their association with the □-glucuronidase-dockerin fusion protein (UidA-dock1). All successfully displayed CipA$_{frags}$ are portrayed as fusions with both NucA and a cell wall anchor, however were also expressed and tested without these two components.

(Sorvig, et al., 2003). In order to facilitate the exchange of scaffold fragments in the expression cassette, *Asc*I-*Not*I restriction sites were engineered just downstream of *nucA* (Fig. 5). To achieve this, an 800-bp fragment containing the *nucA* gene was PCR-amplified from pVE5524 using primers *a* and *b* (Table 3), digested with *Sal*I-*Eco*RV and ligated into similarly digested pVE5524 and pVE5523, yielding pAW004 and pAW005. To facilitate detection of *E. coli* clones that harbor *cipA* fragments, a *lacZ*-α stuffer fragment was PCR-amplified from pUC19 using primers *c* and *d* (Table 3) digested with *Asc*I-*Not*I, and subsequently ligated into similarly cut pAW004 and pAW005, yielding pAW004Z and pAW005Z, respectively. Since *L. lactis htrA*-NZ9000 is resistant to erythromycin, the *ery* marker of the pAW vectors was replaced with the *cat* gene from pSCNIII. The *cat* gene was PCR-amplified using primers *e* and *f* (Table 3), digested with *Afl*II and *Hpa*I, and ligated into similarly digested pAW004Z and pAW005Z, yielding plasmids pAW004ZC and pAW005ZC, respectively. For inducible expression of the scaffolds, I replaced the *P59* promoter with P_{nisA} from pSIP502. The P_{nisA} promoter was isolated using primers *o* and *p* (Table 3), digested with *Apa*I-*Nru*I and ligated to similarly digested pAW004ZC and pAW005ZC, yielding pAW104 and pAW105, respectively.

2.2.3 Cloning of *cipA* fragments from *C. thermocellum*

Five unique *cipA* fragments were PCR-amplified from *C. thermocellum* genomic DNA using primer pairs *g-h, i-j, g-k, l-m* and *n-m* (Table 3), ligated into pGEM-T (Promega) and sequenced to verify the integrity of the gene sequence. The resulting pGEM plasmids

Table 3 - Primers used in this chapter (Restriction enzyme cut sites are in bold).

Primer	Sequence (5' – 3')
a	TAT**AGATCT**TCGATAGCCCGCCTAATGAGC
b	AT**GATATC**GCGGCCGCGGCGCGCCTCGAGATCGATTTG
c	TAGATATC**GGCGCGCC**ATTAGCTATGCGGCATCAGAGC
d	TAGCTAGC**GCGGCCGC**GCCCAATACGCAAACCGCCTC
e	GATCTAGC**CTTAAG**TTCAACAAACTCTAGCGCC
f	CGTAGATC**GTTAAC**CCTTCTTCAACTAACGGGG
g	TCGA**GGCGCGCC**CGGCCACAATGACAGTCGAGA
h	TCGA**GCGGCCGC**CGGTACGGAACTACCAAGAT
i	TA**GGCGCGCC**ATAAGTTGACACTTAAGATAGGCAG
j	TA**GCGGCCGC**AGTTACAAGTACTCCACCATTG
k	TCGA**GCGGCCGC**CGGTGTTGCATTGCCAACGT
l	TCGA**GGCGCGCC**CGGATGATCCGAATGCAATAAAG
m	TCGA**GCGGCCGC**TACTACACTGCCACCGG
n	TGA**GGCGCGCC**CGGCAAATACACCGGTATC
o	ATGC**GGGCCC**GACCTAGTCTTATAACTATACTG
p	ATGTAC**TCGCGA**TTTATTTTGTAGTTCCTTCGAACG
q	AGAACAG**TCATGA**AAAAAAAGATTATCTC
r	ATAT**CTCGAG**ATCGATTTGACCTGAATCA
s	AGTC**ACATGT**TCTTTCCTGCGTTATCCCCTG
t	ATGC**TCGCGA**AGATCTGGGATCAAAAAAAAGCCCGC
u	GCTT**GAATTC**TCTACTAAATTATACGGCGACGTCAATG
v	GCTT**GCGGCCGC**TTTAGTTCTTGTACGGCAATGTATC
w	ATGC**GCTAGC**ATGTTACGTCCTGTAGAAACC

were digested with *Asc*I-*Not*I to release the *cipA* gene fragments and these were ligated into pAW004ZC and pAW005ZC. The *cipA* fragments were chosen on the basis of containing a single cohesin (coh1 or coh9), two cohesins of identical specificity (coh1-coh2), one cohesin and a cellulose-binding domain (coh3-CBM3a) and only a cellulose-binding domain (CBM3a) (Fig. 5). The resulting sp_{Usp45}-*nucA*-$cipA_{\text{frag}}$-cwa_{M6} cassettes were under control of the P_{59} promoter and contained rbs_{usp45}. The same *cipA* fragments were cloned into pAW104 and pAW105 for inducible expression of the scaffold proteins.

For the inducible expression of the fusion proteins under the control of P_{nisA} with an intact ribosome-binding site from the *nisA* gene (rbs_{nisA}), sp_{Usp45}-*nucA* was PCR-amplified from pAW004ZC using primers *q* and *r,* creating a *Bsp*HI cut site at the 5' end of the PCR product. The PCR product was digested with *Bsp*HI and *Xho*I and ligated to pSIP502 digested with *Nco*I-*Xho*I, effectively replacing the *gusA* gene with sp_{Usp45}-*nucA*, retaining the first lysine of the signal peptide, and yielding pSIPSPNUC. For the insertion of an upstream transcriptional terminator and removal of *nucA*, a 1500-bp *Sap*I-*Xba*I fragment was temporarily removed from pAW104, and was ligated to similarly cut pUC19, yielding vector pUC104. To introduce the *E. coli* transcriptional terminator from the tryptophan synthase operon (t_{trpA}) upstream of P_{nisA} and to introduce a *Bgl*II cut site, a 200-bp fragment containing t_{trpA} was PCR-amplified from pVE5524 using primers *s* and *t*, digested with *Afl*III-*Nru*I and ligated to similarly-cut pUC104, yielding pUC104mod. Plasmid pSIPSPNUC was digested with *Bgl*II-*Xho*I and ligated to similarly-digested pUC104mod, yielding vector pUC304. This was the base vector harboring the t_{trpA}-P_{nisA}-rbs_{nisA}-sp_{Usp45}-*nucA* cassette, which was digested with *Apa*I-*Asc*I and ligated into the

pAW100 series of vectors. Inserting this cassette into *Apa*I-*Eco*RV digested pAW110 and pAW111, yielding pAW301 and pAW302, respectively, created controls lacking *cipA* fragments for expression of *nucA* alone. For deletion of the *nucA* reporter and construction of the pAW500 series, pUC304 was digested with *Sal*I-*Xho*I and self-ligated, yielding vector pUC504. The t_{trpA}-P_{nisA}-rbs_{nisA}-sp_{Usp45} cassette was released via digestion with *Apa*I-*Asc*I, gel-purified, and ligated to similarly-cut pAW100 series vectors, yielding the pAW500 series of vectors. This cassette was also ligated into similarly cut pAW104 and pAW105 yielding base vectors containing the *lacZ-α* stuffer fragment. The final expression vectors for this study included the pAW300 series of vectors for inducible expression and targeting of NucA-fused scaffolds, and the pAW500 series of vectors for inducible expression and targeting of scaffolds lacking the N-terminal NucA reporter (Fig. 5).

2.2.4 Expression and localization of $CipA_{frags}$ in *L. lactis*

L. lactis htrA-NZ9000 was transformed with the pAW300 and pAW500 series of vectors for the controlled expression of scaffolds. It contains chromosomal copies of the *nisR* and *nisK* genes necessary for nisin-inducible expression of cassettes under control of the *nisA* promoter, and is deficient in a major extracellular protease, which has been shown previously to be responsible for the proteolysis of exported recombinant proteins (Miyoshi, et al., 2002). Growth curves were used to evaluate whether the over-expressed $CipA_{frag}$ proteins caused growth inhibition. Growth curves were performed in 96 well plates and cells were induced with 10 ng nisin/mL at inoculation (t=0 hrs), 4 hrs post-

inoculation (t=4 hrs) or were not induced. For the expression of $CipA_{frag}$ proteins in *L. lactis htrA*-NZ9000, overnight cultures were diluted 1/50 into fresh GM17 medium and were induced with 10 ng nisin/mL when an $OD_{600} \approx 0.3$ was reached (4 hrs). After 20 hrs growth, successful $CipA_{frag}$ secretion was evaluated using a nuclease assay consisting of spotting cells on TBD-agar and observing pink color formation (Dieye, et al., 2001). For analysis of NucA-$CipA_{frag}$ proteins in various cellular locations, cell fractionation was performed as described previously (Piard, et al., 1997), with the addition of lysostaphin (0.6 mg/mL) (Steidler, et al., 1998). Aliquots of proteins were blotted on TBD-agar plates and formation of a pink color was analyzed after a 1-hr incubation at 37°C.

2.2.5 Expression and purification of $CipA_{frag}$-binding β-glucuronidase

The *E. coli* β-glucuronidase (UidA, GenBank accession no. ZP_03034971.1) was engineered to have a C-terminal dock1 domain for binding onto $CipA_{frag}$ scaffolds, as well as an N-terminal 6 x His-tag for protein purification. The dock1 domain of the *C. thermocellum celS* gene was amplified from *C. thermocellum* genomic DNA using primers *u* and *v* (Table 3). PCR products were digested with *Eco*RI-*Not*I and ligated to similarly-digested pET28(b), yielding pETdock1. The *uidA* gene lacking a stop codon was amplified using primers *w* and *x* and pSIP502 as template. The PCR product was digested with *Nhe*I-*Eco*RI and ligated to similarly-cut pET28(b) and pETdock1, yielding His-tagged UidA proteins with and without a dock1 domain (pETUdock1 and pETU). His-tagged proteins were expressed in *E. coli* BL21(DE3). Cultures were induced at an OD_{600} of 0.5 with 1mM IPTG and incubated for an additional 5 hrs at 37°C. Cells were harvested (1000 x g, 10

min, 4°C) and cell pellets were kept overnight at -80°C. Thawed cell pellets were suspended in 50 mM phosphate buffer, pH 7.5, containing 300 mM NaCl. Samples were subjected to sonication (15 sec pulse, 5 sec pause between pulses, for a total of 2 minutes) and lysates were loaded on approximately 10 mL of Ni-NTA sepharose resin. The resin was washed with phosphate buffer (50 mM, pH 6.0) containing 300 mM NaCl and 20 mM imidazole and eluted using the same buffer containing 250 mM imidazole. Fifty µL of each elution fraction were added to 450 µL GUS buffer containing 50 mM sodium phosphate buffer (pH 7), 10 mM β-mercaptoethanol, 1 mM ethylenediaminetetraacetic acid and 0.1% (v/v) Triton X-100. Samples were heated for 1 min, after which *p*-nitrophenyl-β-D-glucuronide was added to a final concentration of 4 mg/mL (Axelsson, et al., 2003). The UidA-containing fractions were identified by the appearance of a yellow color. Proteins from the elution fractions showing UidA activity were visualized by SDS-PAGE on a 12% (w/v) gel to identify fractions containing the highest purity of enzyme. The specific activities of UidA-dock1 and UidA were determined by colorimetric assays in a thermostated UV-Vis spectrophotometer (Cary 50 WinUv) at 405nm, using a 1 cm (L) cuvette, and the known molar extinction coefficient of *p*-nitrophenol being 18 000 M^{-1} cm^{-1}. Quantification of the proteins was done using a Bradford protein assay kit (Pierce) and BSA as a standard. Specific activities were used to evaluate the amount of enzyme bound to cells in the *in vivo* binding assay described below.

2.2.6 Binding of β-glucuronidase to *L. lactis*

L. lactis htrA-NZ9000 cells harboring the pAW300 or pAW500 series of vectors, as well as the plasmid-free strain were grown overnight in GM17 medium. Cultures were diluted 1/50 in 5 mL of fresh media and grown for an additional 4 hrs (OD_{600} ≈0.3) after which cells were induced with 10 ng nisin/mL for scaffold expression. After 20 hrs of growth, cells from 1-mL of culture were harvested (4,300 x g, 5 min, 4°C) washed once in phosphate buffer (50 mM, pH 6.0) containing 300 mM NaCl and suspended in 100 μL of purified UidA-dock1 or UidA at a concentration 100 μg/mL. To ensure that saturation of all cohesin sites was achieved, binding assay with 200 μg UidA-dock1/mL was tested for *L. lactis* harboring pAW328. Binding was carried out at 4ºC for 10 hrs. Cells were then washed 6 times to eliminate residual enzyme activity and suspended in 100 μL of phosphate buffer (50 mM, pH 6.0) containing 300 mM NaCl for detection of β-glucuronidase activity. For quantification of bound UdiA-dock1, 50 μL of washed cells were analyzed for β-glucuronidase activity. Reactions were stopped with 250 μL of 1 M sodium carbonate once a yellow color appeared, and the duration of each assay was recorded. The specific activities of the purified UidA-dock1 and UidA were used to determine the amount of enzyme bound onto the *L lactis* cells. Using the calculated molecular weight of UidA-dock1 and the known amount of cells present in each sample, the average number of enzyme units bound per cell was estimated. Considering that cohesins and dockerins form complexes in a 1:1 ratio, the number of enzymes present per cell provides an estimate of the number of cohesins on the cell surface. The theoretical molecular weight of the scaffolds was also used to estimate the total amount

of successfully anchored recombinant proteins in each respective culture. Experiments were repeated twice and true biological replicates (independent colonies and cultures) were performed in triplicate for all samples.

2.3 Results

2.3.1 Regulated expression of $CipA_{frags}$ yields the surface-display of scaffold proteins

L. lactis htrA-NZ9000 cells were successfully transformed with either the pAW500 series or pAW300 series of vectors (Fig. 5A), resulting in strains expressing fragments of CipA ($CipA_{frags}$) alone, or as fusions with the NucA export-specific reporter, and/or the cwa_{M6} for anchoring of the scaffold to the cell-surface (Fig. 5B). Growth curves of engineered *L. lactis* strains were used to determine if the expression and secretion of scaffold proteins resulted in growth inhibition. Results from the growth experiments showed a correlation between $cipA_{frag}$ gene expression and growth inhibition (Fig. 6). The constitutive over-production of recombinant proteins targeted to the cell surface in *L. lactis* may interfere with the integrity of the cell wall (Narita, et al., 2006), whereas in *C. thermocellum,* the constitutive expression of CipA is modulated through catabolite repression (Y. H. Zhang & Lynd, 2005). In the absence of the inducer nisin, all $cipA_{frag}$-expressing strains grew similarly to the control *L. lactis htrA*-NZ9000 with a final cell density corresponding to an OD_{600} approaching 0.7 (Fig. 6A, D, G). This indicated that little change in growth profile resulted from any leaky expression of the recombinant proteins. Nisin induction at inoculation resulted in cellular toxicity, as demonstrated by

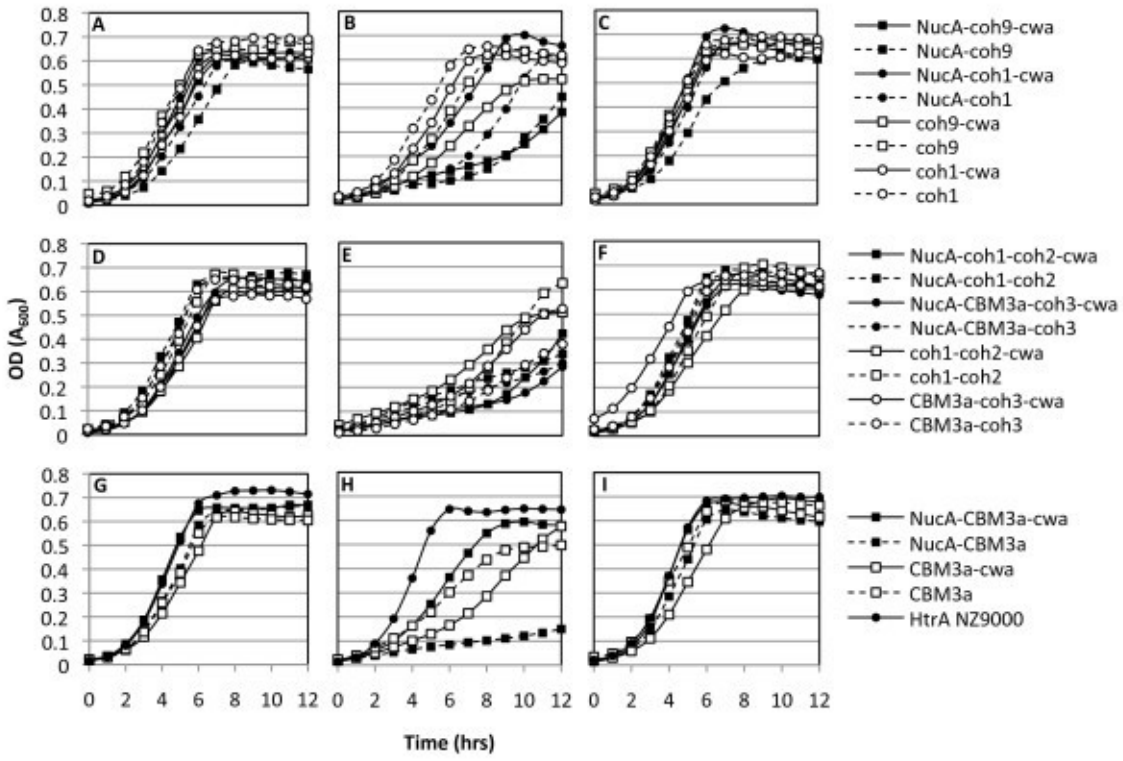

Figure 6. Growth profiles of *L. lactis* expressing $CipA_{frags}$ alone or as fusions with $M6_{cwa}$ and/or NucA. Panels A, D and G represent cultures not induced with nisin, panels B, E, H represent cultures induced with 10 ng/mL nisin at inoculation (t=0hrs), and panels C, F, I represent cultures induced with 10 ng/mL nisin in log phase corresponding to an OD_{600} ≈ 0.3 (t=4 hrs). Constructs were grouped according to their modular nature. Top panels depict constructs containing a single cohesin; Middle panels depict constructs containing two CipA domains; Lower panels depict constructs containing no cohesin domains. Black shapes indicate scaffolds containing a fusion with NucA, and white shapes indicate scaffolds where NucA has been removed. Solid lines represent scaffolds expressed with a cell wall anchor, and dotted lines represent scaffolds lacking the cell wall anchor. Experiments were repeated three times yielding identical trends between growth profiles.

extended lag phases, lower growth rates and in most cases, final cell yields (Fig. 6B, E, H). In all cases, when induction of protein expression was carried out after 4 hrs of growth (corresponding to an OD_{600} ≈0.3), cultures did not display growth retardation and final cell densities were similar to those attained with no induction (Fig. 6C, F, I). Expression of the various *cipA*$_{frags}$ from the constitutive P_{59} promoter consistently resulted in plasmid

rearrangements as observed by restriction digest analysis of the rescued plasmids from both *E. coli* and *L. lactis* (data not shown). From these results, my data suggest that unregulated high-level expression of the $CipA_{frag}$ proteins was toxic to the cells and using a constitutive promoter such as P_{59} induced plasmid rearrangements that abolished or reduced $cipA_{frag}$ expression. These results confirmed the necessity for regulating expression of the proteins, which was achieved using the P_{nisA} promoter. With the exception of cell wall anchored scaffold containing only a cellulose-binding domain (CBM3a-cwa) (Fig. 6H), removal of the NucA lowered or eliminated toxicity to the cells, as observed by improved growth rates and yields.

2.3.2 NucA-$CipA_{frag}$ proteins are localized to the cell wall of *L. lactis*

In order to quickly evaluate my success at recombinant protein secretion in *L. lactis*, a nuclease enzyme was fused to the CipA fragments to be displayed on the cell surface. *L. lactis* cells harboring the pAW300 series of vectors all displayed a $NucA^+$ phenotype on plates overlaid with TBD agar, confirming that all variants of the NucA-$CipA_{frag}$ proteins were successfully secreted and that the nuclease retained its function when expressed as an N-terminal fusion to $CipA_{frags}$. To determine the cellular localization of the expressed $CipA_{frag}$ fusion proteins, cell fractionations were performed, and cytoplasmic, cell wall, and supernatant fractions were spotted on TBD agar. Of the secreted NucA-$CipA_{frag}$ proteins, almost all gave detectable amounts of nuclease activity in the cell wall fractions corresponding to proteins released from lysozyme / lysostaphin treatments, suggesting successful cell wall targeting of the proteins (Fig. 7). $CipA_{frag}$

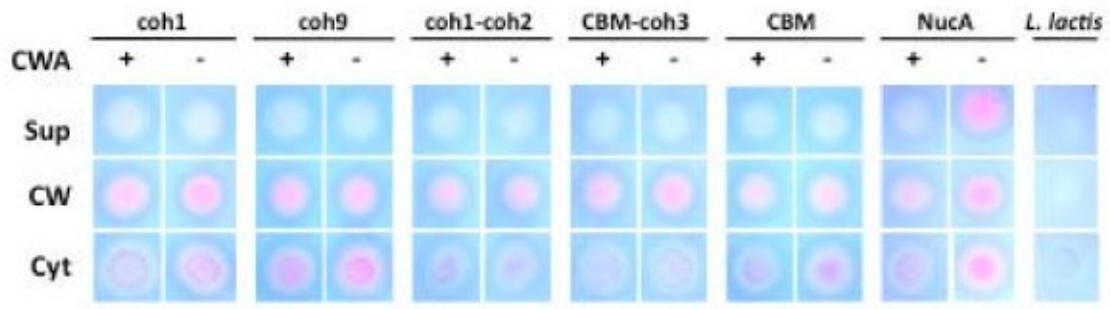

Figure 7. Cellular localization of NucA-CipA$_{frag}$ scaffolds expressed by *L. lactis* with or without M6$_{cwa}$. NucA activity was detected by spotting cell fractions on TBD-agar and analyzing for pink color formation. Fractions analyzed are supernatant (sup), cell wall (cw), and cytoplasm (cyt). Constructs are represented by their respective CipA$_{frag}$ components and were expressed as fusions with NucA with or without cell wall anchor (cwa) domains.

proteins were not detected in the supernatant, suggesting that secreted proteins remained localized to the cell wall due to the activity of lactococcal sortase. Unexpectedly, the NucA- CipA$_{frag}$ fusions lacking the cell wall anchor domain were also detected primarily in the cell wall fractions (Fig. 7) suggesting that fusion of NucA with CipA$_{frags}$ caused the scaffolds to remain associated with the cell wall, even without covalent cross-linking by sortase. All of the cytoplasmic fractions were also found to contain varying levels of expressed scaffolds, a finding consistent with observations previously made while exporting recombinant proteins in *L. lactis* (Dieye, et al., 2003; Dieye, et al., 2001; Miyoshi, et al., 2002; Ribeiro, et al., 2002). My data suggest that these cytoplasmic proteins were either in the process of being synthesized and exported by the cell via cytoplasmic chaperones, or had evaded the sec-pathway due to a lack of recognition of the signal sequence. In certain instances, the net charge of N-terminal residues downstream of the signal peptide can also contribute to the poor secretion

efficiency of recombinant proteins (Langella & Le Loir, 1999). As expected from previous studies (Dieye, et al., 2001; Miyoshi, et al., 2002) in the absence of a cell wall anchor domain, NucA was secreted into the supernatant (Fig. 7).

2.3.3 Cell surface displayed $CipA_{frag}$ scaffolds bind UidA-dock1

In vivo binding assays were performed to determine if a dockerin-containing enzyme could associate with cell surface displayed $CipA_{frag}$ scaffold proteins. *L. lactis* cells expressing cell wall and supernatant-targeted scaffolds were incubated with purified β-glucuronidase enzymes fused to a dockerin domain (UidA-dock1). After incubation, washed cells were assayed for β-glucuronidase activity, allowing a relative comparison of $CipA_{frag}$ display efficiencies between engineered constructs. All constructs containing cohesin domains as part of their scaffolds successfully bound UidA-dock1, while those lacking cohesins as well as the plasmid-free *L. lactis htrA*-NZ9000 failed to do so (Fig. 8). Binding experiments using UidA lacking dock1 resulted in no successful "docking" onto *L. lactis* displaying NucA-CBM3a-coh3 (Fig. 8A) or any other recombinant scaffolds (data not shown). These results demonstrated that functional recombinant scaffolds could be expressed on the surface of *L. lactis* and that cell surface complex formation was dependent on the presence of both cohesin and dockerin domains. Among those strains secreting and displaying functional scaffolds, significant variation in display efficiency was observed. Assuming a 1:1 enzyme-to-cohesin ratio, the approximate number of cohesins and/or scaffolds per cell was determined. The strains that displayed the greatest number of nuclease bearing scaffolds (~$9x10^3$ scaffolds/cell) were those expressing the cohesin 1

domain alone (coh1-cwa and NucA-coh1-cwa) (Fig. 8). Strains expressing coh9-cwa, NucA- coh9-cwa, coh1-coh2-cwa, CBM3a-coh3-cwa and NucA-CBM3a-coh3-cwa, were estimated to display between $5.0x10^3$ and $6.3x10^3$ scaffolds/cell. These results suggested that the size of the $CipA_{frag}$ is not necessarily the limiting factor influencing scaffold display. This was further observed with the relatively lower amount of enzymes binding to *L. lactis* displaying NucA-coh1-coh2-cwa ($1.5x10^3$ UidAdock1/cell). Essentially, NucA-coh1-coh2-cwa is of similar size to NucA-CBM3a-coh3-cwa (approx. 68 kDa), contains twice as many cohesins, yet host cells were able to bind one quarter the amount of UidA-dock1 molecules. The predicted molecular weights of the engineered scaffolds were used in order to estimate the net amount of recombinant protein on the cell surface of strains producing scaffolds with a single cohesin. The culture producing the highest net yield of functional recombinant protein was the strain anchoring NucA-CBM3a-coh3-cwa on its surface. Cultures produced and displayed approximately 0.72 mg/mL of recombinant scaffolds, which remained cell-associated and fully functional. The effect of the N-terminal nuclease reporter on secretion efficiency was also analyzed by comparing the binding capacity of *L. lactis* harboring the pAW300 series (nuclease fusions) with cells harboring the pAW500 (nuclease deficient) series of vectors. Initially included as a reporter to facilitate detection of exported scaffolds, I hypothesized that the nuclease fusion might also increase secretion efficiency, as has been previously observed (Dieye, et al., 2003; Ribeiro, et al., 2002). Removal of NucA had no detrimental effects on scaffold display for all constructs (Fig. 8B), as similar amounts of anchor-containing scaffolds were

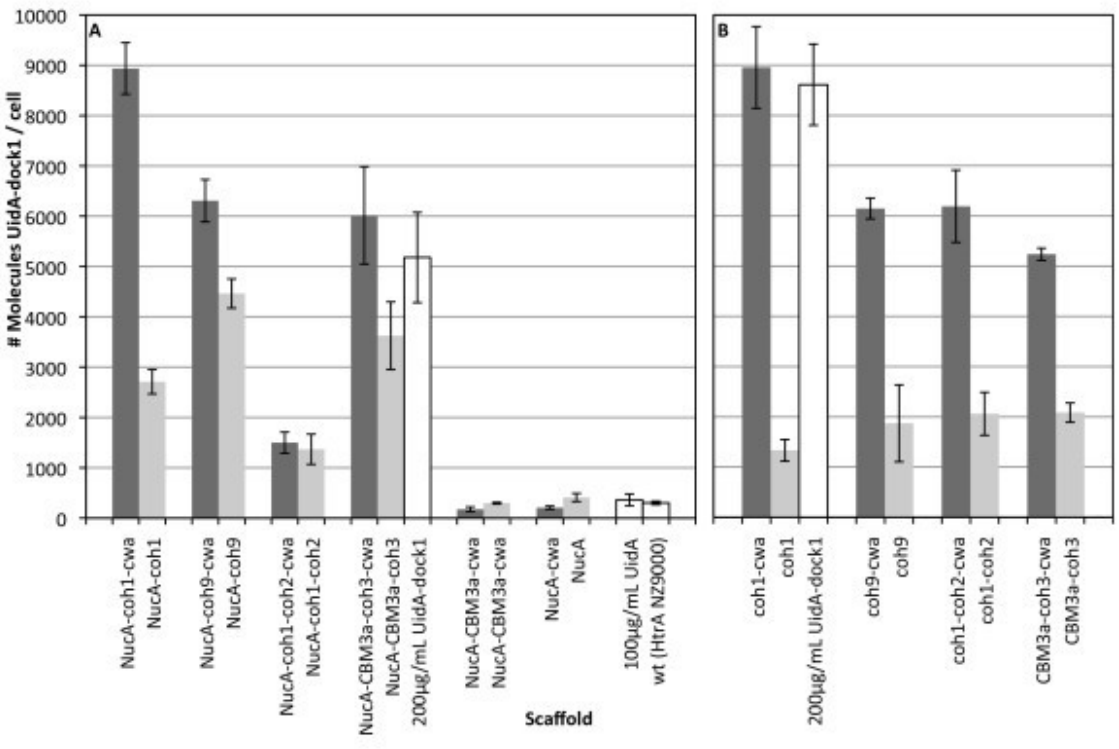

Figure 8. *In vivo* binding of UidAdock1 on live intact *L. lactis* cells displaying CipA$_{frags}$. CipA$_{frags}$ were expressed and anchored as fusions with the NucA reporter enzyme **(A)**, or lacking the NucA reporter **(B)**. Quantification of UidAdock1 molecules bound to *L. lactis* cells corresponds to equivalent amounts of functional cohesin assuming a 1:1 ratio of dockerin-cohesin association. Dark grey bars represent scaffolds containing the C-terminal M6 cell wall anchor motif (cwa), and light grey bars represent their anchor-deficient derivatives. White bars correspond to indicated controls; "200 µg/mL UidA-dock1" represents binding assay carried out with excess enzyme and *L. lactis* pAW328 (NucA-CBM3a-coh3-cwa) to ensure saturation of cohesins. "100µg/mL UidA" represents binding assay carried out in the presence of UidA and *L. lactis* pAW328 (NucA-CBM3a-cwa). Binding assay carried out with UidA and all other constructs resulted in no association with scaffold-expressing strains (data not shown).

located to the cell surface. Furthermore, removal of NucA resulted in a fourfold increase in the amount of coh1-coh2-cwa successfully displayed when compared to its NucA-containing counterpart. The presence of NucA appeared to interfere with the secretion of supernatant-targeted scaffolds from the cell, given that the cwa-deficient variants of coh1, coh9, and CBM3a-coh3 remained associated with the cell to a much larger extent than their NucA-deficient counterparts (Fig. 8).

2.4 Discussion

Several recent studies have reported on the recombinant expression of mini cellulosome scaffold proteins in *Saccharomyces cerevisiae* (Ito, et al., 2009; Lilly, et al., 2009; Tsai, et al., 2009; Wen, et al., 2010). In these examples, the potential application of the engineered strains for the direct conversion of cellulosic biomass to ethanol was the driving factor for choosing *S. cerevisiae* as a host. However, many more platform strains have been or are now being developed that will produce ethanol, biofuels other than ethanol, and non-biofuel chemicals (Atsumi, et al., 2008; Dueber, et al., 2009; S. K. Lee, et al., 2008; L. R. Lynd, et al., 2005; Rittmann, et al., 2008; Rogers, et al., 2007; Shaw, et al., 2008; Steen, et al.; Wu, et al., 2008; M. Zhang, et al., 1995). The economics of these processes would be greatly improved if these engineered microbes could use cellulosic substrates rather than other more easily fermentable carbon sources such as starch, which is also a valuable asset in competing food industries. With this goal in mind, the first logical step in establishing this was the successful secretion and display of cohesin-bearing scaffold proteins. Previous studies have demonstrated that controlled gene expression in *L. lactis* can reduce toxicity and increase net protein yields (Bermudez-Humaran, et al., 2004; de Vos, 1999; Narita, et al., 2006). In my research, the constitutive expression of the scaffold proteins consistently led to cellular toxicity, a problem that was solved by delaying the onset of gene expression until the cells had reached mid log-phase. In cell division, higher concentrations of recombinant cell wall-targeted proteins are localized to the septum, the site of cell wall biosynthesis (Narita, et al., 2006). It is thus likely that over-expression of my scaffold proteins targeted to the extracytoplasmic

space early in the growth phase impaired cell wall biosynthesis and ultimately resulted in cell death. Removal of NucA from the scaffolds decreased or eliminated cellular toxicity for all cohesin-containing constructs (Fig. 6), and I thus suspect that accumulation of NucA in the cytoplasm may also contribute to this observed lag in the onset of growth when induced at t = 0hrs. In addition, as a larger proportion of scaffolds lacking a cell wall anchor remained trapped in the cell wall when fused with NucA, it is also likely that part of this observed reduction in toxicity is due to a decrease in the amounts of recombinant proteins being trapped in the cell wall and ultimately disrupting its integrity.

Quantification of cell surface displayed proteins in lactic acid bacteria was previously reported using fluorescence-activated cell sorting, flow cytometry, or whole-cell ELISA (Leenhouts, et al., 1999). In my assay, functionality of the displayed $CipA_{frag}$ scaffold proteins could be tested directly through binding with a dockerin-containing reporter enzyme, attesting that the number of cohesins detected was a direct quantification of those that retained biochemical function. Preliminary qualitative *in vitro* binding assays were also performed in order to verify the functionality of the cohesin-dockerin interaction (See Appendix A), while the subsequent *in vivo* binding assays were used to estimate relative amounts of scaffolds successfully displayed on the cell surface. Of the four expressed CipA fragments containing at least one cohesin (coh1, coh9, coh1-coh2, CBM3a-coh3), coh1 was displayed with the highest efficiency (~$9x10^3$ scaffolds per cell). Due to its small size and decreased number of domains compared with coh1-coh2 and CBM3a-coh3, it is possible that part of this increase in display is due to the smaller size of the scaffold itself. However, coh1 was also displayed more efficiently than coh9,

which is approximately the same size and similar in primary amino acid sequence. One possible explanation may relate to the position of coh1 relative to coh9 on native CipA scaffold. Coh1 is located at the N-terminus of the 200 kDa scaffold CipA, adjacent to the processing site of the signal peptide by the sec-pathway machinery of *C. thermocellum* (Schwarz, 2001). It is possible that the increase in secretion efficiency of coh1 when compared with coh9 may be in some part due to differences in amino acid content adjacent to the signal peptide, possibly increasing its accessibility to the chaperones involved in its transport to the extracytoplasmic space (Gerngross, et al., 1993). This, however, does not account for the differences in display between NucA-coh1 and NucA-coh9, as in both cases, NucA is downstream to the signal sequence, as opposed to a cohesin. The amount of sequence identity among cohesins perhaps provides a better explanation for these observed differences. Of the nine cohesin domains on CipA, cohesins 3 through 8 show between 96 to 100% sequence identity, whereas among the remaining cohesins, coh1 and coh9 show the least amount of sequence identity (69 and 75%, respectively) (Lytle, et al., 1996). These differences in amino acid content may translate into differences in folding and solubility of the recombinantly expressed domains.

L. lactis was engineered to display a scaffold containing 2 cohesin domains (coh1-coh2). Based on a 1:1 binding ratio of the enzyme-cohesin and assuming equivalent expression and secretion, I expected this strain to bind twice the amount of UidA when compared to scaffolds of similar size but containing a single cohesin domain (i.e. CBM3a-coh3). However, coh1-coh2 bound similar amounts of UidA as CBM3a-coh3 (Fig 8B). This

reduction in UidA binding was not attributed to $CipA_{frag}$ size differences, since both mature scaffolds have a theoretical molecular weight of 68 kDa, suggesting that other factors affected secretion and display efficiency. In fact, protein size is not regarded as a major bottleneck for protein secretion in *L. lactis*, as the size of successfully secreted heterologous proteins ranges from 6.9 kDa to a staggering 165 kDa (Le Loir, et al., 2005). I hypothesize that the substitution of a cohesin domain by CBM3a may have enhanced secretion by increasing the rate of folding of the scaffold into its soluble form. A similar effect was recently reported with the fusion of the highly insoluble *Clostridium cellulovorans* cellulase CelL with the CBM of cellulase CelD, which resulted in dramatic increases in its solubility (Murashima, et al., 2003).

Comparisons between amounts of UidA binding to cells expressing $CipA_{frags}$ with or without the cwa_{M6} domain revealed that the cell wall anchor motif significantly increased the amounts of functional scaffolds displayed on the cell (Fig. 8). With NucA present, $CipA_{frags}$ lacking cwa_{M6} remained associated with whole cells to a larger extent (Fig. 7) and bound UidA (Fig 8), suggesting that NucA fusion proteins remained trapped in the cell wall by non-specific interactions rather than covalent cross-linking by the sortase, and the cohesin domains were accessible to UidA. This phenomenon is well-documented in other studies of protein secretion in *L. lactis*, as in some cases the fusion of two generally well-secreted proteins results in changes in the folding of the hybrid protein, and deficiencies in their release from cells (Bermudez-Humaran, et al., 2002; Miyoshi, et al., 2002). While the exact mechanism of this phenomenon is not clear, hydrophobic

domains resulting from fusing two recombinant proteins may promote cell wall association (Miyoshi, et al., 2002).

2.5 Conclusions

Until now, all attempts to anchor enzymes on the surface of a bacterium have been limited to a single enzyme per anchor (Avall-Jaaskelainen, et al., 2003; Cortes-Perez, et al., 2005; Dieye, et al., 2003; Dieye, et al., 2001; Leenhouts, et al., 1999; Lindholm, et al., 2004; Narita, et al., 2006; Piard, et al., 1997; Raha, et al., 2005; Ramasamy, et al., 2006; Ribeiro, et al., 2002; Yang, et al., 2008). In my system, multiple enzymes could theoretically associate with scaffolds containing a corresponding number of cohesins. I used purified UidA fused to a dockerin domain as a probe to establish proper display and function of the cohesins, but also generated a library of fluorescent probes to target cohesin domains by fusing them with an appropriate dockerin (See Appendix B). I also envision co-expression of enzymes and scaffold in a subsequent development of the strain. I thus envision that further development of this cellulosome-inspired system may contribute to the efficient bioconversion of substrates into industrially relevant fuels and commodity chemicals, and that tailor-designed synthetic macromolecular complexes could be engineered to contain large permutations and combinations of desired enzymes of interest.

Chapter 3

Engineering the cell surface display of cohesins for assembly of cellulosome-inspired enzyme complexes on *Lactococcus lactis*

Wieczorek, A.S. and V.J. Martin, *Engineering the cell surface display of cohesins for assembly of cellulosome-inspired enzyme complexes on Lactococcus lactis*. (Submitted for Publication to Microb Cell Fact. 2012)

3.1 Introduction

Chapter 2 of my thesis reported on the successful display of scaffold proteins containing type 1 cohesins on the surface of *L. lactis* (Wieczorek & Martin, 2010). However, from a biotechnological perspective, the ability to control the spatial organization and stoichiometry of the enzymes is desirable. Engineering scaffold proteins containing cohesins with different specificities for binding dockerin-encoding enzymes provides a means to dictate the architecture of recombinant complexes. I therefore chose to engineer strains of *L. lactis* capable of displaying chimeric scaffold proteins resulting from the fusion of $CipA_{frags}$ encoding type 1 cohesins with fragments of OlpB ($OlpB_{frags}$) or SdbA ($SdbA_{frags}$) encoding type 2 cohesins. The most complex scaffolds contained type 1 and type 2 cohesins, as well as a cellulose-binding domain (CBD), and were composed solely of building blocks of the *C. thermocellum* cellulosome. The effect of protein scaffold architecture on secretion and display, functionality, and ability to bind dockerin-fused enzymes was investigated by using or excluding linker sequences between cohesins, by varying the number and origin of cohesins in the chimeric scaffold, and by changing the order in which enzymes were localized within the complex. The

dockerin-fused reporter enzymes were UidA and *E. coli* β-galactosidase (LacZ). The specificity and efficiency of binding of reporter enzymes to each of the synthetic scaffolds was tested. In addition, I investigated the effects of assembling bi-enzymatic complexes via the simultaneous or sequential addition of enzyme components on the surface-displayed chimeric scaffolds.

3.2 Materials and Methods

3.2.1 Bacterial strains and plasmids used in this chapter

The bacterial strains and plasmids used in this study are listed in Table 4. *E. coli* strains were grown in Luria-Bertani medium at 37°C with shaking (220 rpm). *Lactococcus lactis htrA*-NZ9000 was grown in M17 medium (Terzaghi & Sandine, 1975) supplemented with 1% (w/v) glucose (GM17) at 30°C without agitation. To make competent cells, *L. lactis* was grown in GM17 medium supplemented with 25% (w/v) sucrose and 2% (w/v) glycine and cells were transformed as previously described (Holo & Nes, 1989). *C. thermocellum* was grown in ATCC1191 medium at 55°C with 0.2% (w/v) cellobiose as a carbon source. Genomic DNA was isolated from *C. thermocellum* as previously described (Wang & Wu, 1993). Where appropriate, antibiotics were added as follows: for *E. coli*, ampicillin (100 μg/mL), chloramphenicol (10 μg/mL) and kanamycin (30 μg/mL); for *L. lactis,* erythromycin (5 μg/mL) and chloramphenicol (10 μg/mL). General molecular biology techniques for *E. coli* were performed as previously described (Sambrook & Russell, 2001).

Table 4 - Strains and plasmids used in this chapter. pAW500 series of vectors are designed for the cell-wall targeting of various scaffold protein permutations consisting of cohesins from CipA and OlpB or SdbA. *coh1C*, type 1 cohesin of CipA; *coh2O2*, type 2 cohesin of OlpB; *coh2S1*, type 2 cohesin of SdbA; *PT7*, inducible T7 promoter; P_{nisA}, inducible *nisA* promoter; rbs_{usp45}, Usp45 ribosome-binding site; rbs_{nisA}, *nisA* ribosome-binding site; sp_{Usp45}, signal sequence of Usp45; cwa_{M6}, anchor motif of M6 protein; *tlt2*, transcriptional terminator of *rrnB* operon; t_{trpA}, transcriptional terminator of *trpA*; *Lk, olpB* linker region.

Strain	Genotype / Decription	Source
L. lactis htrA-NZ9000	Mutant MG1363 derivative (*nisRK* genes on the chromosome) lacking *htrA*	(Miyoshi, et al., 2002)
E. coli TG1	*supE thi-1* Δ(*lac-proAB*) Δ(*mcrB-hsdSM*)*5* (rK– mK–) [F′ *traD36 proAB lacI*qZΔ*M15*]	ATCC
E. coli BL21 (DE3)	F^- *ompT gal dcm lon* $hsdS_B(r_B^- m_B^-)$ *λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])*	Novagen
Plasmid		
pET28(b)	Kn^r	Novagen
pAW507	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{usp45}::sp_{Usp45}-*coh1C9-*cwa_{M6}-*tlt2*	(Wieczorek & Martin, 2010)
pAW510	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{usp45}::sp_{Usp45}-*coh1C1-*cwa_{M6}-*tlt2*	(Wieczorek & Martin, 2010)
pAW528	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}- *CBD-coh1C3- -*cwa_{M6}-*tlt2*	(Wieczorek & Martin, 2010)
pAW531	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*coh1C1-coh1C2-*cwa_{M6}-*tlt2*	(Wieczorek & Martin, 2010)
pAW534	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*CBD-*cwa_{M6}-*tlt2*	(Wieczorek & Martin, 2010)
pAW549	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*CBD-coh2O2-*cwa_{M6}-*tlt2*	This Work

pAW564	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*CBD-Lk-coh2O2-cwa*$_{M6}$-*tlt2*	This Work
pAW596	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*CBD-coh2O2-Lk-cwa*$_{M6}$-*tlt2*	This Work
pAW594	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*CBD-Lk-coh2O2-Lk-cwa*$_{M6}$-*tlt2*	This Work
pAW546	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*CBD-coh1C3-coh2O2-cwa*$_{M6}$-*tlt2*	This Work
pAW561	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*CBD-coh1C3-Lk-coh2O2-cwa*$_{M6}$-*tlt2*	This Work
pAW595	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*CBD-coh1C3-coh2O2-Lk-cwa*$_{M6}$-*tlt2*	This Work
pAW592	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*CBD-coh1C3-Lk-coh2O2-Lk-cwa*$_{M6}$-*tlt2*	This Work
pAW579	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*CBD-coh2S1-cwa*$_{M6}$-*tlt2*	This Work
pAW576	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*CBD-coh1C3-coh2S1-cwa*$_{M6}$-*tlt2*	This Work
pAW570	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*coh1C1-coh2S1-cwa*$_{M6}$-*tlt2*	This Work
pAW567	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*coh1C9-coh2S1-cwa*$_{M6}$-*tlt2*	This Work
pAW573	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*coh1C1-coh1C2-coh2S1-cwa*$_{M6}$-*tlt2*	This Work
pETdock1	Kn^r; pET28(b)::with cloned *dock1* from *celS*	(Wieczorek & Martin, 2010)
pETdock2	Kn^r; pET28(b)::with cloned *dock2* from *cipA*	This Work
pETUdock1	Kn^r; pET28(b)::*PT7*::*6xHis-uidA-dock1*	(Wieczorek & Martin, 2010)
pETUdock2	Kn^r; pET28(b)::*PT7*::*6xHis-uidA-dock2*	This Work

pETU	Kn^r; pET28(b)::*PT7*::*6xHis-uidA*	(Wieczorek & Martin, 2010)
pETLdock1	Kn^r; pET28(b)::*PT7*::*6xHis-lacZ-dock1*	This Work
pETLdock2	Kn^r; pET28(b)::*PT7*::*6xHis-lacZ-dock2*	This Work
pETL	Kn^r; pET28(b)::*PT7*::*6xHis-lacZ*	This Work

3.2.2 Assembly of chimeric scaffolds expression cassettes

The *E. coli-L. lactis* shuttle vectors pAW507, pAW510, pAW528, pAW531, and pAW534 all contain gene expression cassettes for the secretion and surface display of the scaffold proteins (Table 4) (Wieczorek & Martin, 2010). Scaffolds are expressed as fusions with the N-terminal signal peptide from the lactococcal Usp45 secreted protein (sp_{Usp45}) and with the C-terminal anchor motif of streprococcal M6 protein cwa_{M6} (Fig. 9). Expression of the cassettes is under the control of the *nisA* nisin-inducible promoter (P_{nisA}) and ribosome-binding site (rbs_{nisA}) from *L. lactis* (Sorvig, et al., 2003). For the construction of cassettes encoding chimeric protein scaffolds, PCR was performed on *C. thermocellum* genomic DNA to amplify regions encoding fragments of the cellulosomal proteins CipA (GenBank accession no. Q06851), OlpB (GenBank accession no. CAA47841.1) and SdbA (GenBank accession no. AAB07763.1). DNA encoding the second cohesin of OlpB *(coh2O2*) was amplified using primers *a* and *b* (Table 5). In order to incorporate the protein linker sequence at the N-terminal end of the Coh2O2 cohesin, *link-coh2O2* was amplified using primers *c* and *b* (Table 5). For incorporation of the C-terminal linker into the recombinant scaffold, *coh2O2-link* was amplified using primers *a*

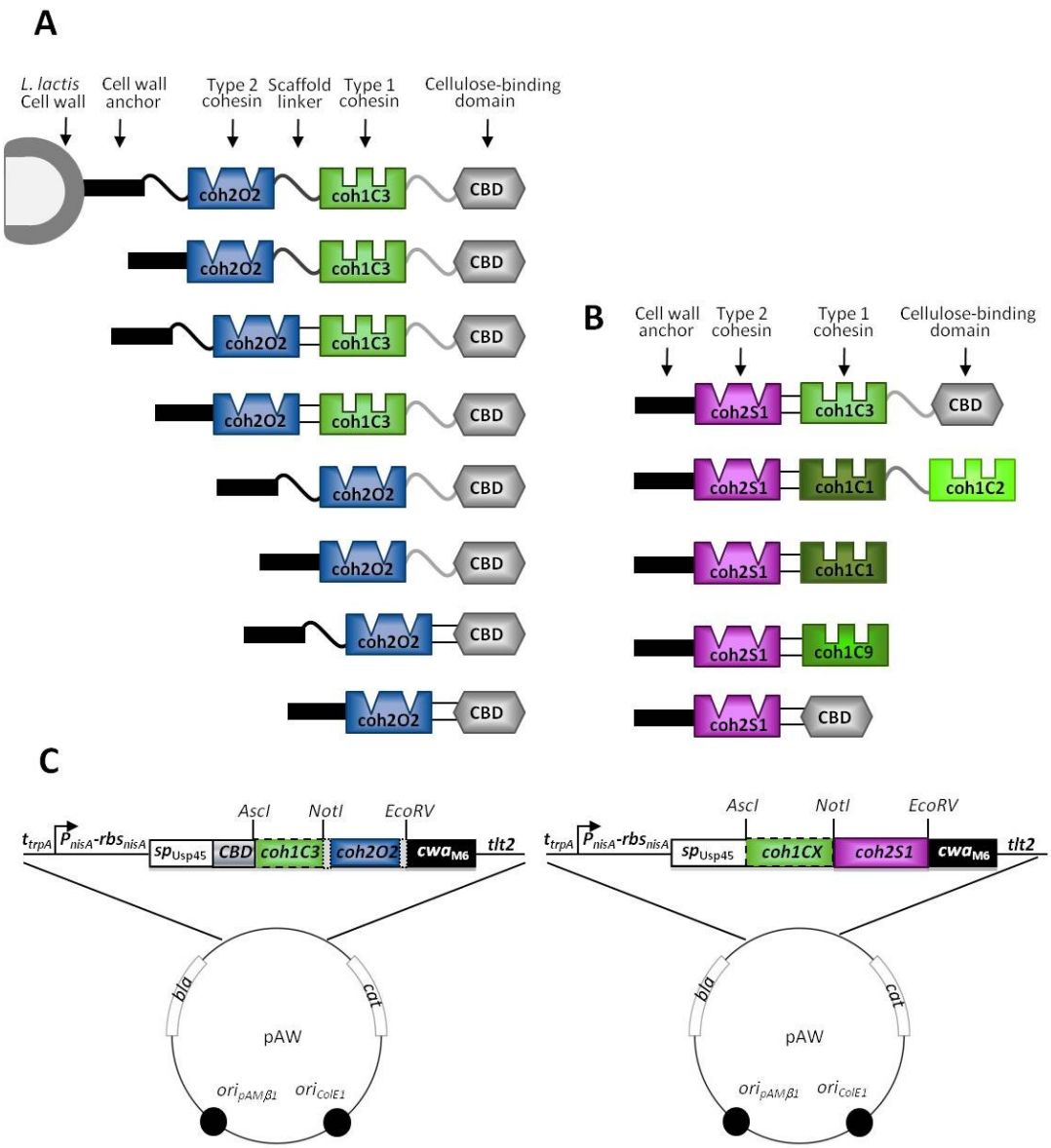

Figure 9. Depiction of chimeric scaffold proteins and expression cassettes. (A) Chimeric protein scaffolds generated as fusions of the CipA type 1 cohesin coh1C3 (green) with the OlpB type 2 cohesin coh2C2 (blue) and cellulose binding domain CBD (grey). Linkers between cohesin domains, the cell anchor, or the CBD are derived from OlpB (black) or CipA (grey). Double lines represent direct fusion of two domains without a linker sequence. **(B)** Chimeric protein scaffolds generated as fusions of CipA type 1 cohesins (each shade of green represents a different cohesin) with the type 2 cohesin of SdbA (purple) and cellulose binding domain CBD (grey). **(C)** Scaffold expression cassettes showing the N-terminal signal peptide from the lactococcal Usp45 secreted protein (sp_{Usp45}) and the cell wall anchor motif of the M6 protein (cwa_{M6}). Expression of the cassettes is under the control of the *nisA* nisin-inducible promoter (P_{nisA}) and ribosome-binding site (rbs_{nisA}) from *L. lactis*. The transcriptional terminators of the *rrnB* operon (*tlt2)* and *trpA* gene (t_{trpA}) are located upstream and downstream of the expression cassette, respectively. DNA encoding optional modules (coh1CX and linkers) is depicted with dotted lines surrounding the module.

and *d* (Table 5). To engineer a scaffold with the Coh2O2 cohesin flanked by two linkers, the *link-coh2O2-link* fragment was amplified using primers *c* and *d* (Table 5). PCR products were purified using a PCR purification kit (Qiagen), digested with *Not*I and *EcoR*I, and ligated into similarly cut pAW528 and pAW531, yielding vectors pAW549 (CBD-coh2O2), pAW546 (CBD-coh1C3-coh2O2), pAW564 (CBD-Link-coh2O2), pAW561 (CBD-coh1C3-Link-coh2O2), pAW596 (CBD-coh2O2-Link), pAW591 (CBD-coh1C3-coh2O2-Link), pAW594 (CBD-Link-coh2O2-Link) and pAW592 (CBD-coh1C3-Link-coh2O2-Link) (Table 4).

Table - 5 Primers used in this chapter (Restriction enzyme cut sites are in bold).

Primer	Sequence (5' – 3')
a	TCGA**GCGGCCGC**GCTGGAACTGGATAAGAC
b	TCGA**GATATC**TTAGGCTGTACTACGCTATAC
c	TCGA**GCGGCCGC**GCTTATAGTTGTAGAGGC
d	ATGC**GATATC**GTCGACTTTATTACATAGGAATCTGGAAG
e	TCGA**GCGGCCGC**GGATAAAGCCTCGAGCATTG
f	TCGA**GATATC**TTATCCGGCTGTATTACCTC
g	ATGC**GAATTC**GGAGACATAGTGAAAGACAATTC
h	ATGC**GCGGCCGC**TTTACTGTGCGTCGTAATCAC
i	ATGC**GCTAGC**ATGACCATGATTACGG
J	GCAT**CAATTG**TTTTTGACACCAGACC

Type 2 cohesin coh2S1 of anchor protein SdbA was PCR-amplified using primers *e* and *f* (Table 5), purified using a PCR purification kit (Qiagen), digested with *Not*I and

*Eco*RV, and ligated to similarly cut pAW507, pAW510, pAW528, pAW531, and pAW534, yielding pAW567 (coh1C9-coh2S1), pAW570 (coh1C1-coh2S1), pAW573 (coh1C1-coh1C2-coh2S1), pAW576 (CBD-coh1C3-coh2S1) and pAW579 (CBD-coh2S1), respectively (Table 4).

3.2.3 Assembly of dockerin-fused UidA and LacZ expression cassettes

E. coli β-glucuronidase (UidA) was previously engineered to contain a C-terminal dock1 domain for binding of the enzyme to type 1 cohesins (Wieczorek & Martin, 2010). In this study, UidA was fused with a dock2 domain from CipA for binding to type 2 cohesins, as well as an N-terminal 6 x His-tag for protein purification. For assembly of the *hisX6-uidA-dock2* cassette, the dock2 sequence of the *cipA* gene was amplified from *C. thermocellum* genomic DNA using primers *g* and *h* (Table 5). The PCR product was digested with *Eco*RI-*Not*I and ligated to similarly-digested pET28(b), yielding pETdock2. To create the UidA-dock2 fusion, pETUdock1 was digested with *Nhe*I-*Eco*RI to isolate the *uidA* gene, which was gel-purified, and ligated to similarly cut pETdock2, yielding pETUdock2. In order to create LacZ-dockerin fusion proteins, DNA encoding the *E. coli* β-galactosidase LacZ (GenBank accession no. EGT70540.1) was PCR amplified from genomic DNA of *E. coli* MG1655 using primers *i* and *j* (Table 5). The resulting PCR product was digested with *Nhe*I-*Mfe*I and ligated into *Nhe*I-*Eco*RI-digested pETU, pETUdock1 and pETUdock2, yielding pETL, pETLdock1 and pETLdock2, respectively (Table 4). All pET vectors described above express cassettes encoding enzymes and enzyme-dockerin fusions with an N-terminal 6XHis tag for purification.

3.2.4 Expression and purification of dockerin-fused UidA and LacZ

All His-tagged enzymes were expressed in *E. coli* BL21(DE3) as previously described (Wieczorek & Martin, 2010). The UidA and LacZ-containing elution fractions were identified by the appearance of a yellow color in a liquid β-glucuronidase and β-galactosidase assay, respectively. Liquid β-glucuronidase assay conditions are previously described (Wieczorek & Martin, 2010). For liquid β-galactosidase assay, 50 µL of each elution fraction were added to 450 µL of Z buffer containing 100 mM phosphate buffer pH7, 5 mM KCl, 1 mM $MgSO_4$, 0.28% (v/v) β-mercaptoethanol. Samples were heated for 1 min, after which *p*-nitrophenyl-β-D-glucuronide was added to a final concentration of 4 mg/mL (Axelsson, et al., 2003). The purity of the elution fractions exhibiting UidA and LacZ activity was assessed by SDS-PAGE (12%, w/v). Proteins were stained using Coomasie Blue Reagent (BioRad) and fractions containing the highest purity of enzyme were pooled. The specific activities of UidA-dock1 and UidA-dock2 were determined by colorimetric assays in a thermostated UV-Vis spectrophotometer (Cary 50 WinUv) at 405 nm, using a 1 cm (L) cuvette, and the molar extinction coefficient of *p*-nitrophenol (PNP) being 18 000 M^{-1} cm^{-1}. A Bradford protein assay kit (Pierce) and BSA as a standard were used in order to quantify net protein amounts, and specific activities were used to evaluate the amount of enzyme bound to cells in the *in vivo* binding assay described below. For simultaneous or sequential binding assays, overall enzymatic activities/cell were calculated by measuring colorimetric changes using *p*-nitrophenyl-β-D-glucuronide as substrate and 405 nm wavelength for UidA activity, and *O*-nitrophenyl-β-galactoside as substrate and 420 nm wavelength for LacZ activity.

3.2.5 Quantitation of UidA-dockerin binding to *L. lactis*-expressed scaffold proteins

L. lactis htrA-NZ9000 was transformed with the expression plasmids encoding permutations of chimeric scaffolds (Fig. 9). The strain is deficient in the HtrA extracellular protease and contains chromosomal copies of the *nisR* and *nisK* genes, which participate in the regulation of expression cassettes under control of the *nisA* promoter (Miyoshi, et al., 2002). *L. lactis* cells harboring the plasmids were grown overnight in GM17 medium and diluted 1/50 in 5 mL of fresh media and grown for an additional 4 hrs (OD_{600} ≈0.3) after which cells were induced with 10 ng nisin/mL for scaffold expression (Wieczorek & Martin, 2010). After 20 hrs growth, 1 mL of cells were washed in phosphate buffer (50 mM, pH 6.0) containing 300 mM NaCl and suspended in 100 μL of purified UidA-dock1 or UidA-dock2 at a concentration of 100 μg/mL. Binding assay conditions and enzyme quantification methods used to determine the amount of enzyme associated with *L. lactis* cells are previously described (Wieczorek & Martin). Using the calculated molecular weights of UidA-dock1 and UidA-dock2 and the known amount of cells present in each sample, the average number of enzyme units bound per cell was estimated. Experiments were performed in triplicate using true biological replicates (independent colonies and cultures).

3.2.6 Simultaneous or sequential binding of UidA- and LacZ-dockerin to cells displaying chimeric protein scaffolds

Enzyme combinations consisting of equimolar amounts of UidA-dock1 and LacZ-dock2 or UidA-dock2 and LacZ-dock1 were mixed to a final enzyme concentration of 100

μg/mL. Cells were incubated in 100 μL of the enzyme mixture, washed 6 times in phosphate buffer (50 mM, pH 6) containing 300 mM NaCl, re-suspended in 100 μL of the same buffer, and analyzed using both the β-glucuronidase assay (Wieczorek & Martin, 2010) and β-galactosidase assay.

For sequential binding assays, cells were incubated with a first test enzyme at a concentration of 100 μg/mL, or no enzyme. After 5 hrs of incubation at 4°C, cells were harvested by centrifugation and suspended in 100 μL phosphate buffer (50 mM, pH 6) containing 300 mM NaCl and an equimolar amount of the second enzyme or no enzyme. After an additional 5 hrs of incubation, cells were harvested, washed 6 times in phosphate buffer (50 mM, pH 6) containing 300 mM NaCl, suspended in 100 μL of the same buffer, and tested for both β-glucuronidase and β-galactosidase activity.

3.3 Results

3.3.1 UidA-dock1 binds to coh2O2-coh1C3 chimeric proteins displayed on *L. lactis*

Chimeric scaffold proteins containing cohesins of different specificity were expressed as fusions with the N-terminal signal peptide from the lactococcal Usp45 secreted protein (sp_{Usp45}) (van Asseldonk, et al., 1990) and under control of the *nisA* nisin-inducible promoter (P_{nisA}) and ribosome-binding site (rbs_{nisA}) from *L. lactis* (Kuipers, et al., 1993; Mierau & Kleerebezem, 2005). For simplicity of scaffold nomenclature, the number preceding the uppercase letter represents the type of cohesin (type 1=coh1 and type

2=coh2), the uppercase letter represents the protein of origin (CipA=C, OlpB=O and SdbA=S) and the number proceeding the uppercase letter represents the relative position of the cohesin from the N-terminus of the protein of origin. The first chimeric protein scaffold architecture tested in *L. lactis* consisted of the cellulose binding domain (CBD), the third type 1 cohesin domain of CipA (coh1C3), as well as the second type 2 cohesin domain of OlpB (coh2O2). My previous work suggested the possibility that the CBD may aid in the secretion of larger scaffolds, so it was included in all coh2O2 fusions (Wieczorek & Martin, 2010). In order to investigate the effects of including linkers between scaffold domains on the efficiency of enzyme binding, the chimeric proteins were constructed with and without linker sequences between the two cohesins and between the cohesin and the cell wall anchor domain (Fig. 9A). *In vivo* binding assays were used to show binding of the dockerin-containing β-glucuronidase (UidA-dock1) to the scaffold and to verify the specificity of dockerin 1 (dock1) to coh1C3 interaction. The dock1 domain used was derived from cellulosomal enzyme CelS, which is capable of binding any of the nine type 1 cohesins of CipA (Lytle, et al., 1996). Cells displaying chimeric scaffolds containing coh1C3 were capable of binding to UidA-dock1 (Fig. 10, black bars), demonstrating the functionality of coh1C3 within recombinant chimeric scaffolds. Wild type UidA is tetrameric, and I considered, due to a 1:1 cohesin:dockerin binding ratio, that only one dockerin of each tetramer would bind a single cohesin of the same type on the chimeric scaffold. Minor differences in enzyme binding were observed based on scaffold architecture. The inclusion of linkers at the C-terminus of coh1C3 seemed to have no significant effect on UidA-dock1 binding. A scaffold containing a single cohesin (CBD-

coh1C3) was used as a reference point and showed that the addition of the coh2O2 cohesin to the synthetic scaffold reduced UidA-dock1 binding by ~two-fold (Fig. 10). A strain displaying CBD alone or CBD-coh2O2 was also used as negative control and failed to bind the UidA-dock1 reporter enzyme (Fig. 10).

To test the functionality of the coh2O2 domain within the chimeric scaffolds, similar binding assays were carried out using UidA fused to a type 2 dockerin domain isolated from CipA (UidA-dock2). It has been previously demonstrated that OlpB is surface displayed on *C. thermocellum* and successfully binds the dock2 domain of CipA (Lemaire, et al., 1995; J. Xu & Smith, 2010). Surprisingly, the chimeric scaffolds containing the coh2O2 domain did not bind UidA-dock2 (Fig 10, white bars). All scaffolds lacking coh2O2 also failed to bind UidA-dock2. From these results, I hypothesized that either coh2O2 or dock2 were incapable of folding into their functional form when fused with CBD (alone or fused to coh1C3) or UidA, respectively. Since all tri-modular chimeric proteins did successfully bind UidA-dock1, I concluded that the lack of interaction of UidA-dock2 with the scaffolds was not due to a complete lack of expression and secretion of the chimeric scaffold proteins. Background β-glucoronidase activity was slightly higher when using UidA-dock2 than when using UidA-dock1 (Fig. 10). This residual β-glucoronidase activity can be attributed to a slightly higher non-specific adherence of UidA-dock2 to cells since binding of the fusion protein to the plasmid-free *L. lactis* strain showed similar levels of activity (data not shown). Based on these results, I sought to test if substitution of the type 2 cohesin domain in the chimeric scaffold would result in successful binding to UidA-dock2.

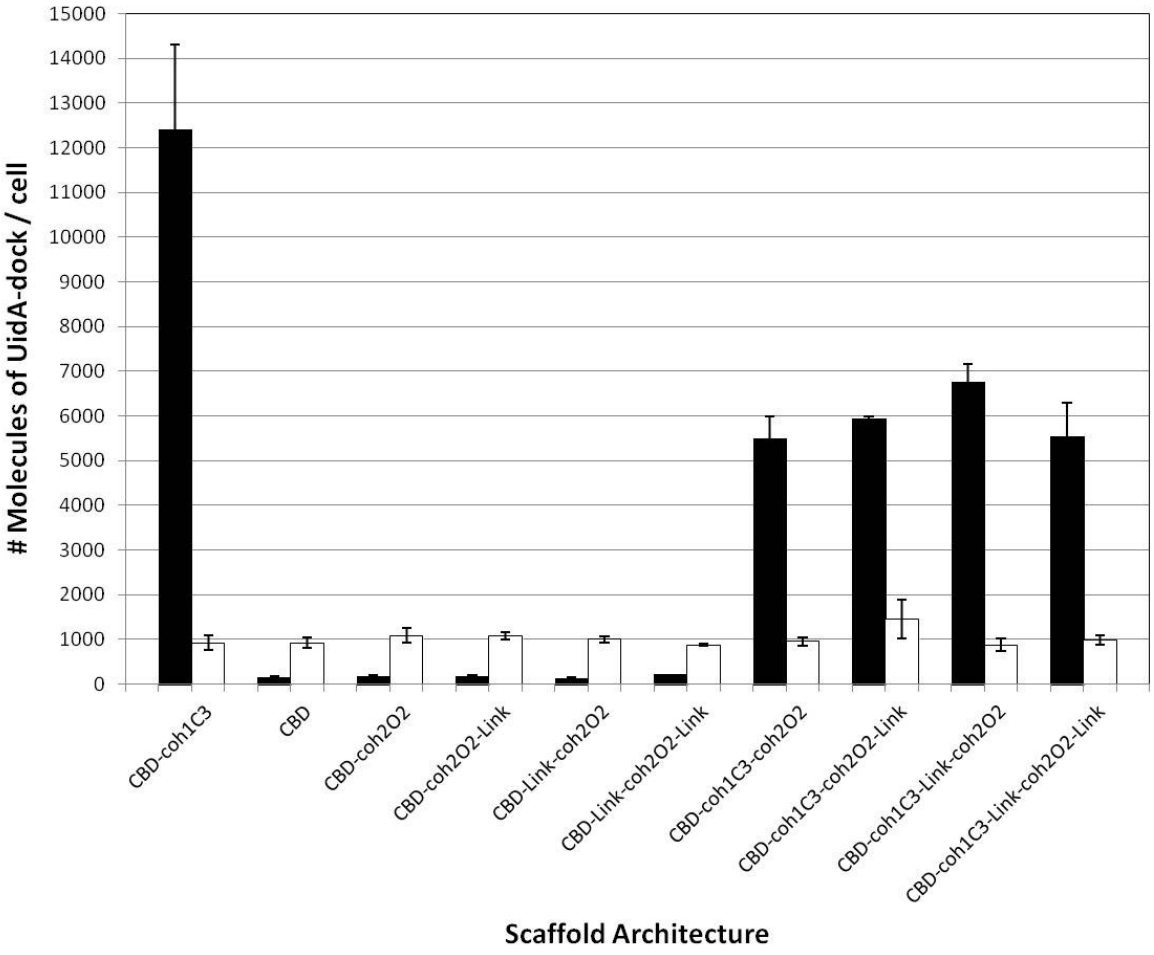

Figure 10. *In vivo* binding of UidA-dock1 and UidA-dock2 on *L. lactis* cells displaying $CipA_{frag}$–$OlpB_{frag}$ chimeric scaffold proteins. Cells displaying chimeric scaffolds were tested for their ability to bind UidA-dock1 or UidA-dock2. Scaffolds were comprised of a type 1 cohesin domain, a type 2 cohesin domain, both a type 1 and type 2 cohesin domain, or no cohesin domain. Quantification of enzymes was carried out using the calculated specific activity of purified enzyme, and the known amount of cells in each sample. The number of molecules bound to *L. lactis* cells corresponds to equivalent amounts of functional cohesin assuming a theoretical 1:1 ratio of dockerin to cohesin binding. Bars represent the number of UidA-dock1 molecules (black) and UidA-dock2 molecules (white) successfully associated with the scaffolds.

3.3.2 coh1CX-coh2S1 chimeric scaffolds bind UidA-dockerin fusion proteins

Scaffolds engineered by replacing coh2O2 with the type 2 cohesin of SdbA (coh2S1) (Fig. 9) were capable of binding both UidA-dock1 and UidA-dock2 (Fig. 11), establishing the functionality of coh2S1 domain incorporated into the chimeric scaffold protein. The coh2S1 domain was fused with fragments of CipA containing a CBD and/or coh1CX, where "X" refers to cohesin 1, 2, 3, or 9 of CipA. Substituting coh2O2 for coh2S1 therefore greatly improved UidA-dock2 binding to cells. Both OlpB and SdbA are anchor proteins that are responsible for binding CipA to the surface of *C. thermocellum*; however, these proteins have two striking differences. First, SdbA contains a single type 2 cohesin rather than four, and second, it contains a unique lysine-rich region at the C-terminus of the coh2S1 domain, which shows a high degree of homology to the streptococcal M proteins (Leibovitz & Beguin, 1996). Both UidA-dockerin fusion proteins were able to bind the CBD-coh1C3-coh2S1 chimeric scaffold protein, demonstrating the functionality of both cohesin domains (Fig. 11). To improve upon those results, alternative coh2S1-containing scaffolds were constructed containing either a single type 2 cohesin (CBD-coh2S1), both a type 1 and type 2 cohesin (coh1C1-coh2S1, coh1C9-coh2S1), or two type 1 cohesins and one type 2 cohesin (coh1C1-coh1C2-coh2S1) (Fig. 9B). The scaffold fragments chosen for fusion with coh2S1 were based on previous

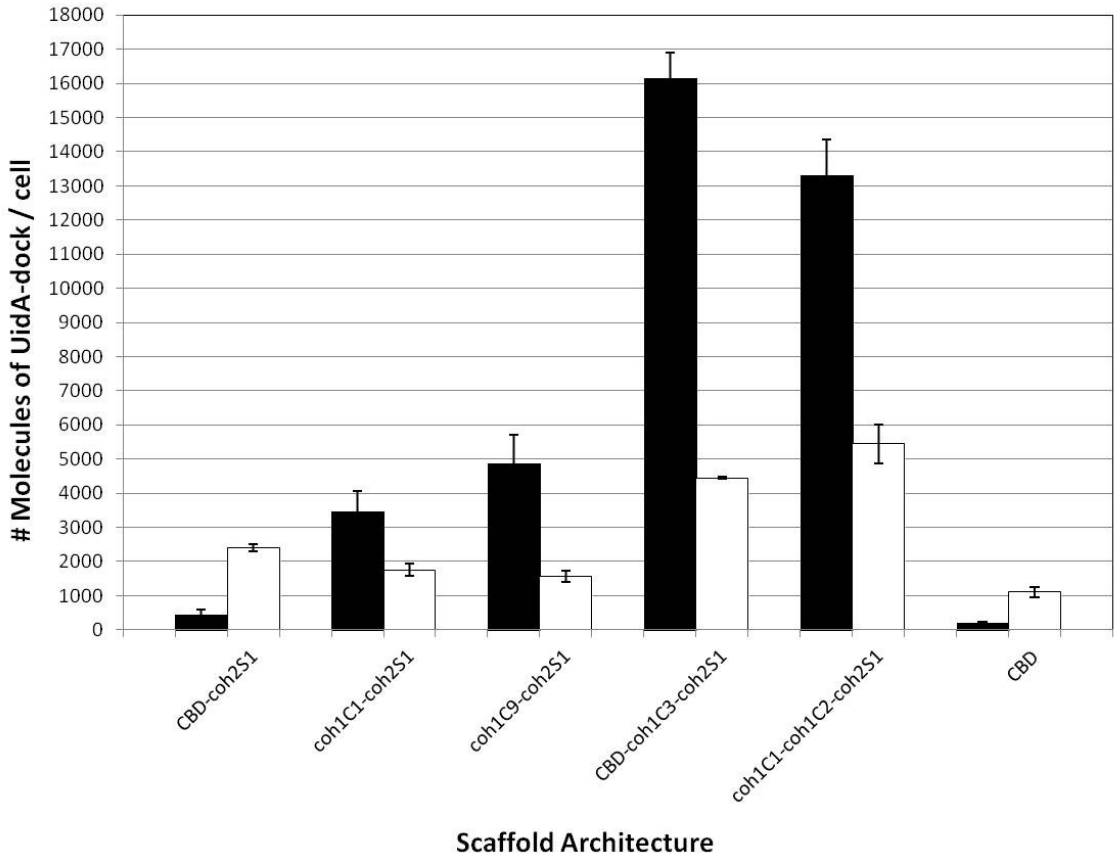

Figure 11. *In vivo* binding of UidA-dock1 and UidA-dock2 on *L. lactis* cells displaying coh1CX-coh2S1 chimeric scaffold proteins. Cells displaying chimeric scaffold proteins were tested for their ability to bind UidA-dock1 or UidA-dock2. Quantification of enzymes was carried out using identical methods as described in the legend of figure 2. Bars represent the number of UidA-dock1 molecules (black bars) and UidA-dock2 molecules (white bars) successfully associated with the scaffolds.

success in their secretion and surface-display (Wieczorek & Martin, 2010). All chimeric scaffolds were capable of binding UidA-dock1, however differences in binding efficiencies were observed. The cells that bound the greatest number of UidA-dock1 molecules were those displaying the larger tri-modular chimeras (CBD-coh1C3-coh2S1 and coh1C1-coh1C2-coh2S1). These strains bound 1.6 x 10^4 and 1.3 x 10^4 molecules of UidA-dock1, respectively (Fig. 11). While I cannot differentiate between binding of UidA-dock1 with

coh1C1 vs. coh1C2 in the latter construct, it is possible each of the two type 1 cohesins may successfully bind an individual reporter enzyme, partially contributing to the observed β-glucuronidase activity. The smaller scaffold coh1C1-coh2S1 bound significantly ($p<0.05$) lower amounts of UidA-dock1 molecules. As expected, UidA-dock1 did not bind to cells expressing CBD-coh2S1. Similarly, scaffold binding assays with UidA-dock2 showed higher UidA-dock2 binding to cells expressing the larger chimeric scaffolds CBD-coh1C3-coh2S1 and coh1C1-coh1C2-coh2S1 when compared to strains expressing the smaller coh1C1-coh2S1 scaffold. A strain expressing coh1C9-coh2S1 did not bind a significant ($p>0.05$) amount of UidA-dock2 above the negative control expressing CBD (Fig. 11). Strains expressing the bi-modular scaffold CBD-coh2S1 were also capable of binding UidA-dock2 (Fig. 11).

3.3.3 coh1C3-coh2S1 chimeric scaffold binds LacZ-dockerin fusion proteins

Having demonstrated the functionality of cell-displayed tri-modular synthetic scaffolds in binding a single enzyme, I sought to test the versatility of the scaffolds by binding a much larger enzyme. The *E. coli* β-galactosidase was fused to dock1 (LacZ-dock1) or dock2 (LacZ-dock2), and the resulting enzyme fusions were tested for their ability to bind the chimeric scaffolds. The LacZ-dockerin fusions were tested for their ability to bind cells displaying chimeric scaffold CBD-coh1C3-coh2S1 or CBD alone. Similar to UidA, LacZ is tetrameric, and I considered for quantification purposes, that due to a 1:1 cohesin:dockerin binding ratio only one dockerin would bind a single cohesin of the same type on the chimeric scaffold. LacZ-dock1 and LacZ-dock2 were both capable of

binding the coh1C3 and coh2S1 sites on the chimeric scaffolds, respectively, and did not bind CBD (Fig. 12B, D). This clearly indicated that much like the UidA-dockerin fusion proteins, the dockerin-containing LacZ was binding to its corresponding cohesin partner. LacZ lacking a dockerin domain did not bind to any of the strains described (data not shown). Having confirmed the functionality of the UidA and LacZ dockerin fusions, as well as their ability to bind to the chimeric scaffolds I sought to further probe the versatility of the scaffolds by binding two enzymes, simultaneously or sequentially.

3.3.4 Simultaneous binding of UidA- and LacZ-dockerin fusions to chimeric protein scaffolds

The incubation of cells displaying the CBD-coh1C3-coh2S1 scaffold with an enzyme mixture consisting of equimolar amounts of both UidA-dock1 and LacZ-dock2 resulted in the successful assembly of a two-enzyme complex tethered to the surface of *L. lactis*. These results demonstrated that the architecture of the synthetic scaffold could accommodate both enzymes at the respective coh1C3 and coh2S1 sites. Comparisons in activity were made when each enzyme was targeted to the displayed scaffold independently, or when the two enzymes were bound simultaneously. Binding UidA-dock1 to the coh1C3 domain on the scaffold resulted in increased activity when compared with UidA-dock2 binding to the coh2S1 domain (Fig. 12A, C), and this result was also observed when binding LacZ-dock1 and LacZ-dock2 to these same cohesin domains (Fig. 12B, D). Since in my construct, coh1C3 is closer to the N-terminus of the protein, whereas coh2S1 is adjacent to the C-terminal cwa, it is possible that a greater

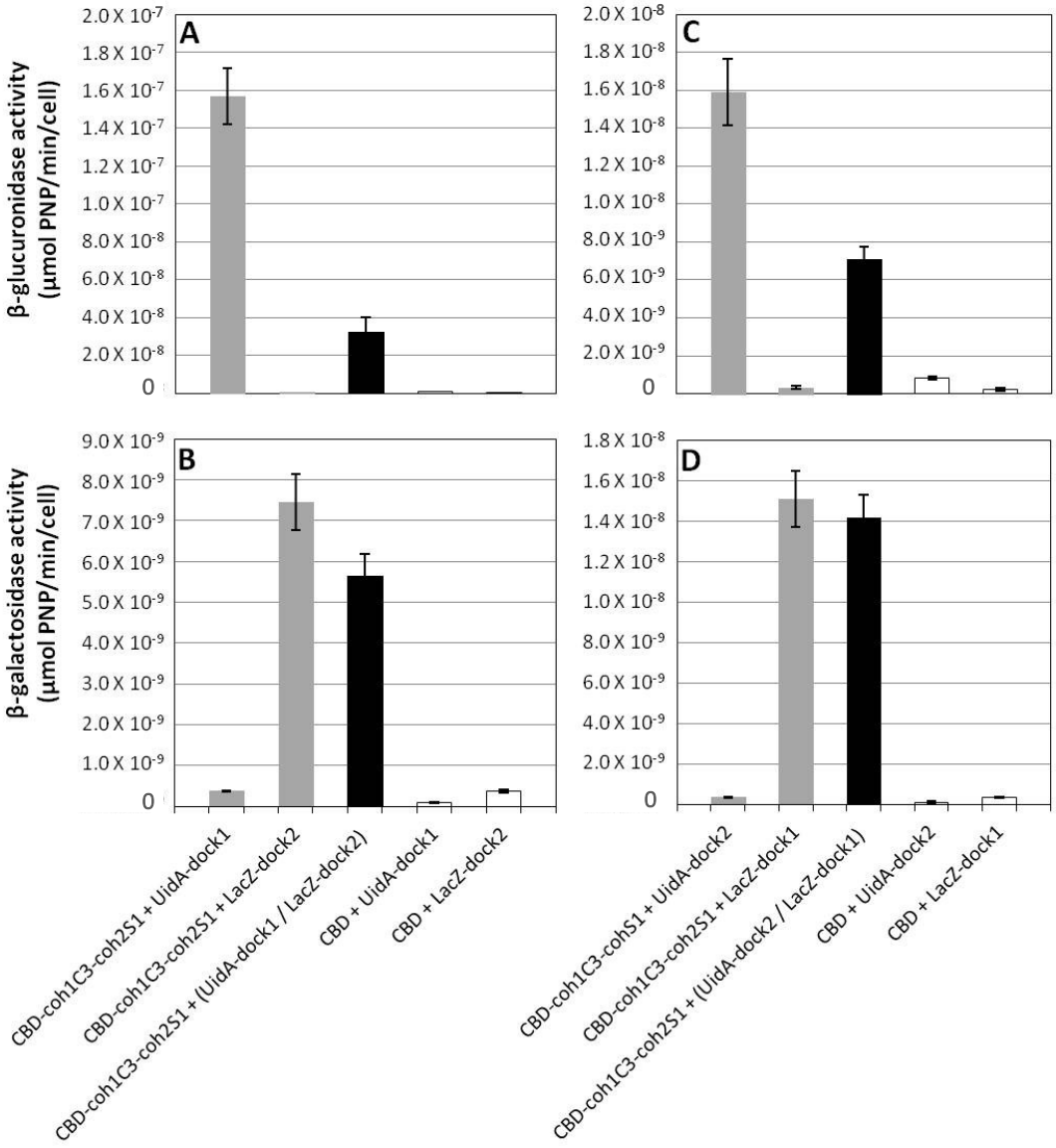

Figure 12. Enzymatic profiles of whole cells with anchored multi-enzyme complexes assembled via simultaneous targeting. (A and C) β-glucuronidase and **(B and D)** β-galactosidase activity of multi-enzyme complexes resulting from the simultaneous binding of **(A and B)** UidA-dock1 and LacZ-dock2 with a surface-displayed chimeric scaffold or **(C and D)** resulting from the simultaneous binding of UidA-dock2 and LacZ-dock1 with a surface-displayed chimeric scaffold. Grey bars represent catalytic activity of enzyme complex when a single enzyme is present, black bars represent catalytic activity of enzyme complex when both enzymes are present, and white bars represent binding assays carried out with scaffolds lacking cohesins. Expermiments were conducted in triplicate and a student T-test was performed.

protruding length of the scaffold exposing coh1C3 may have improved binding at this cohesin domain. The simultaneous binding of both UidA-dock1 and LacZ-dock2 resulted in a fivefold decrease in UidA activity compared to complexes containing UidA-dock1 alone (Fig. 12A). In contrast, complexes containing both enzymes showed no significant ($p>0.05$) decrease in LacZ activity when compared to complexes containing LacZ-dock2 alone (Fig. 12B). To gain insight into the drop in enzyme activity observed for the two-enzyme complex, cells expressing the same scaffold were incubated with equimolar amounts of LacZ-dock1 and UidA-dock2, targeting the same enzymes to opposite cohesins. As observed previously, the simultaneous docking of both enzymes resulted in a decrease (two-fold) in UidA activity compared to the scaffolds to which only UidA-dock2 was bound (Fig. 12C). Once more, no significant ($p>0.05$) decrease in LacZ activity was observed for complexes containing both enzymes when compared with complexes containing LacZ-dock1 alone (Fig. 12D).

3.3.5 Sequential binding of UidA- and LacZ-dockerin fusions to chimeric protein scaffolds

The order in which the chimeric scaffold was "loaded" with UidA and LacZ resulted in the assembly of two-enzyme complexes with different enzyme activities (Fig. 13). When LacZ-dock2 was bound onto the scaffold CBD-coh1C3-coh2S1 prior to UidA-dock1, the result was a two-fold decrease in UidA activity when compared to similar complexes containing UidA-dock1 alone (Fig. 13A). When this order of assembly was reversed, and UidA-dock1 was bound onto the scaffold chimera prior to LacZ-dock2, UidA

activity was not significantly affected ($p>0.05$) (Fig. 13A). In the same experiment, β-galactosidase activity was also measured to determine the effect of the sequential incorporation of the two enzymes on LacZ activity. When UidA-dock1 was bound to the scaffold prior to LacZ-dock2, a similar result was observed where LacZ activity decreased approximately 1.8 fold compared to complexes containing LacZ-dock2 alone (Fig. 13B). Contrarily, when LacZ-dock2 was bound to the scaffold prior to UidA-dock1, the resulting complex exhibited a similar level of LacZ activity when compared with complexes containing LacZ-dock2 alone (Fig. 13B). When LacZ-dock1 was targeted to the scaffold prior to UidA-dock2, a 4.7-fold decrease in UidA activity was observed, compared to complexes containing UidA-dock2 alone (Fig. 13C). However when the order was reversed and UidA-dock2 was incorporated prior to LacZ-dock1, much of the β-glucuronidase activity was regained, with an approximate 1.5 fold decrease in UidA activity compared with complexes containing UidA-dock2 alone (Fig. 13C). Interestingly, the observed β-galactosidase activity did not significantly ($p>0.05$) change when the order of assembly was switched. When UidA-dock2 was incorporated into the complex prior to LacZ-dock1, the result was only a marginal decrease in LacZ activity when compared to complexes containing LacZ-dock1 alone (Fig. 13D). In addition, when LacZ-dock1 was incorporated into the complex prior to UidA-dock2, LacZ activity was identical when compared with complexes containing LacZ-dock1 alone.

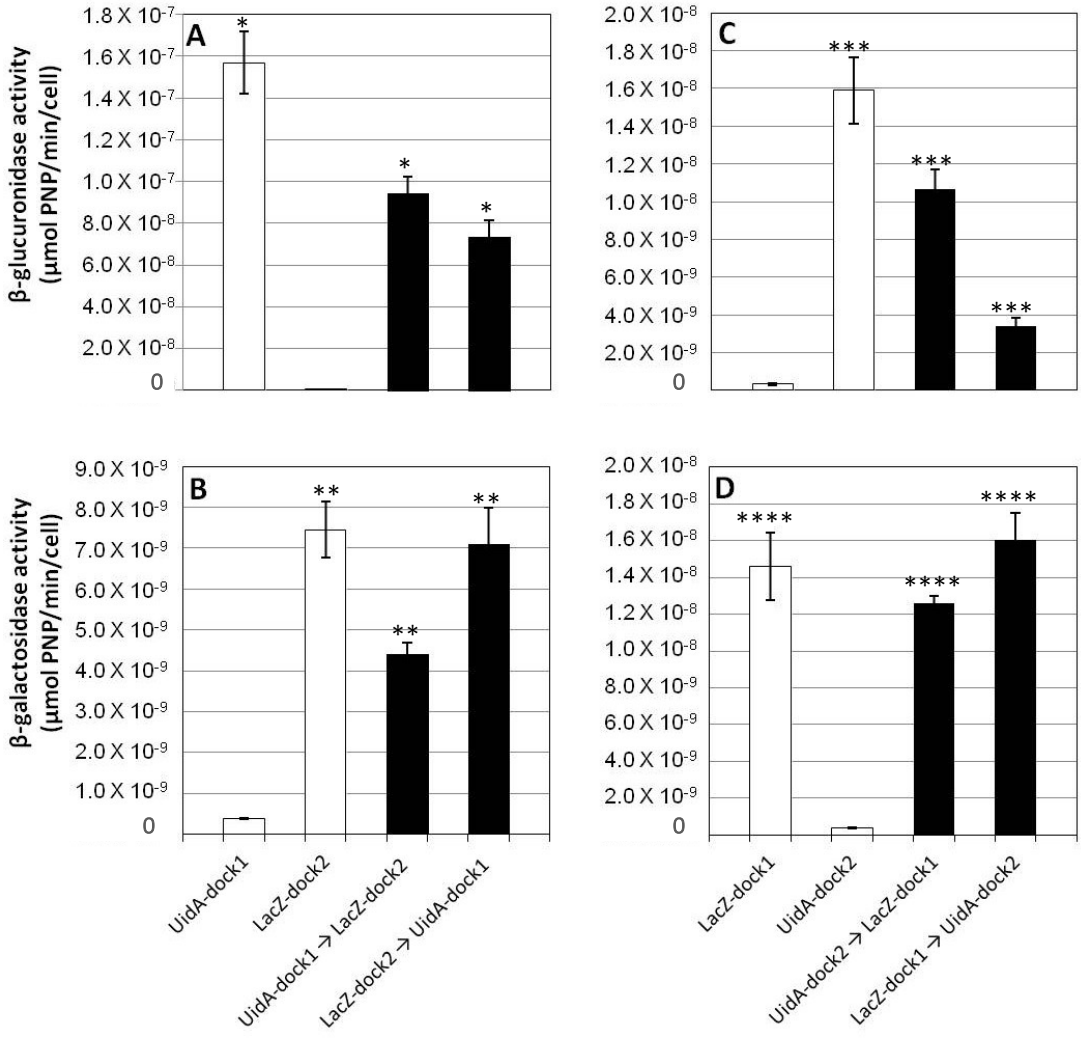

Figure 13. Enzymatic profiles of whole cells with anchored multi-enzyme complexes assembled via sequential targeting. (A and C) β-glucuronidase and **(B and D)** β-galactosidase activity of multi-enzyme complexes resulting from the sequential association of test enzymes onto the chimeric scaffold CBD-coh1C3-coh2S1. **(A and B)** Sequential targeting of UidA-dock1 and LacZ-dock2 and (**C and D)** sequential targeting of UidA-dock2 and LacZ-dock1. Enzyme activities are reported for a single enzyme bound to the scaffolds (white bars) and when both enzymes are bound (black bars). Expermiments were conducted in triplicate and a student T-test was performed. Single factor ANOVA was performed on each sample set: (* p=0.07, **p=0.28, ***p=0.02, ****p=0.58).

3.4 Discussion

In chapter 2, I reported on strains of *L. lactis* that successfully displayed type 1 cohesins on their surface, and demonstrated their ability to bind UidA-dock1 (Wieczorek & Martin, 2010). In this chapter, chimeric scaffold proteins consisting of cohesins from CipA and OlpB or SdbA were successfully displayed on the surface of *L. lactis*, however only CipA-SdbA chimeric scaffolds were capable of binding both UidA-dock1 and UidA-dock2, suggesting that either improper folding or inaccessibility of coh2O2 may have prevented its association with UidA-dock2. Previous studies have demonstrated that scaffold proteins derived from bacteria that anchor their cellulosome to the cell surface such as *C. thermocellum, Ruminococcus flavifaciens,* and *Acetivibrio cellulolyticus*, contain long inter-cohesin linkers (50-550 residues) compared to cellulosomes from organisms lacking such anchoring strategies such as *Clostridium cellulolyticum* (10 residues) (Gerngross, et al., 1993; Rincon, et al., 2003; Q. Xu, et al., 2003). It has also been proposed that linkers joining cohesins within CipA may increase the protein's conformational flexibility (Hammel, et al., 2005). With the goal of improving coh2O2 accessibility for dockerin binding, scaffold-derived linkers were engineered in my synthetic scaffolds (Fig. 1A), however no significant difference in enzyme binding at either cohesin was observed (Fig 2). Since the scaffolds were successfully displayed on the cell surface, I hypothesize that improper folding of the scaffold protein may have prevented important intramolecular ionic interactions or buried the coh2O2 domain within protein aggregates (Miot & Betton, 2004; Vallejo & Rinas, 2004). In addition, deletion of the HtrA housekeeping protease in this strain may account for the misfolded

proteins remaining associated with the cell surface (Poquet, et al., 2000). It has also been previously demonstrated that targeting recombinant fusion proteins to the cell wall of *L. lactis* can cause problems with secretion, anchoring, and/or folding (Linares, et al., 2010).

Since the inclusion of linkers exterior to the coh2O2 domain did not result in binding of UidA-dock2 to the chimeric scaffolds, I replaced coh2O2 with coh2S1 and found that the resulting scaffold could bind UidA-dock1 and UidA-dock2 demonstrating that both cohesin domains were accessible and functional. SdbA differs from OlpB in that it contains one rather than four cohesins, as well as a lysine-rich region downstream of coh2S1 that shares a high degree of homology to streptococcal M proteins (Leibovitz & Beguin, 1996). I postulate that incorporating coh2S1 adjacent to the anchor motif of streptococcal M6 protein may emulate some structural characteristics found in the native SdbA anchor protein of the *C. thermocellum* cellulosome, resulting in improved accessibility for UidA-dock2 binding.

Having successfully generated a chimeric scaffold capable of binding both UidA-dock1 and UidA-dock2, I sought to determine which scaffolds, among a total of four varints, could be displayed most efficiently and bind the greatest amount of each reporter enzyme. Interestingly, cells binding the most UidA-dock1 and UidA-dock2 molecules were those displaying the larger tri-modular scaffolds CBD-coh1C3-coh2S1 and CBD-coh1C1-coh1C2-coh2S1. In a previous study, I also demonstrated that increased scaffold protein size did not reduce the efficiency of scaffold display or functionality (Wieczorek & Martin, 2010). It was observed that fusion of the highly insoluble

Clostridium cellulovorans cellulase CelL with the CBD of cellulase CelD resulted in a significant increase in the recombinant enzyme's overall solubility (Murashima, et al., 2003), and I hypothesize that this may account in part for the higher efficiency of one larger scaffold, CBD-coh1C3-coh2S1.

Having determined which scaffolds bound the greatest amounts of dockerin-fused enzymes, I also analyzed their relative abundance within the assembled complexes, since protein ratios can ultimately have an effect on enzyme synergy and substrate-channeling (Dueber, et al., 2009; Murashima, Kosugi, et al., 2002). Assuming a 1:1 cohesin to dockerin binding ratio, it should be expected that the coh1C1-coh1C2-coh2S1 scaffold would bind double the number of UidA-dock1 compared to CBD-coh1C3-coh2S1 as it contains twice as many type 1 cohesins. However, I observed that cells expressing coh1C1-coh1C2-coh2S1 were able to bind less of the UidA-dock1 reporter enzyme. This result is in accordance with my previous observations made while comparing scaffolds CBD-coh1C3 and coh1C1-coh1C2 (Wieczorek & Martin, 2010). Assuming a 1:1 cohesin to dockerin binding ratio, coh1C1-coh1C2-coh2S1 should also theoretically bind UidA-dock1 and UidA-dock2 in a 2:1 ratio while CBD-coh1C3-coh2S1 should bind equimolar amounts. The resulting ratios deviated from those predicted, with a ratio of 2.5 for UidA-dock1 / UidA-dock2 binding to coh1C1-coh1C2-coh2S1, and 3.5 for UidA-dock1 / UidA-dock2 binding to CBD-cohC3-cohS1. In a previous study, the assembly of chimeric scaffold-derived enzyme complexes on the surface of *Saccharomyces cerevisiae* also resulted in deviations from expected ratios of enzymes, as cellobiohydrolase CBHII associated with scaffolds at lower levels than other enzymes (Wen, et al., 2010). I therefore suggest that

variability in the proper folding and/or accessibility of individual cohesin domains within a chimeric scaffold may yield differences in the number of each respective enzyme successfully becoming included in the resulting complexes.

To gain further insight into factors affecting protein binding to my synthetic scaffold proteins, I "docked" individual enzymes simultaneously or sequentially onto the chimeric CBD-coh1C3-coh2S1 protein. When simultaneously binding UidA-dock1 and LacZ-dock2 to the scaffold, an approximate five-fold decrease in UidA activity was observed compared to the binding of UidA-dock1 alone whereas no significant decrease in LacZ activity was observed in these assays (Fig. 12). My data suggest that the different effects on UidA and LacZ binding and/or activity may be due to either the location of the cohesin within the scaffold, to the size of each enzyme relative to the other, or differences in binding affinities between the two recombinant cohesion-dockerin interactions. Therefore, a similar binding assay was performed where the location of the cohesins on the scaffold protein was reversed. Similarly, UidA activity was two-fold lower when incorporated in the presence of LacZ-dock1, and once again, no significant change in LacZ activity was observed when incorporated in the presence of UidA-dock2 (Fig. 12). Since LacZ is significantly larger than UidA (480 kDa vs 280 kDa), this suggests that enzyme size may result in steric factors inhibiting the binding of one enzyme partner, and that the relative location of each enzyme did not seem to play a role in the resulting activities when enzymes were incorporated simultaneously.

Sequential enzyme binding assays gave similar results as simultaneous binding assays where more than a two-fold decrease in UidA activity resulted when LacZ-dock2 was bound to the scaffold prior to UidA-dock1 addition. Contrarily, although LacZ activity decreased significantly when UidA-dock1 was bound to the scaffold protein prior to LacZ-dock2, reversing this order resulted in the same LacZ activity as when LacZ-dock2 alone was targeted to the scaffold (Fig. 13). To verify if enzyme location also affected the overall resulting activity of the complex, the location of each enzyme partner was reversed. UidA activity decreased when LacZ-dock1 was incorporated prior to UidA-dock2, and this activity was only partially regained when the order of assembly was reversed (Fig. 13C). LacZ activity was not affected by the order in which LacZ-dock1 and UidA-dock2 were bound into such complexes (Fig. 13D). In addition, when UidA-dock1 was targeted to the coh1C3 cohesin (Fig. 13A), the order in which LacZ was targeted to coh2S1 also had less of an effect on resulting UidA activity compared to when UidA-dock2 was targeted to coh2S1 (Fig. 13C). From these results, it appears that when a fusion enzyme is targeted to the outermost position on the scaffold, distal to the cell surface, its binding to the scaffold may be less affected by enzyme partners, compared to when it is targeted to the innermost position, proximal to the cell surface.

3.5 Conclusions

I describe the first successful display of engineered chimeric scaffolds containing type 1 and type 2 cohesins on the surface of *L. lactis,* and the ability for the scaffolds to

support the assembly of multi-enzyme complexes. Traditional modes of enzyme display in this bacterium were generally limited to fusing a single enzyme with an appropriate anchor (Avall-Jaaskelainen, et al., 2003; Cortes-Perez, et al., 2005; Dieye, et al., 2003; Dieye, et al., 2001; Leenhouts, et al., 1999; Lindholm, et al., 2004; Narita, et al., 2006; Piard, et al., 1997; Raha, et al., 2005; Ramasamy, et al., 2006; Ribeiro, et al., 2002; Yang, et al., 2008). In this study, I expand this capacity to two enzymes with the simultaneous or sequential incorporation of the two enzymes resulting in differences in the enzymatic profile of the assembled complexes. These results suggest that the size and location of each enzyme within each complex should be carefully taken into consideration when further developing this system of enzyme display.

The strategy adopted in this chapter of my thesis involved two enzymes binding to specific sites on a single surface displayed scaffold. It is noteworthy to also mention that less enzyme was successfully bound to scaffolds when appended with a type 2 dockerin. I envisioned that possible improvements on this system would include multi-scaffold complexes for binding larger amounts of enzyme, and investigation of means to improve the type 2 cohesin-dockerin interaction.

Chapter 4

Engineering and *in vivo* assembly of two-level bi-enzymatic complexes on *Lactococcus lactis*

4.1 Introduction

To date, the design of recombinant cellulosome has focused mainly on the engineering of single scaffolds capable of binding up to four enzymes as described in chapter 1. In one example, however, novel cellulosome architectures including the covalent attachment of all domains as well as the assembly of a "cyclic" cellulosome were investigated (Mingardon, et al., 2007). Of all architectures explored, however, the cellulosome demonstrating the greatest activity was still the "hybrid cellulosome", consisting of a single chimeric scaffold upon which enzymes could bind by means of dockerin-cohesin interactions (Mingardon, et al., 2007). In the results presented in this chapter, I sought to engineer novel scaffolds to increase the amount of enzymes present, potentially increasing the overall activity as well. I first drew inspiration from strategies adopted for metropolitain housing, where the construction of vertical high-rises could increase the number of inhabitants within a dedicated surface area once ground-level saturation has been achieved. I then searched for examples analogous to this in natural cellulosome systems, and discovered that certain bacteria such as *R. flavefaciens* (Rincon, et al., 2003) and *A. cellulolyticus* (Q. Xu, et al., 2003) both display cellulosomes on their surface which are composed of secondary "adapter" scaffolds that increase the amount of enzymes localized within each complex (Fig. 2). In addition, in the case of *A. cellulolyticus,* certain adapter scaffolds contain not only cohesins and dockerins, but a

catalytic module as well, potentially contributing to the cellulosome's overall activity (Q. Xu, et al., 2003).

The following chapter of my thesis describes the engineering of multi-enzyme complexes containing a primary surface-displayed scaffold, a secondary adapter scaffold with catalytic activity, as well as dockerin-fused enzyme(s). I also describe my attempt to generate "self-assembling" scaffolds with the goal of increasing the overall amount of enzymes present within each complex. Chapter 3 also demonstrated that significantly less enzyme activity was observed when UidA was targeted to the type 2 cohesin site, and I hypothesized that this may be due in part to the absence of the "X" module in my constructs. This module has been shown to contribute electrostatic interactions necessary for the integrity of the type 2 dockerin-cohesin complex (J. Xu & Smith, 2010). I therefore investigated the effects of incorporating the X module in a UidA-dock2 fusion. Finally, I describe how a dual-plasmid system could be used for the simultaneous expression and secretion of both scaffolds and enzymes in *L. lactis*, a step that may prove necessary for the full *in vivo* assembly of extracellular complexes on this organism.

4.2 Materials and Methods

4.2.1. Bacterial strains and plasmids used in this chapter

The bacterial strains and plasmids used in this chapter are listed in Table 6. *E. coli* strains were grown in Luria-Bertani medium at 37°C with shaking (220 rpm). *L. lactis htrA*-NZ9000 was grown in M17 medium (Terzaghi & Sandine, 1975) supplemented with

Table 6 - Strains and plasmids used in this chapter. pAW500 series of vectors are designed for the cell-wall targeting of various scaffold permutations consisting of Cip_{frags} or $CipA_{frag}$-$SdbA_{frag}$ fusions. *cohC1*, type 1 cohesin of CipA; *coh2S1*, type 2 cohesin of SdbA; *PT7*, inducible T7 promoter; P_{nisA}, inducible *nisA* promoter; rbs_{usp45}, Usp45 ribosome-binding site; rbs_{nisA}, *nisA* ribosome-binding site; sp_{Usp45}, signal sequence of Usp45; cwa_{M6}, anchor motif of M6 protein; *llt2*, transcriptional terminator of *rrnB* operon; t_{trpA}, transcriptional terminator of *trpA*; *Lk, olpB* linker region. pST series of vectors are *L. lactis-E. coli* shuttle vectors used for the *in vivo* production of enzymes or secondary adapter scaffolds in *L. lactis*.

Strain	Genotype / Decription	Source
L. lactis htrA-NZ9000	Mutant MG1363 derivative (*nisRK* genes on the chromosome) lacking *htrA*	(Miyoshi, et al., 2002)
E. coli TG1	*supE thi-1* Δ(*lac-proAB*) Δ(*mcrB-hsdSM*)*5* (rK– mK–) [F´ *traD36 proAB lacI*q*Z*Δ*M15*]	ATCC
E. coli BL21 (DE3)	*F⁻ ompT gal dcm lon* $hsdS_B(r_B^- m_B^-)$ *λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])*	Novagen
Plasmid		
pSCNIII	Cm^r	J. Seegers[a]
pAW528	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}- *CBD-coh1C3-*cwa_{M6}*-tlt2*	(Wieczorek & Martin, 2010)
pAW531	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*CBD-*cwa_{M6}*-tlt2*	(Wieczorek & Martin, 2010)
pAW579	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*CBD-coh2S1-*cwa_{M6}*-tlt2*	S3[b]
pAW576	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*CBD-coh1C3-coh2S1-*cwa_{M6}*-tlt2*	S3[b]
pAW573	Cm^r, Amp^r; pBS::pIL252::t_{trpA}::P_{nisA}::rbs_{nisA}::sp_{Usp45}-*coh1C1-coh1C2-coh2S1-*cwa_{M6}*-tlt2*	S3[b]
pETUdock1	Kn^r; pET28(b)::*PT7*::*6xHis-uidA-dock1*	(Wieczorek & Martin, 2010)
pETUdock2	Kn^r; pET28(b)::*PT7*::*6xHis-uidA-dock2*	S3[b]
pETU	Kn^r; pET28(b)::*PT7*::*6xHis-uidA*	(Wieczorek & Martin, 2010)

pETLdock1	Kn^r; pET28(b)::*PT7*::*6xHis-lacZ-dock1*	S3[b]
pETUXD2	Kn^r; pET28(b)::*PT7*::*6xHis-uidA-X-dock1*	This Work
pETULC9XD2	Kn^r; pET28(b)::*PT7*::*6xHis-uidA-link-coh1C9-X-dock2*	This Work
pET73D1	Kn^r; pET28(b)::*PT7*::*6xHis-coh1C1-coh1C2-coh2S1-dock1*	This Work
pETPRSLD1	Kn^r; pET28(b):: P_{nisA}::rbs_{nisA}::sp_{Usp45}-*lacZ-dock1*	This Work
pETPRSUXD2	Kn^r; pET28(b):: P_{nisA}::rbs_{nisA}::sp_{Usp45}-*uidA-X-dock2*	This Work
pETRSUXD2	Kn^r; pET28(b):: rbs_{nisA}::sp_{Usp45}-*uidA-X-dock2*	This Work
pETPRSULC9XD2	Kn^r; pET28(b):: P_{nisA}::rbs_{nisA}::sp_{Usp45}-*uidA-link-coh1C9-X-dock2*	This Work
pETPRS73D1	Kn^r; pET28(b):: P_{nisA}::rbs_{nisA}::sp_{Usp45}-*coh1C1-coh1C2-coh2S1-dock1*	This Work
pST	Tet^r: pSCNIII	This Work
pSTe	Tet^r; pST::Empty Vector	This Work
pSTPRSLD1	Tet^r; pST:: P_{nisA}::rbs_{nisA}::sp_{Usp45}-*lacZ-dock1*	This Work
pSTPRSUXD2	Tet^r; pST:: P_{nisA}::rbs_{nisA}::sp_{Usp45}-*uidA-X-dock2*	This Work
pSTPRSULC9XD2	Tet^r; pST:: P_{nisA}::rbs_{nisA}::sp_{Usp45}-*uidA-link-coh1C9-X-dock2*	This Work
pSTPRSLD1-RSUXD2	Tet^r; pST:: P_{nisA}::rbs_{nisA}::sp_{Usp45}-*lacZ-dock1*:: rbs_{nisA}::sp_{Usp45}-*uidA-X-dock2*	This Work
pST73D1	Tet^r; pST:: *coh1C1-coh1C2-coh2S1-dock1*	This Work
pSTPRS73D1	Tet^r; pST:: P_{nisA}::rbs_{nisA}::sp_{Usp45}-*coh1C1-coh1C2-coh2S1-dock1*	This Work

[a]Vector pSCNIII was a gift provided by Jos Seegers (unpublished data)
[b]Refer to section 3 of this thesis for the construction of these vectors

1% (w/v) glucose (GM17) at 30°C without agitation. *C. thermocellum* was grown in ATCC1191 medium at 55°C with 0.2% (w/v) cellobiose as a carbon source. Where appropriate, antibiotics were added as follows: for *E. coli*, ampicillin (100 µg/mL), chloramphenicol (10 µg/mL), tetracycline (6 µg/mL), and kanamycin (30 µg/mL); for *L. lactis,* erythromycin (5 µg/mL), tetracycline (3 µg/mL), and chloramphenicol (10 µg/mL).

General molecular biology techniques for *E. coli* were performed as previously described (Sambrook & Russell, 2001). Genomic DNA was isolated from *C. thermocellum* as previously described (Wang & Wu, 1993). Genomic DNA from *L. lactis* was isolated using a DNeasy Blood and Tissue kit (Qiagen), after cells were incubated for 1hr in the presence of 10mg/mL lysozyme to degrade the cell wall. To make competent cells, *L. lactis* was grown in M17 medium (Terzaghi & Sandine, 1975) supplemented with 1% (w/v) glucose, 25% (w/v) sucrose and 2% (w/v) glycine and cells were transformed as previously described (Holo & Nes, 1989). M17 media was supplied by Oxoid, LB media was supplied by Novagen, all antibiotics, *p*-nitrophenyl-β-D-glucuronide, *O*-nitrophenyl-β-galactoside and nisin were provided by Sigma, and X-gal and IPTG were supplied by Fermentas.

4.2.2 Assembly of an expression cassette for the overproduction of UidA-X-dock2

E. coli β-glucuronidase was previously engineered to contain a HisX6 tag for protein purification as well as a C-terminal dock1 (Wieczorek & Martin, 2010) or dock2 domain for targeting type 1 or type 2 cohesins (chapter 3), respectively. In this chapter, UidA was fused with the C-terminal X-dock2 fragment of CipA (Fig. 14A) in order to test the effects of the X module on enzyme loading on scaffold-displaying cells, and on the resulting β-glucuronidase activity of the complexes. DNA encoding the X-dock2 region of *cipA* was PCR-amplified from *C. thermocellum* genomic DNA using primers *a* and *b* (Table 7), digested with *EcoRI* and *XhoI*, and ligated to similarly-cut pETUidAdock1, yielding vector pETUXD2. The integrity of the expression cassette was verified by DNA sequencing and pETUXD2 was transformed in *E. coli* BL21 (DE3) by electroporation.

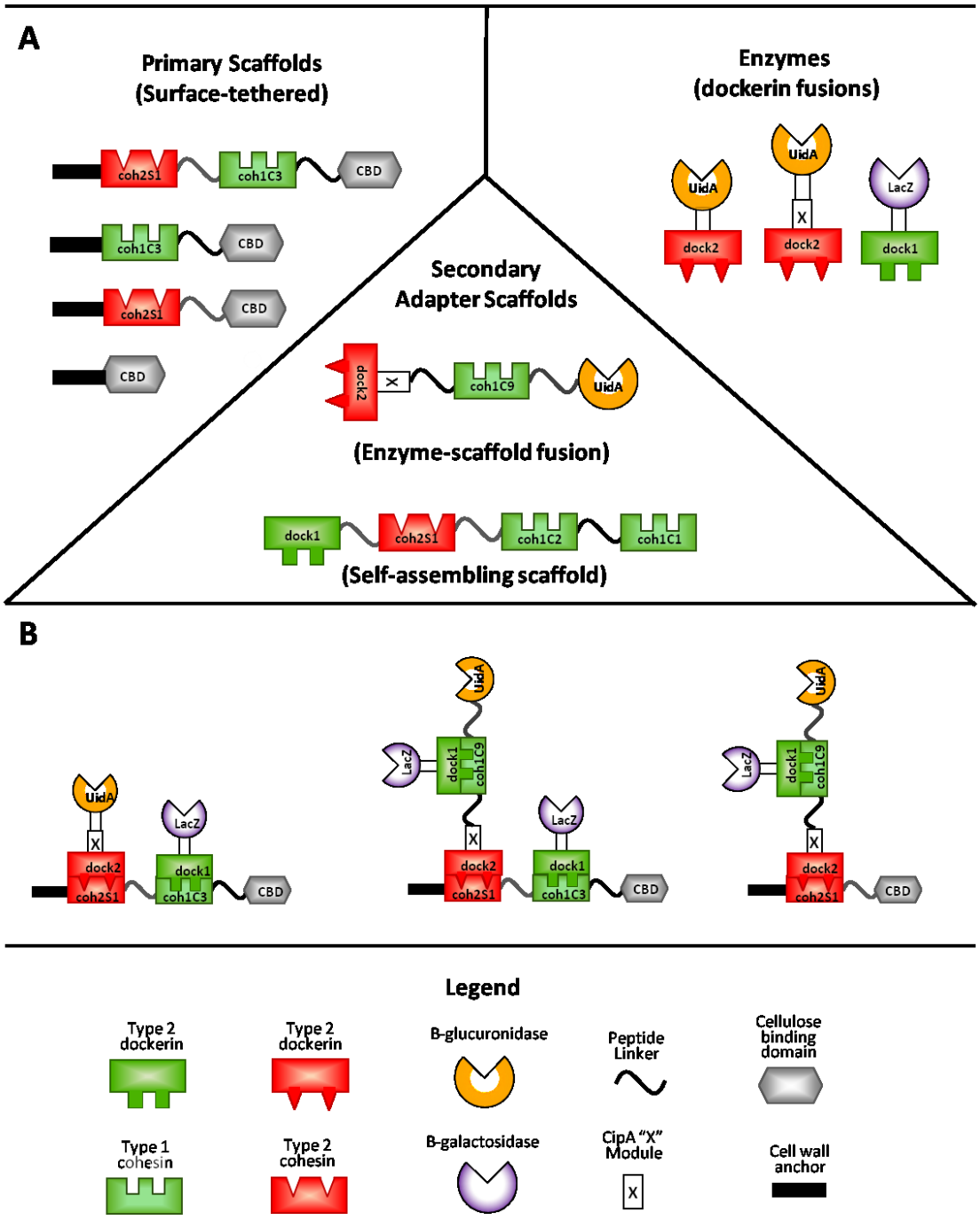

Figure 14. Architectural design of enzymes and scaffolds engineered to assemble into multi-enzyme complexes. (A) Primary scaffolds produced *in vivo* in *L. lactis* are covalently tethered to the cell surface by means of the cwa_{M6}. Enzyme-dockerin fusions consist of LacZ fused with dock1 and UidA fused with either dock2 or X-dock2. Secondary adapter scaffolds include UidA fused with a $CipA_{frag}$ containing both a coh1 and dock2 domain for binding both LacZ-dock1 and primary scaffolds, respectively, and a self-assembling adapter scaffold containing both coh1 and coh2 domains, as well as a dock2 domain. (B) Schematic depiction of the strategy involved towards designing two-level enzymatic complexes whose assembly is mediated by primary scaffolds CBD-coh2S1 or CBD-coh1C3-coh2S1.

Table 7 - Primers used in this chapter (restriction enzyme cut sites are in bold).

Primer	Sequence (5' – 3')
a	ATGC**GAATTC**AATAAACCTGTAATAGAAGGATATAAAG
b	ATGC**CTCGAG**TTTAATATTTACTGTGCGTCGTAATCAC
c	ATGC**GAATTC**GATACAACAGTACCTACAACATCGCCGAC
d	ATGC**GCTAGC**GAGGCGCGCCCGGCCACAATGAC
e	TAGC**GAATTC**CGATATCTTATCCGGCTGTATTAC
f	ATGC**GAATTC**ATGCTTTGTATACCTATGGTTATG
g	GCAT**AGGCCT**CTAAGTTATTTTATTGAACATATATCG
h	TCACTAAAGGGAACAAAAGCTGGGTACCGGGCC
i	ATGC**GCTAGC**CGCATCTTGTTTAGCAATATCTGAG
J	ATGC**AGATCTGGGGCCT**TACAAAATAAATTATAAGGAGGC

4.2.3 Assembly of an expression cassette for the overproduction of secondary adapter scaffold UidA-Link-coh1C9-X-dock2

In order to generate an expression cassette for the overproduction of a UidA-fused adapter scaffold containing a type 1 cohesin and a type 2 dockerin (Fig. 14A), DNA encoding the C-terminal Link-coh1C9-X-dock2 fragment of CipA was PCR-amplified from *C. thermocellum* genomic DNA using primers *c* and *b* (Table 7), digested with *EcoRI* and *XhoI*, and ligated to similarly-cut pETUidAdock1, yielding vector pETULC9XD2. The expression cassette was sequenced in order to verify its integrity and pETULC9XD2 was transformed in *E. coli* BL21 (DE3) by electroporation.

4.2.4 Assembly of an expression cassette for the overproduction of "vertically" self-assembling scaffold coh1C1-coh1C2-coh2S1-dock1

In order to generate an expression cassette for the production of a self-assembling scaffold protein (Fig. 14A), the *coh1C1-coh1C2-coh2S1* cassette was PCR-amplified from vector pAW573 using primers *d* and *e* (Table 7), digested with *NheI* and *EcoRI*, and ligated to similarly-cut pETUidAdock1, yielding vector pET73D1. The expression cassette was sequenced in order to verify its integrity and pET73D1 was transformed in *E. coli* BL21 (DE3) by electroporation.

4.2.5 Expression and purification of recombinant enzymes, adapter scaffolds, and vertically self-assembling scaffolds

All His-tagged fusion enzymes were over-expressed in *E. coli* BL21(DE3) as previously described (Wieczorek & Martin, 2010). Fractions containing UidA and LacZ fusion proteins were identified based on the appearance of a yellow color in a liquid β-glucuronidase or β-galactosidase assay, respectively. Elution fractions exhibiting enzyme activity were analyzed by SDS-PAGE (12% w/v). For detection of the non-enzymatic scaffold coh1C1-coh1C2-coh2S1-dock1, multiple elution fractions were loaded on 12% (w/v) SDS-PAGE for detection of purified proteins. Proteins were stained using Coomasie Blue Reagent (BioRad) and fractions containing the highest purity of enzyme or scaffold were identified and pooled.

4.2.6 Assembly of enzyme/scaffold complexes on the surface of *L. lactis* by exogenous addition of components

L. lactis htrA-NZ9000 strains harboring the vectors pAW531 (CBD), pAW579 (CBD-coh2S1) and pAW576 (CBD-coh1C3-coh2S1) were grown overnight in GM17 media containing appropriate antibiotics, diluted 1/50 in 5 mL of fresh media, grown for 4 hrs (OD_{600} ≈0.3) and induced with 10 ng/mL nisin. One mL of cell cultures were harvested by centrifugation, washed in phosphate buffer (50 mM, pH 6.0) containing 300 mM NaCl and suspended in 100 μL of purified enzyme or adapter scaffold UidA-Link-coh1C9-X-dock2 at a concentration of 100 μg/mL. For the assembly of a two-level complex (Fig. 14B), the binding assay was carried out in identical fashion to the sequential binding assays (see chapter 3) with UidA-Link-coh1C9-X-dock2 and LacZ-dock1 added in a 1:2 molar ratio. Incubation and enzyme assay conditions are described in chapters 2 and 3. Experiments were performed in triplicate using true biological replicates (independent colonies and cultures).

4.2.7 Construction of vectors for the *in vivo* production of recombinant enzymes, adapter scaffolds, and self-assembling scaffolds in *L. lactis*

pSCNIII is an *E. coli-L. lactis* shuttle vector which contains a chloramphenicol resistance gene (*cat*). For the simultaneous expression of both primary scaffolds and enzymes or adapter scaffolds, a dual-plasmid system was used (Fig. 15). To achieve this, the *cat* gene on pSCNIII was replaced with tetracycline resistance gene (tet^r) from *Streptococcus pneumoniae*. A mutant *L. lactis* strain was previously constructed by replacing the gene encoding an endogenous recombinase RecA with the tet^r gene

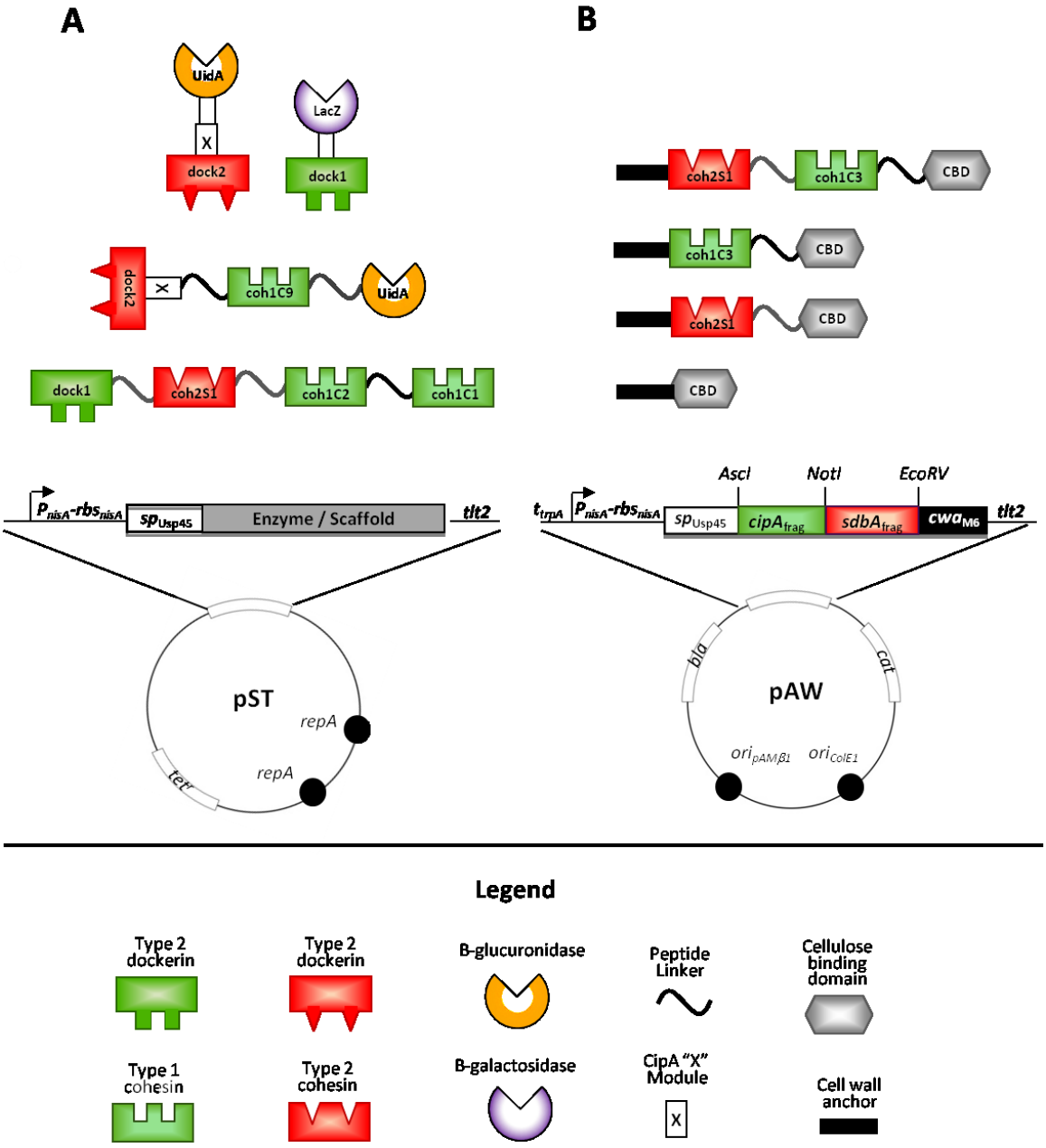

Figure 15. Dual-plasmid system for the simultaneous *in vivo* production of all complex components in *L. lactis*. (A) The pST series of vectors use tetracycline resistance as a selective marker and contain expression cassettes encoding enzyme-dockerin fusions or secondary adapter scaffolds. Expression is under control of the inducible *nisA* promoter while secretion is mediated by an N-terminal signal peptide. (B) pAW series of vectors use chloramphenicol resistance as a selective marker and contain expression cassettes for the secretion and surface-display of primary scaffolds which mediate complex assembly.

(Duwat, et al., 1995), and was kindly provided by Alexandra Gruss. The *tet*^{r} gene lacking it's native promoter was PCR-amplified from *L. lactis recA* chromosomal DNA using primers *f* and *g* (Table 7), digested with *EcoRI* / *StuI* and ligated to *MfeI* / *StuI* digested pSCNIII, yielding vector pST. This strategy resulted in the use of the endogenous promoter and ribosome-binding site of the *cat* gene, followed by the ORF of the *tet*^{r} gene. For the expression of enzymes and adapter scaffolds in *L. lactis*, DNA encoding the P_{nisA}-RBS_{nisA}-SP_{USp45} (PRS) cassette was PCR amplified from vector pAW576 using primers *h* and *I* (Table 7), digested with *BglII* and *NheI*, and ligated into similarly-cut pETLacZdock1, pETUXD2 and pETULC9XD2, yielding vectors pETPRSLD1, pETPRSUXD2 and pETPRSULC9XD2 respectively. Digestion of these vectors with *BglII* and *XhoI* and subsequent ligation into similarly-cut pST yielded vectors pSTPRSLD1, pSTPRSUXD2 and pSTPRSUlc9XD2, respectively (Fig. 15A).

For the simultaneous expression of LacZ-dock1 and UidA-X-dock2, an operon was engineered within the pST vector. To achieve this, the RBS_{nisA}-SP_{USp45} (RS) cassette was PCR-amplified using primers *j* and *I* (Table 7), digested with *BglII* and *NheI*, and ligated into similarly-cut pETUXD2, yielding vector pETRSUXD2. Vector pETRSUXD2 was digested with *EcoO109I* and *BlpI* and ligated to similarly-cut pSTPRSLD1. The resulting vector pSTPRSLD1-RSUXD2 was verified using restriction digest analysis. For construction of a vector expressing coh1C1-coh1C2-coh2S1-dock1, pET73D1 was digested with *BglII and BlpI* and the released fragment was ligated to similarly-cut pST, yielding vector pST73D1. Vector pAW573 was digested with *BglII* and *SphI*, and the released fragment containing

the P_{nisA}-RBS_{nisA}-SP_{USp45} cassette was ligated to similarly-cut pST73D1, yielding pSTPRS73D1.

For the construction of a control vector lacking an expression cassette, pST was digested with *EagI* and the backbone was self-ligated generating vector pSTe. All resulting vectors were established in scaffold-producing strains of *L. lactis* by electroporation.

4.2.8 *In vivo* expression of scaffolds and enzymes in *L. lactis* for assembly of surface-displayed complexes

L. lactis electrocompetent cells harboring vectors pAW576, pAW528, pAW579 or pAW531 (Fig. 15B), as well as plasmid-free strain *htrA*-NZ9000 were transformed with the desired pST series vectors (Fig. 15A) and plated on SR solid media (10 g/L tryptone, 5 g/L yeast extract, 200 g/L sucrose, 10 g/L glucose, 15 g/L agar, 2.5 mM $MgCl_2$, and 2.5 mM $CaCl_2$, pH 6.8), supplemented with appropriate antibiotics. After 2 days growth, colonies were streak-purified on SR media containing appropriate antibiotics. Individual colonies were streaked on GM17 media containing 0.5% glucose and appropriate antibiotics. Colonies were transferred to 5 mL of liquid GM17 media containing 0.5% glucose and grown overnight. Cultures were diluted 1/50 and grown until mid-log phase was reached corresponding to an OD_{600} of ≈0.3, after which they were induced with 10 ng/mL nisin. After 20 hrs of growth, cells from 1 mL of culture were harvested by centrifugation (4300 X g), washed 6 times in phosphate buffer (50 mM, pH 6.0) containing 300 mM NaCl, and suspended in 100 µL of the same buffer. In order to verify

that protein production was not occurring after the wash steps, one additional sample was washed with the addition of 12 µg/mL lincomycin in the suspension buffer to inhibit protein synthesis. For quantification of β-glucuronidase and β-galactosidase activity, 50 µL of cells were subjected to the appropriate enzyme assays.

A co-culture was also used to assemble multi-level enzyme complexes comprising secondary scaffold coh1C1-coh1C2-coh2S1-dock1 on the cell surface. To achieve this, scaffold-displaying cultures of *L. lactis* harboring pSTPRSUXD2 and pSTPRS73D1 were diluted 1/50 and when cells reached mid-log phase, the cultures were diluted to an exact OD_{600} of 0.3, after which the mixed cultures were induced with 10 ng/mL nisin.

4.3 Results

4.3.1 The CipA "X" module in a UidA-dock2 fusion contributes to an increase in enzyme activity when bound to enzyme/scaffold complexes

The effects of the X module insertion on β-glucuronidase activity of surface-bound complexes were tested using binding assays on cells displaying the scaffold CBD-coh1C3-coh2S1. Insertion of the X module in the UidA-dock2 fusion increased β-glucuronidase activity by approximately 2.5-fold when compared with cells bound with UidA-dock2 without the X module (Fig. 16). The X module was shown to contribute electrostatic interactions, which are necessary for the integrity of the type 2 dockerin-cohesin interaction (J. Xu & Smith, 2010). For negative controls, similar binding assays were conducted where either the cohesin or dockerin partner was removed. Incubation

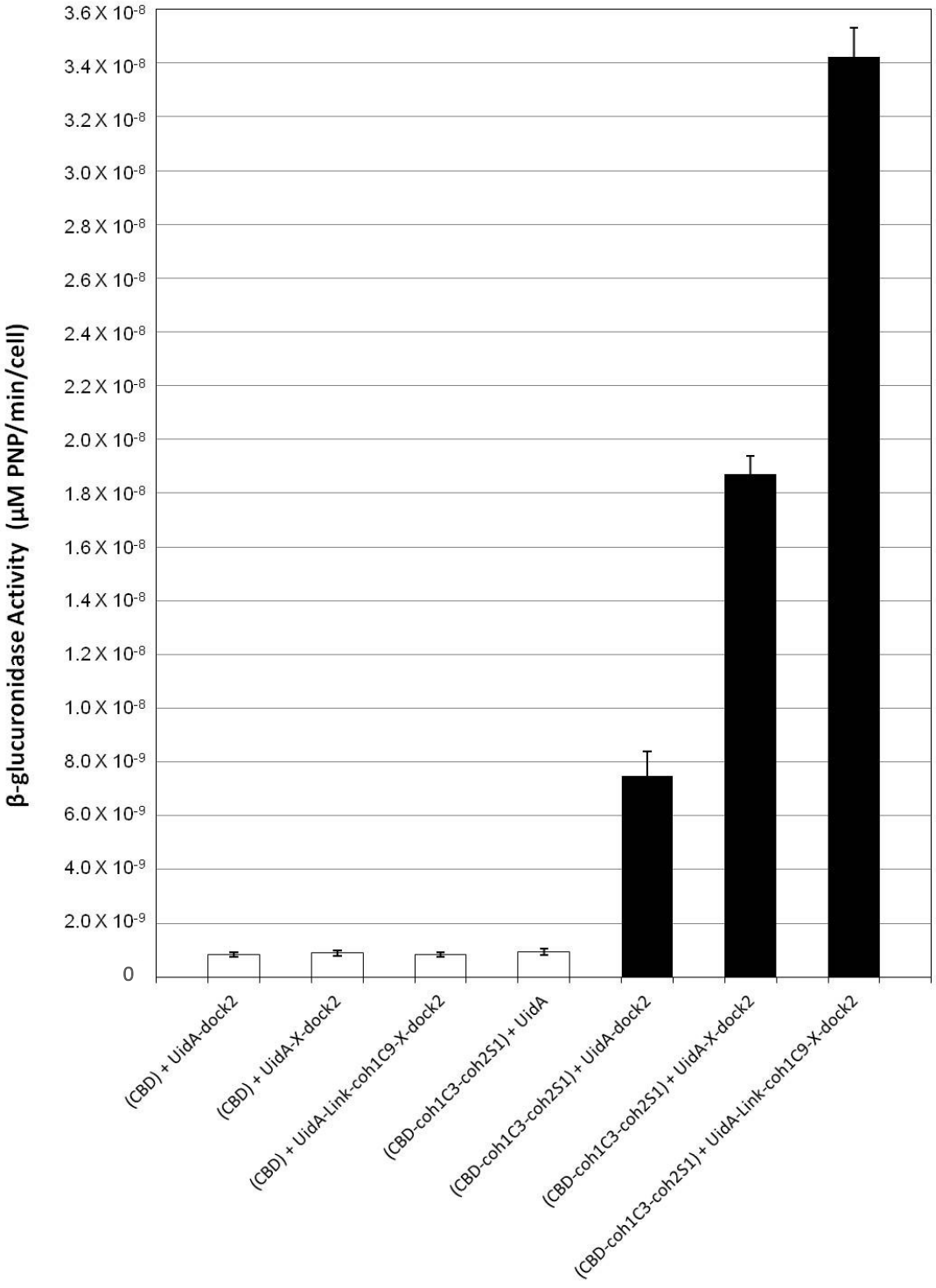

Figure 16. Effects of the CipA X-module on complex assembly mediated by type 2 cohesin-dockerin interactions. β-glucuronidase activity of cells displaying primary scaffolds bound with either UidA-dock2, UidA-X-dock2 or UidA-Link-coh1C9-X-dock2 (black bars). Controls include binding assays using UidA with no dockerin fusion, or scaffolds containing no cohesin domain (white bars).

of cells displaying CBD alone were incapable of binding both UidA-dock2 and UidA-X-dock2, while cells displaying the chimeric scaffold CBD-coh1C3-coh2S1 were incapable of binding UidA lacking a dockerin domain (Fig. 16).

4.3.2 Secondary adapter scaffold UidA-Link-coh1C9-X-dock2 binds surface-displayed primary scaffolds and exhibits catalytic activity

Synthetic secondary adapter scaffold UidA-Link-coh1C9-X-dock2 was designed based on characteristics of the cellulosome of *A. cellulyticus.* It contains a type 1 cohesin domain for binding dock1-fused enzymes, a C-terminal dock2 domain for binding type 2 cohesin domains on primary scaffolds, as well as the entire UidA enzyme at its N-terminus (Fig. 14A). I first sought to verify if the increase in distance between the UidA and dock2 domains in this construct would result in increased β-glucuronidase activity using cell binding assays. When targeted to cells displaying CBD-coh1C3-coh2S1, I observed an approximate 5-fold increase in enzyme activity when compared with cells bound with UidA-dock2, and an approximate 2-fold increase in activity when compared with cells bound with UidA-X-dock2 (Fig. 16). This suggests that both the presence of the "X" module, as well as an increase in distance between the UidA and dock2 domains may contribute to the observed increase in activity. As expected, UidA-Link-coh1C9-X-dock2 failed to bind cells displaying CBD alone (Fig. 16).

4.3.3 UidA-Link-coh1C9-X-dock2 serves as a catalytic adapter scaffold capable of binding LacZ-dock1

I was successful in assembling a two-scaffold bi-enzymatic complex on the surface of *L. lactis* using UidA-Link-coh1C9-X-dock2 as a secondary adapter scaffold capable of binding a type 2 cohesin as well as dock1-containing enzyme. Targeting UidA-Link-coh1C9-X-dock2 to cells displaying CBD-coh1C3-coh2S1 resulted in greater β-glucuronidase activity than when it was targeted to those displaying CBD-coh2S1 (Fig. 17A). This is probably due to a greater number of CBD-coh1C3-coh2S1 scaffolds being functionally displayed on the cell surface compared with CBD-coh2S1, as was previously determined by binding UidA-dock2 to cells displaying these same scaffolds (chapter 3). I assembled a complex consisting of a primary cell-bound scaffold (CBD-coh1C3-coh2S1), a secondary adapter scaffold (UidA-Link-coh1C9-X-dock2) and a dockerin-fused test enzyme (LacZ-dock1) (Fig. 14B). Cells displaying the complexes exhibited an overall decrease in β-glucuronidase activity when LacZ-dock1 was present as compared to cells displaying identical complexes lacking LacZ-dock1 (Fig. 17A). This effect was observed both on cells displaying CBD-coh2S1 as well as those displaying CBD-coh1C3-coh2S1. β-galactosidase activity was three-fold lower when LacZ-dock1 was incorporated into the CBD-coh1C3-coh2S1 / UidA-Link-coh1C9-X-dock2 two-scaffold complex when compared to the activity resulting from its targeting to CBD-coh1C3-coh2S1 alone (Fig. 17B). In order to verify that in fact LacZ-dock1 was targeting the coh1C9 domain on the adapter scaffold, a similar experiment was performed where the primary scaffold contains only a type 2 cohesin (CBD-coh2S1) (Fig. 14B). This experiment demonstrated that LacZ-dock1

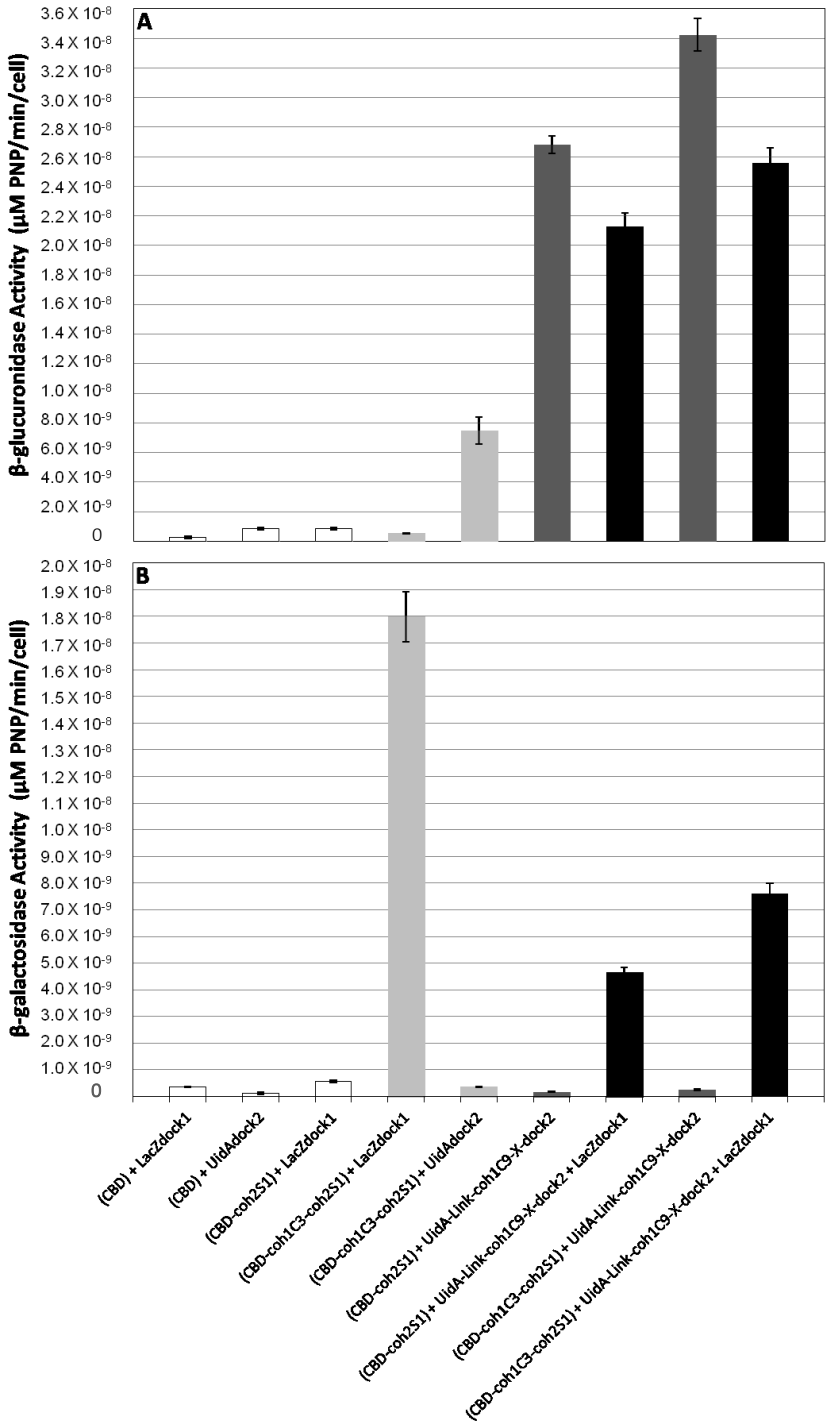

Figure 17. β-glucuronidase (A) and β-galactosidase (B) activity of cells displaying two-level multi-enzyme complexes. Cells displaying primary scaffolds CBD-coh2S1 or CBD-coh1C3-coh2S1 were incubated with secondary adapter scaffold UidA-Link-coh1C9-X-dock2 (dark grey bars) or with both UidA-Link-coh1C9-X-dock2 and LacZ-dock1 (black bars). Cells displaying primary scaffolds incubated with LacZ-dock1 or UidA-dock2 alone (light grey bars) were used as positive controls. Cells displaying CBD (no cohesin) incubated with LacZ-dock1 or UidA-dock2, and cells displaying CBD-coh2S1 incubated with LacZ-dock1 (white bars) were used as negative controls.

would bind the complex only in the presence of UidA-Link-coh1C9-X-dock2 (Fig. 17B), demonstrating the functionality of the coh1C9 domain on the adapter scaffold.

A second "vertically-assembling" adapter scaffold, coh1C1-coh1C2-coh2S1-dock1 (Fig. 14A), was constructed to generate a multi-level enzyme complex by virtue of its ability to self-assemble. The goal was to generate a large number of type 2 cohesin domains for binding UidA-X-dock2 within the multi-level complex. Unfortunately, I was unable to isolate this scaffold by affinity chromatography after it was overproduced in *E. coli* BL21(DE3).

4.3.4 *In vivo* assembly of enzyme complexes on the surface of *L. lactis* by use of a dual-plasmid system

The exogenous addition of scaffolds and enzymes requires purification steps that increase the costs associated with assembling enzyme complexes on the bacterial cell surface. Consolidating enzyme production and complex assembly into a single step would also be advantageous if this system would be used in small and large-scale fermentations. Therefore, I attempted to engineer *L. lactis* to produce both surface displayed scaffolds and secreted enzymes by using a dual-plasmid system where expression of each component would be carried out on respective vectors. Strains of *L. lactis* harboring both the pAW series of vectors expressing primary scaffolds, as well as the pST series of vectors for secretion of enzymes or scaffold-enzyme fusions were analyzed for extracellular enzyme activity (Fig. 18). Vectors designed for the expression and secretion of UidA-X-dock2, UidA-Link-coh1C9 or LacZ-dock1 (Fig. 15A) were transformed in *L. lactis* that displayed CBD-coh1C3-coh2S1, CBD-coh1C3, CBD-coh2S1, or CBD (Fig. 15B), as well

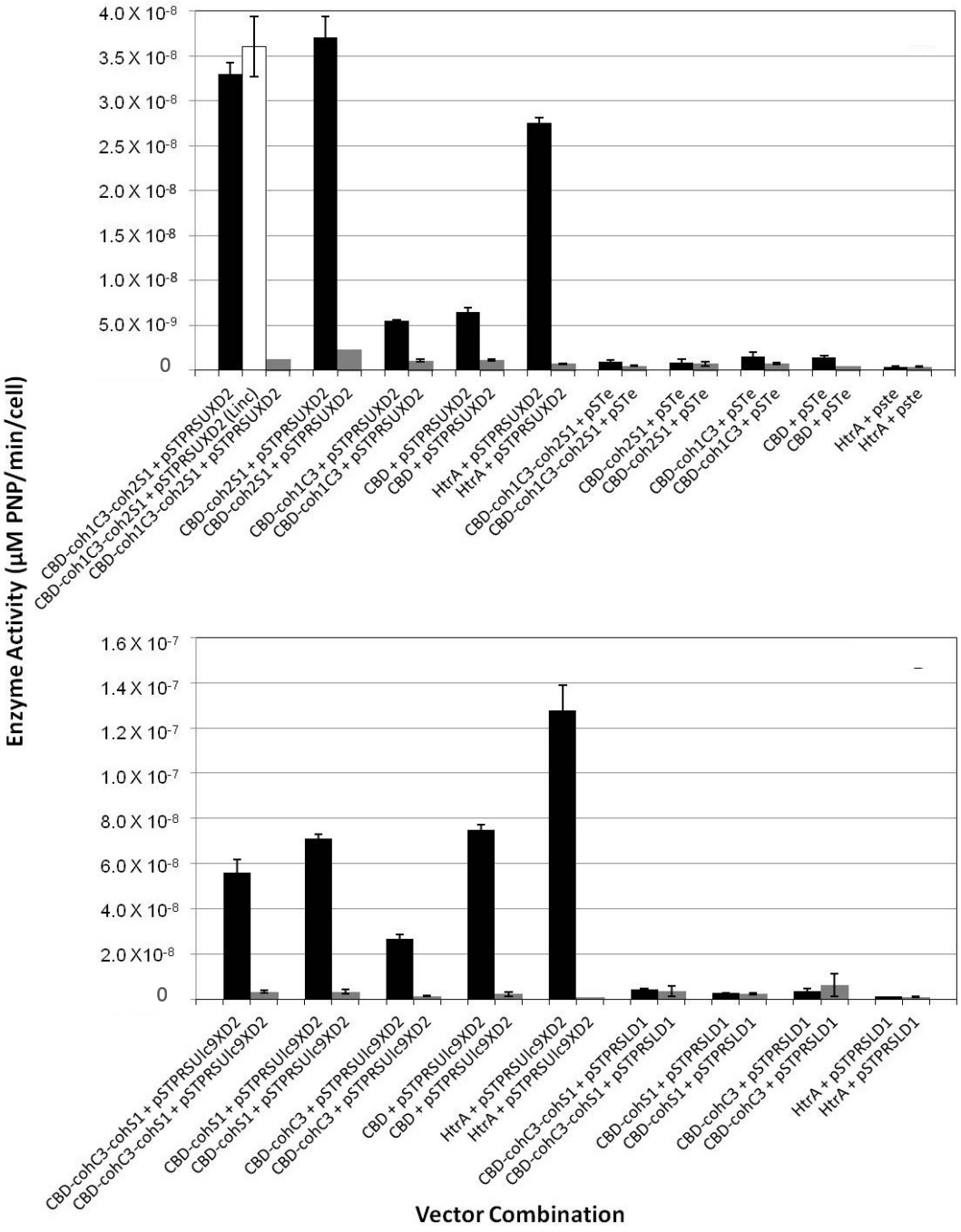

Figure 18. *In vivo* assembly of scaffold/enzyme complexes on the surface of *L. lactis* using a dual-plasmid system. β-glucuronidase (black bars) and β-galactosidase (grey bars) activity of cells both anchoring primary scaffolds and secreting enzyme or secondary scaffold components. A total of four strains displaying different primary scaffolds as well as the *htrA*-NZ9000 plasmid free-strain were transformed with the pST series of vectors for the expression of enzyme-dockerin fusions, secondary adapter scaffolds, or containing no expression cassette (pSTe). A strain displaying a chimeric scaffold and producing UidA-X-dock2 was washed in the presence of 12µg/mL lincomycin to ensure that protein production had stopped during the wash steps following cell harvest (white bars).

as the plasmid-free *htrA*-NZ9000 strain. Although I also sought to co-express LacZ-dock1 and UidA-X-dock2 in *L. lactis* by generating an operon on the pST vector, I was unable to obtain viable transformants, probably due to cellular toxicity.

Strains of *L. lactis* co-expressing UidA-X-dock2 and surface displayed CBD-coh1C3-coh2S1 or CBD-coh2S1 (each of which contains a type 2 cohesin) both demonstrated β-glucuronidase activity on the cell surface (Fig. 18A). Strains co-expressing UidA-X-dock2 and CBD-coh1C3 or CBD (each lacking a type 2 cohesin) did not demonstrate significant levels of β-glucuronidase activity. This suggested that UidA-X-dock2 was specifically binding surface displayed scaffolds only when they contained a corresponding coh2S1 domain. However, strain *htrA*-NZ9000 that lacks a pAW vector also exhibited β-glucuronidase activity approaching 2.8 X 10^{-8} µmol PNP/min/cell when expressing UidA-X-dock2, suggesting that UidA-X-dock2 was non-specifically associating with the cell surface in this strain (Fig. 18A). The higher growth rate of this control strain lacking a plasmid may contribute to excess enzymes becoming "trapped" in the cell wall, a phenomenon that was previously observed when targeting certain proteins for secretion (Wieczorek & Martin, 2010).

UidA-Link-coh1C9-X-dock2 was also successfully co-expressed in *L. lactis* with surface-displayed scaffolds. Strains producing this enzyme-scaffold fusion with surface displayed CBD-coh1C3-coh2S1 or CBD-coh2S1 both demonstrated β-glucuronidase activity. Comparable activity however was also detected on a strain displaying CBD, which lacks a coh2S1 domain, suggesting that this secreted protein remained trapped in the cell

wall of *L. lactis* (Fig. 18B). Once again, a strain lacking a pAW series vector but producing UidA-Link-coh1C9-X-dock2 demonstrated significant levels of β-glucuronidase activity.

L. lactis strains were also engineered to co-express surface displayed scaffolds with LacZ-dock1, however in all cases, no significant β-galactosidase activity was observed (Fig. 18B). I hypothesize that LacZ-dock1 was not successfully secreted since cell fractionation revealed β-galactosidase activity in the cytoplasm but not in the culture medium (data not shown).

The *in vivo* production of both scaffold and enzyme components gives rise to the possibility that residual enzyme may be secreted during the final harvest following the sequential wash steps, prior to the enzyme assay. To determine if any enzyme was being produced by the harvested cells following the wash steps, lincomycin, a protein synthesis inhibitor, was included in the wash buffer of one sample (Fig. 18A, white bar). As expected, no difference in enzyme activity was observed, suggesting that protein production in late stationary phase was not occurring due to resource exhaustion.

Co-cultures consisting of two strains, each producing different primary scaffolds, as well as UidA-X-dock2 or coh1C1-coh1C2-coh2S1-dock1 were mixed in early exponential-phase (OD_{600} of 0.3) and grown for an additional 20 hrs. The goal was to assemble multi-level scaffold complexes on one population of cells, while incorporating UidA-X-dock2 within these complexes via intercellular complementation. In these experiments, UidA activity could not be observed when compared with a monoculture of cells displaying CBD and secreting UidA-X-dock2 (Fig. 19). It is possible that coh1C1-

coh1C2-coh2S1-dock1 was not efficiently secreted or that the cohesin domains remained buried preventing access to UidA-X-dock2.

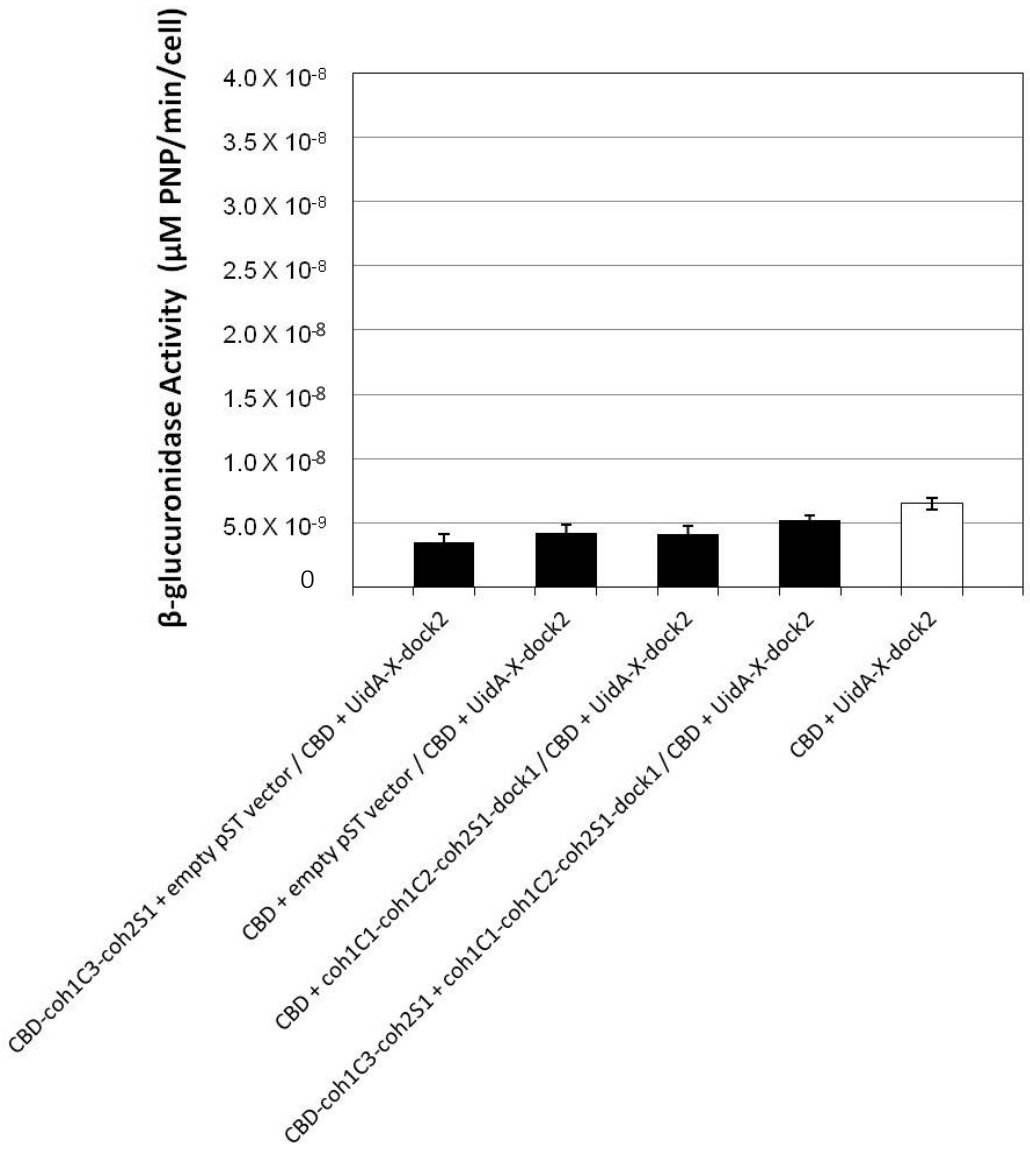

Figure 19. Co-cultures of *L. lactis* towards the production of multi-level enzyme complexes by intercellular complementation. β-glucuronidase activity of resulting co-cultures after wash steps (Black bars). A first strain displayed primary scaffold CBD-coh1C3-coh2S1 or CBD, and secreted either a self-assembling scaffold coh1C1-coh1C2-coh2S1-dock1 or no scaffold (empty vector). A second strain anchored CBD on the cell surface and secreted UidA-X-dock2. As a control, a monoculture of cells anchoring CBD and secreting UidA-X-dock2 was included (white bar).

4.4 Discussion

Engineering microbes to produce recombinant multi-enzyme complexes is an important step in optimizing the efficiency of biochemical processes via substrate channeling or enzyme synergy (Bayer, et al., 2004; Conrado, et al., 2008; Dueber, et al., 2009). In chapter 3, the surface-display of a scaffold chimera consisting of type 1 and type 2 cohesins allowed the targeting of dockerin-fused enzymes to corresponding sites on the scaffold, however less enzyme activity was observed when the enzyme-dockerin fusion was targeted to the type 2 cohesin. In an effort to increase the efficiency of the type 2 cohesin-dockerin interaction, a test enzyme UidA-X-dock2 was targeted to the surface of cells displaying CBD-coh1C3-coh2S1. It has been previously demonstrated that the X module contributes electrostatic interactions necessary for the integrity of the cohS1/X-dock2 complex. This was elucidated based primarily on a computational analysis of the crystal structure of the complex (J. Xu & Smith, 2010). In agreement with these findings, inclusion of the X module in the UidA-X-dock2 fusion and its subsequent binding with cells displaying CBD-coh1C3-coh2S1, resulted in a 2.5-fold increase in activity when compared with identical cells bound with UidA-dock2 (Fig. 16). This strongly suggests that unlike recombinant complexes assembled by type 1 cohesin-dockerin interactions, the X module is an important component to be included in a type 2 dockerin fusion.

I engineered UidA-Link-coh1C9-X-dock2 as a secondary adapter scaffold capable of binding both dock1-fused enzymes and type 2 cohesin-containing scaffolds, and it differs from UidA-X-dock2 by an approximate 200 additional amino acids between the catalytic and dockerin domains, which comprise the coh1C9 domain (Fig. 14A). UidA-

Link-cohC9-X-dock2 successfully bound cells displaying CBD-coh1C3-coh2S1, and the resulting enzyme activity corresponded to a 5- and 2-fold increase when compared with identical cells bound with UidA-dock2 and UidA-X-dock2, respectively (Fig. 16). This suggests that in addition to the presence of the X module, some increase in activity may also be due to a longer distance between the catalytic and dockerin domain in the fusion protein. In a previous study it was observed that the length of a linker between the catalytic module and dock1 domain of enzyme-dockerin fusions had no effect on enzyme activity (Caspi, et al., 2009). In contrast, I observed that two factors may possibly be contributing to the increase in activity measured when binding enzyme-dockerin fusions to the type 2 cohesin found on primary scaffolds: improvement of the type 2 dockerin-cohesin interaction by virtue of the X module, as well as an increase in distance between the catalytic and dockerin domains.

In an effort to increase the number of available type 1 cohesin sites for docking enzymes fused with dock1, the functionality of UidA-Link-coh1C9-X-dock2 as a secondary adapter scaffold was tested. This architecture is inspired by the bacterial cellulosome of *Acetivibrio cellulolyticus*, which is characterized by secondary adapter scaffolds exhibiting catalytic activity (Q. Xu, et al., 2003). The authors who elucidated the structure of the *A. cellulolyticus* cellulosome use the term “secondary” and “adapter” to describe ScaB since it binds only scaffolds and not enzymes, however I use these terms to describe a scaffold which binds both “primary” surface-displayed scaffolds and enzyme components. ScaA is a unique secondary adapter scaffold in that it contains a catalytic domain, cohesin domains for binding dockerin-containing enzymes, as well as an X module and dockerin

domain at its C-terminus which show high degrees of homology with components of CipA from *C. thermocellum* (Q. Xu, et al., 2003). The specificity of the dockerin/cohesin interaction between ScaA and ScaB may therefore be analogous to the type 2 interaction in *C. thermocellum*, and further supports the hypothesis that the X module is important for certain dockerin-cohesin interactions (Q. Xu, et al., 2003). UidA-Link-coh1C9-X-dock2 successfully bound cells displaying CBD-coh1C3-coh2S1 or CBD-coh2S1, however greater β-glucuronidase activity was observed when the adapter bound cells displaying coh1C3-coh2S1. This is in accordance with my previous observation that *L. lactis* secretes and displays more functional CBD-coh1C3-coh2S1 scaffolds on its surface compared with CBD-coh2S1 (chapter 3). To assemble a two-level complex where enzymes can bind both primary and secondary scaffolds, a trimeric complex consisting of CBD-coh1C3-coh2S1, UidA-Link-coh1C9-X-dock2, and LacZ-dock1 was assembled on the cell surface (Fig. 14B). The presence of LacZ-dock1 within the surface displayed complex resulted in less UidA activity (Fig. 17A) when compared with an identical complex lacking LacZ-dock1. When assembling a trimeric complex consisting of CBD-coh2S1 as the primary surface displayed scaffold, a similar decrease in UidA activity was observed when LacZ-dock1 was included in the complex. This suggests that this effect was independent of the primary scaffold chosen, but rather resulted from the interaction between UidA-Link-coh1C9-X-dock2 and LacZ-dock1. I hypothesize this was due to molecular crowding of the two catalytic modules, where steric hindrance of UidA may result from the presence of the larger LacZ enzyme. This is reminiscent of previous findings where simultaneously targeting LacZ and UidA fusions to a common scaffold resulted in decreased UidA activity (chapter 3). It is

also possible that conformational changes within the secondary adapter scaffold may occur upon binding LacZ-dock1, ultimately contributing to this decrease in UidA activity.

The ultimate objective of incorporating a second scaffold into such complexes was to increase the number of "docking" sites for LacZ-dock1 to bind. I hypothesized that such an increase in docking sites for LacZ-dock1 would also result in an increase in LacZ activity. Interestingly, the result was a threefold decrease in LacZ activity (Fig. 17B). It is possible that the type 1 cohesin-dockerin interaction on UidA-Link-coh1C3-X-dock2 is less efficient than on CBD-coh1C3-coh2S1, and in addition, it is likely that the large adapter scaffold may have resulted in decreased accessibility of coh1C3 to LacZ-dock1. Using cells displaying CBD-coh1C3-coh2S1, it is not possible to differentiate between LacZ-dock1 binding with the primary vs. secondary scaffold since both contain appropriate docking sites. I therefore assembled a two-level complex using CBD-coh2S1 as a primary scaffold so that any observed β-galactosidase activity would be resulting from LacZ-dock1 associating with UidA-Link-coh1C9-X-dock2 (Fig. 14B). Targeting UidA-Link-coh1C9-X-dock2 and LacZ-dock1 to surface-displayed CBD-coh2S1 yielded cells exhibiting both β-glucuronidase (Fig. 17A) and β-galactosidase activity (Fig. 17B), demonstrating that LacZ-dock1 was capable of binding the coh1C9 site on the secondary adapter scaffold. This provides, to my knowledge, the first description of a recombinant multi-enzyme complex using a two-level scaffold platform.

To increase the number of type 2 cohesin sites with the complex, I engineered an adapter scaffold capable of self-assembling via intermolecular interactions into multi-level complexes. Since these scaffolds assemble perpendicular to the cell surface, I

describe these as “vertically-assembling” scaffold complexes. I attempted to over-express and purify scaffold coh1C1-coh1C2-coh2S1-dock1 from *E. coli*, however the protein was not detected in any of the elution fractions after binding to the Ni-NTA resin. Although it was not verified if over-expressed proteins were present in the total cell lysate, I hypothesize that since over-expression of proteins using the pET28a vector results in their accumulation in the cytoplasm, the scaffolds may have self-assembled prematurely or generated inclusion bodies, whereby such protein aggregates prevented their ability to be isolated.

The costs associated with over-producing and purifying enzyme components before their exogenous addition to scaffold displaying cells generate much appeal to have a strain capable of producing all components of the complex *in vivo*. A next logical step was therefore to express enzymes and adapter scaffolds *in vivo* in strains of *L. lactis* already displaying primary scaffolds, using a dual-plasmid system (Fig. 15). I also hypothesized that the *in vivo* expression of cohC1-cohC2-cohS1-dock1 in *L. lactis* may possibly circumvent the shortcomings observed when attempting to isolate this protein in *E. coli*, since intracellular chaperones of the secretion pathway should prevent folding until export occurs. Strains displaying CBD-coh1C3-coh2S1, CBD-coh1C3, CBD-coh2S1 or CBD (Fig. 15B), as well as the plasmid free strain *L. lactis htrA*-NZ9000 were engineered to express either LacZ-dock1, UidA-X-dock2, UidA-Link-coh1C9-X-dock2 or coh1C1-coh1C2-coh2S1-dock1 (Fig. 15A) by means of this dual-plasmid system. All strains containing both a pAW and pST series vector demonstrated slower growth on solid and liquid media, reaching lower final optical densities when compared with strains harboring only one

vector or strains containing a pAW series vector and pSTe (empty vector). This suggested a larger physiological burden on cells from over-exertion of the secretion machinery leading to toxicity, possibly due to problems with cell-wall biosynthesis at the cellular septum (Narita, et al., 2006). Unfortunately, I was unable to successfully transform *L. lactis* with a vector containing an operon encoding both UidA-X-dock2 and LacZ-dock1, most probably due to cellular toxicity resulting from the simultaneous over-expression of these two fusion enzymes.

Strains of *L. lactis* co-expressing UidA-X-dock2 and surface displayed scaffolds containing a coh2S1 domain demonstrated β-glucuronidase activity on the cell surface, confirming that the enzyme component was successfully produced and secreted. Strains expressing UidA-X-dock2 and scaffolds lacking coh2S1 did not demonstrate significant levels of β-glucuronidase activity. This observation suggested the possibility that UidA-X-dock2 was specifically bound to scaffolds on the cell surface by means of the coh2S1 domain. However, strain *htrA*-NZ9000, which lacks a pAW vector, also exhibited β-glucuronidase activity approaching 2.8 X 10^{-8} µmol PNP/min/cell when expressing UidA-X-dock2, demonatrating that the enzyme fusion was non-specifically associating with the cell surface of this strain (Fig. 18A). This pAW vector-lacking strain exhibited significantly higher growth rates and final optical densities due to less physiological burden, a characteristic that may have resulted in excess enzymes becoming “trapped” in the cell wall. This effect has been documented previously in this organism (Wieczorek & Martin, 2010).

UidA-Link-coh1C9-X-dock2 was successfully co-expressed with surface displayed scaffolds, as demonstrated by β-glucuronidase activity on the surface of cells. Once again, strain *htrA*-NZ9000, which lacks a pAW vector, demonstrated β-glucuronidase activity on the cell surface, suggesting that the enzyme/scaffold fusion protein remained trapped in the cell wall due to higher growth rates. In this case, cells co-expressing surface displayed CBD and UidA-Link-coh1C9-X-dock2 also demonstrated significant levels of β-glucuronidase activity on the cell surface, suggesting that UidA-Link-coh1C9-X-dock2 is potentially more prone to non-specifically associating with the cell wall when compared to UidA-X-dock2, or that its expression and/or secretion efficiency is higher. It was previously observed that certain modules included within secreted recombinant proteins may contribute to their remaining trapped in the cell wall of *L. lactis* (Wieczorek & Martin, 2010). In that study, fusion of fragments of CipA with enzyme NucA resulted in their becoming trapped in the cell wall, whereas removal of NucA alleviated the effect of the non-specific binding. Nonetheless, the goal of co-expressing both scaffold and UidA-fusions in a single strain of *L. lactis* was achieved. As an additional control to verify if residual proteins were actively being produced and secreted during the wash steps prior to cell harvest, one strain co-expressing CBD-coh1C3-coh2S1 and UidA-X-dock2 was washed in the presence of lincomycin, an inhibitor of protein biosynthesis. No significant differences in β-glucuronidase activity were observed, suggesting that as expected, protein synthesis halted in late stationary phase.

Strains displaying scaffolds on the cell surface and harboring a pST vector containing the ORF encoding Lac-dock1 fusion protein did not demonstrate any

observable β-galactosidase activity. Cell fractionation revealed that β-galactosidase activity was detectable in the cytoplasm but not in the culture medium, suggesting that LacZ-dock1 was not successfully secreted. In an additional proof-of-concept experiment, I attempted to generate vertically-assembling complexes by intercellular complementation (Arai, et al., 2007; Tsai, et al., 2010). Co-cultures contained one population of cells displaying primary scaffold CBD-coh1C3-coh2S1 or CBD, and secreting coh1C1-coh1C2-coh2S1-dock1 or harboring empty vector pSTe, while a second culture displayed CBD and secreted UidA-X-dock2. In all cases, no significant β-glucuronidase activity was observed when compared with a monoculture of cells displaying CBD and secreting UidA-X-dock2 (Fig. 19). Possible explanations include coh1C1-coh1C2-coh2S1-dock1 not being functionally expressed, secreted, and/or anchored in one population of cells, or that UidA-X-dock2 was not successfully secreted and released from the second population of cells. It is also possible that scaffold coh1C1-coh1C2-coh2S1-dock1 generated a cyclic scaffold incapable of interacting with other scaffolds or UidA-X-dock2 fusions by folding upon itself, whereby dock1 may have interacted with either coh1C1 or coh1C2 on the same scaffold. Further investigation into the possibility of generating such multi-level scaffolds by intercellular complementation in this organism is required.

To my knowledge, this study describes the first successful assembly of a multi-enzyme complex on the surface of an organism whose assembly is mediated by both primary anchored scaffolds as well as secondary adapter scaffolds. In addition, this work provides insights into the function of the X module in a recombinant type 2 cohesin-dockerin interaction. Although the use of secondary scaffolds did not increase the

amount of enzymes within the complex, I envision that improvements and manipulation of this strategy will yield novel architectural permutations of enzyme complexes, which may be useful in a number of biotechnological applications.

4.5 Conclusions

The past decade has been characterized by significant advances in the field of recombinant cellulosome production by microorganisms towards the direct conversion of cellulosic substrates to bioethanol. In some cases, microorganisms such as yeast and *B. subtilis* have been engineered to utilize cellulosic substrates as a carbon source for hydrolysis and subsequent fermentation of released sugars into this desirable biofuel. While optimizing the combinations of enzymes within such complexes has been a successful strategy to improving the overall activity of recombinant cellulosomes, no recombinant cellulosomes produced by host organisms have demonstrated increased activity than their natural counterparts (Fierobe, et al., 2002; Fierobe, et al., 2005; Mingardon, et al., 2007). Therefore, the possibility of generating novel architectures of recombinant cellulosomes offers another avenue towards engineering complexes, which demonstrate activity equal to or greater than such naturally occurring ones. In addition, the metabolic diversity of microorganisms that produce commodity chemicals other than bioethanol has made the development of multi-enzyme systems in such organisms of particular appeal. This study therefore demonstrates that it is feasible to generate two-level enzyme complexes on the surface of an industrially relevant bacterium, *L. lactis*. I envision that this area of research may prove beneficial in the further engineering of

organisms to carry out complex processes requiring the benefits associated with enzyme-substrate and enzyme-microbe synergy.

Chapter 5

Conclusions and Suggestions for Future Work

This thesis describes the stepwise approach towards anchoring multi-enzyme complexes on the surface of *L. lactis* inspired by native cellulosome structures. I was successful in demonstrating that it is feasible to anchor fragments of CipA on the cell surface and that the controlled expression of the proteins results in decreased cellular toxicity. I next demonstrated that by fusing a reporter enzyme, UidA, with a dock1 domain I could successfully target this enzyme fusion to the displayed scaffolds. Next, the engineering of scaffold chimeras consisting of both type 1 and type 2 cohesin domains was achieved, and fusion of UidA with dock 1 or dock2 resulted in its successful targeting to specific corresponding sites on the displayed scaffolds. By fusing a second enzyme, LacZ, with either a dock1 or dock2 domain, I was able to assemble bi-enzymatic complexes, and further investigated the parameters of assembly by targeting enzyme pairs simultaneously or sequentially to the displayed scaffolds. By creating a UidA-X-dock2 fusion, I also generated insights into the possible importance of the X module in maintaining the integrity of a recombinant type 2 dockerin-cohesin interaction. By use of a secondary adapter scaffold with catalytic activity, I also demonstrated that it is feasible to target a dockerin-enzyme fusion to such an adapter scaffold, resulting in bi-enzymatic complexes. By using a dual-plasmid system, both enzyme and scaffold components were succesfully produced and secreted *in vivo*.

Type 1 dockerins interact with any of the nine type 1 cohesins located on CipA. Therefore, the mechanisms governing how enzymes assemble onto such scaffolds in

nature yielding the resulting enzymatic composition and organization of such complexes remains poorly understood. The results observed in this study, when loading enzymes of different size and character either sequentially or simultaneously, suggests that steric factors may play a role in complex assembly. It is possible that similar factors govern the assembly of cellulosomes in *C. thermocellum*.

I envision that a major avenue to explore further would be developing adapter scaffolds to accommodate more enzymes, resulting in the assembly of multi-level recombinant cellulosomes. Changing the type of dockerin domain present on the adapter scaffold and/or changing the order of the cohesin domains located within it could possibly achieve this. The use of a third dockerin-cohesin pair of divergent specificity may also prove useful in this effort. In addition, the fusion of cellulases with dock1 or X-dock2 and subsequent expression by means of the dual-plasmid system could be pursued in order to verify if such a strategy could be optimized in order to bestow cellulolytic activity in *L. lactis*. In order to achieve this, mesophilic cellulases could potentially be fused with appropriate dockerins and incorporated into such complexes. Combining cellulases with different complementary activities within such complexes could potentially yield a strain of *L. lactis* capable of degrading cellulose in order to release fermentable sugars. Such an achievement may yield great potential in utilizing cellulose as feedstock towards the production of commodity chemicals such as organic acids in order to dramatically decrease costs of production. The development of an industrial strain to perform this role would probably require increasing the amount of total enzyme complexes on the cell surface. Therefore, I envision that engineering the strains to increase overall secretion

efficiency would be beneficial. Developing this system in an HtrA+ strain may also help increase secretion efficiency, however proteolysis of the displayed recombinant complexes should be analyzed if this strategy should be adopted.

Another possible avenue to explore would be to intracellularly produce such scaffolds alongside enzyme-dockerin fusions in order to optimize the metabolic flux of desirable pathways towards the efficient production of commodity chemicals in this organism. The system and platform strains described in this thesis offer numerous possibilities for exploration, in addition to providing the necessary building blocks for the assembly of custom-designed complexes with precise enzymatic compositions simply by fusing appropriate enzymes with corresponding dockerins.

References

Adams, J. J., Pal, G., Jia, Z., & Smith, S. P. (2006). Mechanism of bacterial cell-surface attachment revealed by the structure of cellulosomal type II cohesin-dockerin complex. *Proc Natl Acad Sci U S A, 103*(2), 305-310.

Anderson, T. D., Robson, S. A., Jiang, X. W., Malmirchegini, G. R., Fierobe, H. P., Lazazzera, B. A., et al. (2011). Assembly of minicellulosomes on the surface of *Bacillus subtilis*. *Appl Environ Microbiol, 77*(14), 4849-4858.

Arai, T., Matsuoka, S., Cho, H. Y., Yukawa, H., Inui, M., Wong, S. L., et al. (2007). Synthesis of *Clostridium cellulovorans* minicellulosomes by intercellular complementation. *Proc Natl Acad Sci U S A, 104*(5), 1456-1460.

Atsumi, S., Hanai, T., & Liao, J. C. (2008). Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature, 451*(7174), 86-89.

Avall-Jaaskelainen, S., Kyla-Nikkila, K., Kahala, M., Miikkulainen-Lahti, T., & Palva, A. (2002). Surface display of foreign epitopes on the *Lactobacillus brevis* S-layer. *Appl Environ Microbiol, 68*(12), 5943-5951.

Avall-Jaaskelainen, S., Lindholm, A., & Palva, A. (2003). Surface display of the receptor-binding region of the *Lactobacillus brevis* S-layer protein in *Lactococcus lactis* provides nonadhesive lactococci with the ability to adhere to intestinal epithelial cells. *Appl Environ Microbiol, 69*(4), 2230-2236.

Axelsson, L., Lindstad, G., & Naterstad, K. (2003). Development of an inducible gene expression system for *Lactobacillus sakei*. *Lett Appl Microbiol, 37*(2), 115-120.

Bayer, E. A., Belaich, J. P., Shoham, Y., & Lamed, R. (2004). The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annu Rev Microbiol, 58*, 521-554.

Bayer, E. A., Kenig, R., & Lamed, R. (1983). Adherence of *Clostridium thermocellum* to cellulose. *J Bacteriol, 156*(2), 818-827.

Bayer, E. A., Setter, E., & Lamed, R. (1985). Organization and distribution of the cellulosome in *Clostridium thermocellum*. *J Bacteriol, 163*(2), 552-559.

Bermudez-Humaran, L. G., Cortes-Perez, N. G., Le Loir, Y., Alcocer-Gonzalez, J. M., Tamez-Guerra, R. S., de Oca-Luna, R. M., et al. (2004). An inducible surface presentation system improves cellular immunity against human papillomavirus type 16 E7 antigen in mice after nasal administration with recombinant lactococci. *J Med Microbiol, 53*(Pt 5), 427-433.

Bermudez-Humaran, L. G., Langella, P., Commissaire, J., Gilbert, S., Le Loir, Y., L'Haridon, R., et al. (2003). Controlled intra- or extracellular production of staphylococcal nuclease and ovine omega interferon in *Lactococcus lactis*. *FEMS Microbiol Lett, 224*(2), 307-313.

Bermudez-Humaran, L. G., Langella, P., Miyoshi, A., Gruss, A., Guerra, R. T., Montes de Oca-Luna, R., et al. (2002). Production of human papillomavirus type 16 E7 protein in *Lactococcus lactis*. *Appl Environ Microbiol, 68*(2), 917-922.

Bolotin, A., Wincker, P., Mauger, S., Jaillon, O., Malarme, K., Weissenbach, J., et al. (2001). The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. lactis IL1403. *Genome Res, 11*(5), 731-753.

Callegari, M. L., Riboli, B., Sanders, J. W., Cocconcelli, P. S., Kok, J., Venema, G., et al. (1998). The S-layer gene of *Lactobacillus helveticus* CNRZ 892: cloning, sequence and heterologous expression. *Microbiology, 144 (Pt 3)*, 719-726.

Caspi, J., Barak, Y., Haimovitz, R., Irwin, D., Lamed, R., Wilson, D. B., et al. (2009). Effect of linker length and dockerin position on conversion of a *Thermobifida fusca* endoglucanase to the cellulosomal mode. *Appl Environ Microbiol, 75*(23), 7335-7342.

Chatel, J. M., Langella, P., Adel-Patient, K., Commissaire, J., Wal, J. M., & Corthier, G. (2001). Induction of mucosal immune response after intranasal or oral inoculation of mice with *Lactococcus lactis* producing bovine beta-lactoglobulin. *Clin Diagn Lab Immunol, 8*(3), 545-551.

Cho, H. Y., Yukawa, H., Inui, M., Doi, R. H., & Wong, S. L. (2004). Production of minicellulosomes from *Clostridium cellulovorans* in *Bacillus subtilis* WB800. *Appl Environ Microbiol, 70*(9), 5704-5707.

Choi, S. K., & Ljungdahl, L. G. (1996). Dissociation of the cellulosome of *Clostridium thermocellum* in the presence of ethylenediaminetetraacetic acid occurs with the formation of trucated polypeptides. *Biochemistry, 35*(15), 4897-4905.

Ciruela, A., Gilbert, H. J., Ali, B. R., & Hazlewood, G. P. (1998). Synergistic interaction of the cellulosome integrating protein (CipA) from *Clostridium thermocellum* with a cellulosomal endoglucanase. *FEBS Lett, 422*(2), 221-224.

Conrado, R. J., Varner, J. D., & DeLisa, M. P. (2008). Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. *Curr Opin Biotechnol, 19*(5), 492-499.

Cortes-Perez, N. G., Azevedo, V., Alcocer-Gonzalez, J. M., Rodriguez-Padilla, C., Tamez-Guerra, R. S., Corthier, G., et al. (2005). Cell-surface display of E7 antigen from human papillomavirus type-16 in *Lactococcus lactis* and in *Lactobacillus plantarum* using a new cell-wall anchor from lactobacilli. *J Drug Target, 13*(2), 89-98.

Cortes-Perez, N. G., Poquet, I., Oliveira, M., Gratadoux, J. J., Madsen, S. M., Miyoshi, A., et al. (2006). Construction and characterization of a *Lactococcus lactis* strain deficient in intracellular ClpP and extracellular HtrA proteases. *Microbiology, 152*(Pt 9), 2611-2618.

de Vos, W. M. (1999). Gene expression systems for lactic acid bacteria. *Curr Opin Microbiol, 2*(3), 289-295.

Delebecque, C. J., Lindner, A. B., Silver, P. A., & Aldaye, F. A. (2011). Organization of intracellular reactions with rationally designed RNA assemblies. *Science, 333*(6041), 470-474.

Dieye, Y., Hoekman, A. J., Clier, F., Juillard, V., Boot, H. J., & Piard, J. C. (2003). Ability of *Lactococcus lactis* to export viral capsid antigens: a crucial step for development of live vaccines. *Appl Environ Microbiol, 69*(12), 7281-7288.

Dieye, Y., Usai, S., Clier, F., Gruss, A., & Piard, J. C. (2001). Design of a protein-targeting system for lactic acid bacteria. *J Bacteriol, 183*(14), 4157-4166.

Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., Ullal, A. V., et al. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. *Nat Biotechnol, 27*(8), 753-759.

Duwat, P., Ehrlich, S. D., & Gruss, A. (1995). The *recA* gene of *Lactococcus lactis*: characterization and involvement in oxidative and thermal stress. *Mol Microbiol, 17*(6), 1121-1131.

Enouf, V., Langella, P., Commissaire, J., Cohen, J., & Corthier, G. (2001). Bovine rotavirus nonstructural protein 4 produced by *Lactococcus lactis* is antigenic and immunogenic. *Appl Environ Microbiol, 67*(4), 1423-1428.

Farhi, M., Marhevka, E., Masci, T., Marcos, E., Eyal, Y., Ovadis, M., et al. (2011). Harnessing yeast subcellular compartments for the production of plant terpenoids. *Metab Eng, 13*(5), 474-481.

Farrell, A. E., Plevin, R. J., Turner, B. T., Jones, A. D., O'Hare, M., & Kammen, D. M. (2006). Ethanol can contribute to energy and environmental goals. *Science, 311*(5760), 506-508.

Fierobe, H. P., Bayer, E. A., Tardif, C., Czjzek, M., Mechaly, A., Belaich, A., et al. (2002). Degradation of cellulose substrates by cellulosome chimeras. Substrate targeting versus proximity of enzyme components. *J Biol Chem, 277*(51), 49621-49630.

Fierobe, H. P., Mechaly, A., Tardif, C., Belaich, A., Lamed, R., Shoham, Y., et al. (2001). Design and production of active cellulosome chimeras. Selective incorporation of dockerin-containing enzymes into defined functional complexes. *J Biol Chem, 276*(24), 21257-21261.

Fierobe, H. P., Mingardon, F., Mechaly, A., Belaich, A., Rincon, M. T., Pages, S., et al. (2005). Action of designer cellulosomes on homogeneous versus complex substrates: controlled incorporation of three distinct enzymes into a defined trifunctional scaffoldin. *J Biol Chem, 280*(16), 16325-16334.

Foucaud-Scheunemann, C., & Poquet, I. (2003). HtrA is a key factor in the response to specific stress conditions in *Lactococcus lactis*. *FEMS Microbiol Lett, 224*(1), 53-59.

Freitas, D. A., Leclerc, S., Miyoshi, A., Oliveira, S. C., Sommer, P. S., Rodrigues, L., et al. (2005). Secretion of *Streptomyces tendae* antifungal protein 1 by *Lactococcus lactis*. *Braz J Med Biol Res, 38*(11), 1585-1592.

Garcia-Campayo, V., & Beguin, P. (1997). Synergism between the cellulosome-integrating protein CipA and endoglucanase CelD of *Clostridium thermocellum*. *J Biotechnol, 57*(1-3), 39-47.

Gerngross, U. T., Romaniec, M. P., Kobayashi, T., Huskisson, N. S., & Demain, A. L. (1993). Sequencing of a *Clostridium thermocellum* gene (cipA) encoding the cellulosomal SL-protein reveals an unusual degree of internal homology. *Mol Microbiol, 8*(2), 325-334.

Gold, N. D., & Martin, V. J. (2007). Global view of the *Clostridium thermocellum* cellulosome revealed by quantitative proteomic analysis. *J Bacteriol, 189*(19), 6787-6795.

Gottesman, S. (1996). Proteases and their targets in *Escherichia coli*. *Annu Rev Genet, 30*, 465-506.

Hammel, M., Fierobe, H. P., Czjzek, M., Kurkal, V., Smith, J. C., Bayer, E. A., et al. (2005). Structural basis of cellulosome efficiency explored by small angle X-ray scattering. *J Biol Chem, 280*(46), 38562-38568.

Hernandez, I., Molenaar, D., Beekwilder, J., Bouwmeester, H., & van Hylckama Vlieg, J. E. (2007). Expression of plant flavor genes in *Lactococcus lactis*. *Appl Environ Microbiol, 73*(5), 1544-1552.

Holo, H., & Nes, I. F. (1989). High-Frequency Transformation, by Electroporation, of *Lactococcus lactis* subsp. cremoris Grown with Glycine in Osmotically Stabilized Media. *Appl Environ Microbiol, 55*(12), 3119-3123.

Holo, H., & Nes, I. F. (1995). Transformation of Lactococcus by electroporation. *Methods Mol Biol, 47*, 195-199.

Ito, J., Kosugi, A., Tanaka, T., Kuroda, K., Shibasaki, S., Ogino, C., et al. (2009). Regulation of the display ratio of enzymes on the *Saccharomyces cerevisiae* cell surface by the immunoglobulin G and cellulosomal enzyme binding domains. *Appl Environ Microbiol, 75*(12), 4149-4154.

Kataeva, I., Guglielmi, G., & Beguin, P. (1997). Interaction between *Clostridium thermocellum* endoglucanase CelD and polypeptides derived from the cellulosome-integrating protein CipA: stoichiometry and cellulolytic activity of the complexes. *Biochem J, 326 (Pt 2)*, 617-624.

Kleerebezem, M., Boels, I. C., Groot, M. N., Mierau, I., Sybesma, W., & Hugenholtz, J. (2002). Metabolic engineering of *Lactococcus lactis*: the impact of genomics and metabolic modelling. *J Biotechnol, 98*(2-3), 199-213.

Kosugi, A., Amano, Y., Murashima, K., & Doi, R. H. (2004). Hydrophilic domains of scaffolding protein CbpA promote glycosyl hydrolase activity and localization of cellulosomes to the cell surface of *Clostridium cellulovorans*. *J Bacteriol, 186*(19), 6351-6359.

Kruus, K., Lua, A. C., Demain, A. L., & Wu, J. H. (1995). The anchorage function of CipA (CelL), a scaffolding protein of the *Clostridium thermocellum* cellulosome. *Proc Natl Acad Sci U S A, 92*(20), 9254-9258.

Kuipers, O. P., Beerthuyzen, M. M., Siezen, R. J., & De Vos, W. M. (1993). Characterization of the nisin gene cluster nisABTCIPR of *Lactococcus lactis*. Requirement of expression of the *nisA* and *nisI* genes for development of immunity. *Eur J Biochem, 216*(1), 281-291.

Langella, P., & Le Loir, Y. (1999). Heterologous protein secretion in *Lactococcus lactis*: a novel antigen delivery system. *Braz J Med Biol Res, 32*(2), 191-198.

Le Loir, Y., Azevedo, V., Oliveira, S. C., Freitas, D. A., Miyoshi, A., Bermudez-Humaran, L. G., et al. (2005). Protein secretion in *Lactococcus lactis* : an efficient way to increase the overall heterologous protein production. *Microb Cell Fact, 4*(1), 2.

Le Loir, Y., Gruss, A., Ehrlich, S. D., & Langella, P. (1998). A nine-residue synthetic propeptide enhances secretion efficiency of heterologous proteins in *Lactococcus lactis*. *J Bacteriol, 180*(7), 1895-1903.

Le Loir, Y., Nouaille, S., Commissaire, J., Bretigny, L., Gruss, A., & Langella, P. (2001). Signal peptide and propeptide optimization for heterologous protein secretion in *Lactococcus lactis*. *Appl Environ Microbiol, 67*(9), 4119-4127.

Lee, H., Deloache, W. C., & Dueber, J. E. (2011). Spatial organization of enzymes for metabolic engineering. *Metab Eng, 14*(3), 242-251.

Lee, P., & Faubert, G. M. (2006). Expression of the *Giardia lamblia* cyst wall protein 2 in *Lactococcus lactis*. *Microbiology, 152*(Pt 7), 1981-1990.

Lee, S. K., Chou, H., Ham, T. S., Lee, T. S., & Keasling, J. D. (2008). Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. *Curr Opin Biotechnol, 19*(6), 556-563.

Leenhouts, K., Buist, G., & Kok, J. (1999). Anchoring of proteins to lactic acid bacteria. *Antonie Van Leeuwenhoek, 76*(1-4), 367-376.

Leibovitz, E., & Beguin, P. (1996). A new type of cohesin domain that specifically binds the dockerin domain of the *Clostridium thermocellum* cellulosome-integrating protein CipA. *J Bacteriol, 178*(11), 3077-3084.

Lemaire, M., Ohayon, H., Gounon, P., Fujino, T., & Beguin, P. (1995). OlpB, a new outer layer protein of *Clostridium thermocellum*, and binding of its S-layer-like domains to components of the cell envelope. *J Bacteriol, 177*(9), 2451-2459.

Liao, H., Zhang, X. Z., Rollin, J. A., & Zhang, Y. H. (2011). A minimal set of bacterial cellulases for consolidated bioprocessing of lignocellulose. *Biotechnol J, 6*(11), 1409-1418.

Lilly, M., Fierobe, H. P., van Zyl, W. H., & Volschenk, H. (2009). Heterologous expression of a Clostridium minicellulosome in *Saccharomyces cerevisiae*. *FEMS Yeast Res, 9*(8), 1236-1249.

Linares, D. M., Geertsma, E. R., & Poolman, B. (2010). Evolved *Lactococcus lactis* strains for enhanced expression of recombinant membrane proteins. *J Mol Biol, 401*(1), 45-55.

Lindholm, A., Smeds, A., & Palva, A. (2004). Receptor binding domain of *Escherichia coli* F18 fimbrial adhesin FedF can be both efficiently secreted and surface displayed in a functional form in *Lactococcus lactis*. *Appl Environ Microbiol, 70*(4), 2061-2071.

Lowell, G. H., Ballou, W. R., Smith, L. F., Wirtz, R. A., Zollinger, W. D., & Hockmeyer, W. T. (1988). Proteosome-lipopeptide vaccines: enhancement of immunogenicity for malaria CS peptides. *Science, 240*(4853), 800-802.

Lowell, G. H., Smith, L. F., Seid, R. C., & Zollinger, W. D. (1988). Peptides bound to proteosomes via hydrophobic feet become highly immunogenic without adjuvants. *J Exp Med, 167*(2), 658-663.

Lu, Y., Zhang, Y. H., & Lynd, L. R. (2006). Enzyme-microbe synergy during cellulose hydrolysis by *Clostridium thermocellum*. *Proc Natl Acad Sci U S A, 103*(44), 16165-16169.

Lynd, L. (1996). Overview and evaluation of fuel ethanol from cellulosic biomass: technology, economics, the environment, and policy. *Annu. Rev. Energy. Environ., 21*, 403-465.

Lynd, L. R., van Zyl, W. H., McBride, J. E., & Laser, M. (2005). Consolidated bioprocessing of cellulosic biomass: an update. *Curr Opin Biotechnol, 16*(5), 577-583.

Lynd, L. R., Weimer, P. J., van Zyl, W. H., & Pretorius, I. S. (2002). Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol Mol Biol Rev, 66*(3), 506-577, table of contents.

Lytle, B., Myers, C., Kruus, K., & Wu, J. H. (1996). Interactions of the CelS binding ligand with various receptor domains of the *Clostridium thermocellum* cellulosomal scaffolding protein, CipA. *J Bacteriol, 178*(4), 1200-1203.

Mayer, F., Coughlan, M. P., Mori, Y., & Ljungdahl, L. G. (1987). Macromolecular Organization of the Cellulolytic Enzyme Complex of *Clostridium thermocellum* as Revealed by Electron Microscopy. *Appl Environ Microbiol, 53*(12), 2785-2792.

Mierau, I., & Kleerebezem, M. (2005). 10 years of the nisin-controlled gene expression system (NICE) in *Lactococcus lactis*. *Appl Microbiol Biotechnol, 68*(6), 705-717.

Mingardon, F., Chanal, A., Tardif, C., Bayer, E. A., & Fierobe, H. P. (2007). Exploration of new geometries in cellulosome-like chimeras. *Appl Environ Microbiol, 73*(22), 7138-7149.

Miot, M., & Betton, J. M. (2004). Protein quality control in the bacterial periplasm. *Microb Cell Fact, 3*(1), 4.

Miron, J., Ben-Ghedalia, D., & Morrison, M. (2001). Invited review: adhesion mechanisms of rumen cellulolytic bacteria. *J Dairy Sci, 84*(6), 1294-1309.

Mitchell, W. J. (1998). Physiology of carbohydrate to solvent conversion by clostridia. *Adv Microb Physiol, 39*, 31-130.

Miyoshi, A., Bermudez-Humaran, L. G., Ribeiro, L. A., Le Loir, Y., Oliveira, S. C., Langella, P., et al. (2006). Heterologous expression of *Brucella abortus* GroEL heat-shock protein in *Lactococcus lactis*. *Microb Cell Fact, 5*, 14.

Miyoshi, A., Poquet, I., Azevedo, V., Commissaire, J., Bermudez-Humaran, L., Domakova, E., et al. (2002). Controlled production of stable heterologous proteins in *Lactococcus lactis*. *Appl Environ Microbiol, 68*(6), 3141-3146.

Morais, S., Barak, Y., Caspi, J., Hadar, Y., Lamed, R., Shoham, Y., et al. (2010a). Cellulase-xylanase synergy in designer cellulosomes for enhanced degradation of a complex cellulosic substrate. *MBio, 1*(5).

Morais, S., Barak, Y., Caspi, J., Hadar, Y., Lamed, R., Shoham, Y., et al. (2010b). Contribution of a xylan-binding module to the degradation of a complex cellulosic substrate by designer cellulosomes. *Appl Environ Microbiol, 76*(12), 3787-3796.

Murashima, K., Chen, C. L., Kosugi, A., Tamaru, Y., Doi, R. H., & Wong, S. L. (2002). Heterologous production of *Clostridium cellulovorans* engB, using protease-deficient *Bacillus subtilis*, and preparation of active recombinant cellulosomes. *J Bacteriol, 184*(1), 76-81.

Murashima, K., Kosugi, A., & Doi, R. H. (2002). Synergistic effects on crystalline cellulose degradation between cellulosomal cellulases from *Clostridium cellulovorans*. *J Bacteriol, 184*(18), 5088-5095.

Murashima, K., Kosugi, A., & Doi, R. H. (2003). Solubilization of cellulosomal cellulases by fusion with cellulose-binding domain of noncellulosomal cellulase engd from *Clostridium cellulovorans*. *Proteins, 50*(4), 620-628.

Narita, J., Okano, K., Kitao, T., Ishida, S., Sewaki, T., Sung, M. H., et al. (2006). Display of alpha-amylase on the surface of *Lactobacillus casei* cells by use of the PgsA anchor protein, and production of lactic acid from starch. *Appl Environ Microbiol, 72*(1), 269-275.

Neubauer, H., Bauche, A., & Mollet, B. (2003). Molecular characterization and expression analysis of the dextransucrase DsrD of *Leuconostoc mesenteroides* Lcc4 in homologous and heterologous *Lactococcus lactis* cultures. *Microbiology, 149*(Pt 4), 973-982.

Ng, T. K., Weimer, T. K., & Zeikus, J. G. (1977). Cellulolytic and physiological properties of *Clostridium thermocellum*. *Arch Microbiol, 114*(1), 1-7.

Nolling, J., Breton, G., Omelchenko, M. V., Makarova, K. S., Zeng, Q., Gibson, R., et al. (2001). Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J Bacteriol, 183*(16), 4823-4838.

Nouaille, S., Bermudez-Humaran, L. G., Adel-Patient, K., Commissaire, J., Gruss, A., Wal, J. M., et al. (2005). Improvement of bovine beta-lactoglobulin production and secretion by *Lactococcus lactis*. *Braz J Med Biol Res, 38*(3), 353-359.

Perret, S., Casalot, L., Fierobe, H. P., Tardif, C., Sabathe, F., Belaich, J. P., et al. (2004). Production of heterologous and chimeric scaffoldins by *Clostridium acetobutylicum* ATCC 824. *J Bacteriol, 186*(1), 253-257.

Petrov, K., Urshev, Z., & Petrova, P. (2008). L+-lactic acid production from starch by a novel amylolytic *Lactococcus lactis* subsp. lactis B84. *Food Microbiol, 25*(4), 550-557.

Piard, J. C., Hautefort, I., Fischetti, V. A., Ehrlich, S. D., Fons, M., & Gruss, A. (1997). Cell wall anchoring of the *Streptococcus pyogenes* M6 protein in various lactic acid bacteria. *J Bacteriol, 179*(9), 3068-3072.

Poquet, I., Ehrlich, S. D., & Gruss, A. (1998). An export-specific reporter designed for gram-positive bacteria: application to *Lactococcus lactis*. *J Bacteriol, 180*(7), 1904-1912.

Poquet, I., Saint, V., Seznec, E., Simoes, N., Bolotin, A., & Gruss, A. (2000). HtrA is the unique surface housekeeping protease in *Lactococcus lactis* and is required for natural protein processing. *Mol Microbiol, 35*(5), 1042-1051.

Ragauskas, A. J., Williams, C. K., Davison, B. H., Britovsek, G., Cairney, J., Eckert, C. A., et al. (2006). The path forward for biofuels and biomaterials. *Science, 311*(5760), 484-489.

Raha, A. R., Varma, N. R., Yusoff, K., Ross, E., & Foo, H. L. (2005). Cell surface display system for *Lactococcus lactis*: a novel development for oral vaccine. *Appl Microbiol Biotechnol, 68*(1), 75-81.

Raman, B., Pan, C., Hurst, G. B., Rodriguez, M., Jr., McKeown, C. K., Lankford, P. K., et al. (2009). Impact of pretreated Switchgrass and biomass carbohydrates on *Clostridium thermocellum* ATCC 27405 cellulosome composition: a quantitative proteomic analysis. *PLoS One, 4*(4), e5271.

Ramasamy, R., Yasawardena, S., Zomer, A., Venema, G., Kok, J., & Leenhouts, K. (2006). Immunogenicity of a malaria parasite antigen displayed by *Lactococcus lactis* in oral immunisations. *Vaccine, 24*(18), 3900-3908.

Ribeiro, L. A., Azevedo, V., Le Loir, Y., Oliveira, S. C., Dieye, Y., Piard, J. C., et al. (2002). Production and targeting of the *Brucella abortus* antigen L7/L12 in *Lactococcus lactis*: a first step towards food-grade live vaccines against brucellosis. *Appl Environ Microbiol, 68*(2), 910-916.

Rincon, M. T., Cepeljnik, T., Martin, J. C., Lamed, R., Barak, Y., Bayer, E. A., et al. (2005). Unconventional mode of attachment of the *Ruminococcus flavefaciens* cellulosome to the cell surface. *J Bacteriol, 187*(22), 7569-7578.

Rincon, M. T., Ding, S. Y., McCrae, S. I., Martin, J. C., Aurilia, V., Lamed, R., et al. (2003). Novel organization and divergent dockerin specificities in the cellulosome system of *Ruminococcus flavefaciens*. *J Bacteriol, 185*(3), 703-713.

Rittmann, D., Lindner, S. N., & Wendisch, V. F. (2008). Engineering of a glycerol utilization pathway for amino acid production by *Corynebacterium glutamicum*. *Appl Environ Microbiol, 74*(20), 6216-6222.

Rogers, P. L., Jeon, Y. J., Lee, K. J., & Lawford, H. G. (2007). *Zymomonas mobilis* for fuel ethanol and higher value products. *Adv Biochem Eng Biotechnol, 108*, 263-288.

Sabathe, F., Belaich, A., & Soucaille, P. (2002). Characterization of the cellulolytic complex (cellulosome) of *Clostridium acetobutylicum*. *FEMS Microbiol Lett, 217*(1), 15-22.

Sabathe, F., & Soucaille, P. (2003). Characterization of the CipA scaffolding protein and *in vivo* production of a minicellulosome in *Clostridium acetobutylicum*. *J Bacteriol, 185*(3), 1092-1096.

Saier, M. H., Jr. (2006). Protein secretion and membrane insertion systems in gram-negative bacteria. *J Membr Biol, 214*(2), 75-90.

Sambrook, J., & Russell, D. W. (2001). *Molecular cloning : a laboratory manual* (3rd ed.). Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Sampson, E. M., & Bobik, T. A. (2008). Microcompartments for B12-dependent 1,2-propanediol degradation provide protection from DNA and cellular damage by a reactive metabolic intermediate. *J Bacteriol, 190*(8), 2966-2971.

Schwarz, W. H. (2001). The cellulosome and cellulose degradation by anaerobic bacteria. *Appl Microbiol Biotechnol, 56*(5-6), 634-649.

Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E., & Tsien, R. Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma sp*. red fluorescent protein. *Nat Biotechnol, 22*(12), 1567-1572.

Shanmugam, K. T., & Ingram, L. O. (2008). Engineering biocatalysts for production of commodity chemicals. *J Mol Microbiol Biotechnol, 15*(1), 8-15.

Shaw, A. J., Podkaminer, K. K., Desai, S. G., Bardsley, J. S., Rogers, S. R., Thorne, P. G., et al. (2008). Metabolic engineering of a thermophilic bacterium to produce ethanol at high yield. *Proc Natl Acad Sci U S A, 105*(37), 13769-13774.

Simonen, M., & Palva, I. (1993). Protein secretion in *Bacillus* species. *Microbiol Rev, 57*(1), 109-137.

Sorvig, E., Gronqvist, S., Naterstad, K., Mathiesen, G., Eijsink, V. G., & Axelsson, L. (2003). Construction of vectors for inducible gene expression in *Lactobacillus sakei* and *L plantarum*. *FEMS Microbiol Lett, 229*(1), 119-126.

Sriraman, K., & Jayaraman, G. (2008). HtrA is essential for efficient secretion of recombinant proteins by *Lactococcus lactis*. *Appl Environ Microbiol, 74*(23), 7442-7446.

Steen, E. J., Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., McClure, A., et al. (2010). Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature, 463*(7280), 559-562.

Steidler, L., Viaene, J., Fiers, W., & Remaut, E. (1998). Functional display of a heterologous protein on the surface of *Lactococcus lactis* by means of the cell wall anchor of *Staphylococcus aureus* protein A. *Appl Environ Microbiol, 64*(1), 342-345.

Terzaghi, B. E., & Sandine, W. E. (1975). Improved medium for lactic streptococci and their bacteriophages. *Appl Microbiol, 29*(6), 807-813.

Thomson, J. A. (1993). Molecular biology of xylan degradation. *FEMS Microbiol Rev, 10*(1-2), 65-82.

Tsai, S. L., Goyal, G., & Chen, W. (2010). Surface display of a functional minicellulosome by intracellular complementation using a synthetic yeast consortium and its application to cellulose hydrolysis and ethanol production. *Appl Environ Microbiol, 76*(22), 7514-7520.

Tsai, S. L., Oh, J., Singh, S., Chen, R., & Chen, W. (2009). Functional assembly of minicellulosomes on the *Saccharomyces cerevisiae* cell surface for cellulose hydrolysis and ethanol production. *Appl Environ Microbiol, 75*(19), 6087-6093.

Tyurin, M. V., Desai, S. G., & Lynd, L. R. (2004). Electrotransformation of *Clostridium thermocellum*. *Appl Environ Microbiol, 70*(2), 883-890.

Vallejo, L. F., & Rinas, U. (2004). Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microb Cell Fact, 3*(1), 11.

van Asseldonk, M., Rutten, G., Oteman, M., Siezen, R. J., de Vos, W. M., & Simons, G. (1990). Cloning of usp45, a gene encoding a secreted protein from *Lactococcus lactis* subsp. lactis MG1363. *Gene, 95*(1), 155-160.

Wang, W. K., & Wu, J. H. (1993). Structural features of the *Clostridium thermocellum* cellulase SS gene. *Appl Biochem Biotechnol, 39-40*, 149-158.

Watanabe, H., & Tokuda, G. (2010). Cellulolytic systems in insects. *Annu Rev Entomol, 55*, 609-632.

Wen, F., Sun, J., & Zhao, H. (2010). Yeast surface display of trifunctional minicellulosomes for simultaneous saccharification and fermentation of cellulose to ethanol. *Appl Environ Microbiol, 76*(4), 1251-1260.

Wieczorek, A. S., & Martin, V. J. (2010). Engineering the cell surface display of cohesins for assembly of cellulosome-inspired enzyme complexes on *Lactococcus lactis*. *Microb Cell Fact, 9*, 69.

Wong, S. L. (1995). Advances in the use of *Bacillus subtilis* for the expression and secretion of heterologous proteins. *Curr Opin Biotechnol, 6*(5), 517-522.

Wu, C. H., Mulchandani, A., & Chen, W. (2008). Versatile microbial surface-display for environmental remediation and biofuels production. *Trends Microbiol, 16*(4), 181-188.

Xu, J., & Smith, J. C. (2010). Probing the mechanism of cellulosome attachment to the *Clostridium thermocellum* cell surface: computer simulation of the Type II cohesin-dockerin complex and its variants. *Protein Eng Des Sel, 23*(10), 759-768.

Xu, Q., Gao, W., Ding, S. Y., Kenig, R., Shoham, Y., Bayer, E. A., et al. (2003). The cellulosome system of *Acetivibrio cellulolyticus* includes a novel type of adaptor protein and a cell surface anchoring protein. *J Bacteriol, 185*(15), 4548-4557.

Yang, Z., Liu, Q., Wang, Q., & Zhang, Y. (2008). Novel bacterial surface display systems based on outer membrane anchoring elements from the marine bacterium *Vibrio anguillarum*. *Appl Environ Microbiol, 74*(14), 4359-4365.

Yanisch-Perron, C., Vieira, J., & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene, 33*(1), 103-119.

You, C., Zhang, X. Z., Sathitsuksanoh, N., Lynd, L. R., & Zhang, Y. H. (2012). Enhanced microbial utilization of recalcitrant cellulose by an *ex vivo* cellulosome-microbe complex. *Appl Environ Microbiol, 78*(5), 1437-1444.

Zhang, M., Eddy, C., Deanda, K., Finkelstein, M., & Picataggio, S. (1995). Metabolic Engineering of a Pentose Metabolism Pathway in Ethanologenic *Zymomonas mobilis*. *Science, 267*(5195), 240-243.

Zhang, Y. H., & Lynd, L. R. (2004). Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. *Biotechnol Bioeng, 88*(7), 797-824.

Zhang, Y. H., & Lynd, L. R. (2005). Regulation of cellulase synthesis in batch and continuous cultures of *Clostridium thermocellum*. *J Bacteriol, 187*(1), 99-106.

Zverlov, V. V., Kellermann, J., & Schwarz, W. H. (2005). Functional subgenomics of *Clostridium thermocellum* cellulosomal genes: identification of the major catalytic components in the extracellular complex and detection of three new enzymes. *Proteomics, 5*(14), 3646-3653.

Zverlov, V. V., Klupp, M., Krauss, J., & Schwarz, W. H. (2008). Mutations in the scaffoldin gene, cipA, of *Clostridium thermocellum* with impaired cellulosome formation and cellulose hydrolysis: insertions of a new transposable element, IS1447, and implications for cellulase synergism on crystalline cellulose. *J Bacteriol, 190*(12), 4321-4327.

Appendix A

In vitro cohesin-dockerin binding assay

I used a scaffold-enzyme binding assay to verify that the dock1 motif of UidA-dock1 retained its ability to interact with the type 1 cohesin motif of the recombinant NucA-coh1 scaffold. Cohesin-dockerin binding experiments were performed on Ni-NTA sepharose columns to form scaffold-enzyme complexes, and elution fractions were tested for both nuclease and β-glucuronidase activity. Whole-cell lysates of *E. coli* BL21(DE3) expressing UidA-dock1 or UidA were loaded onto Ni-NTA sepharose columns which were subsequently washed with 50 mM phosphate buffer, pH 6.0, containing 300 mM NaCl and 20 mM imidazole. Cells from 250 mL of nisin-induced cultures of *L. lactis* pAW310 (NucA-coh1-cwa) and *L. lactis* pAW301 (NucA-cwa) were suspended in 50 mM phosphate buffer, pH 7.5, containing 300 mM NaCl, to which lysozyme was added at a final concentration of 10 mg/mL. After digestion for 1 hr at 37°C to degrade the cell wall and release associated scaffold proteins, samples were sonicated and lysates were loaded onto columns with either pre-bound UidA-dock1 or UidA. Columns were further washed with phosphate buffer (50 mM, pH 6.0) containing 300 mM NaCl and 20 mM imidazole, and samples were eluted using the same buffer containing 250 mM imidazole. The theoretical expected results are portrayed in Figure 20. Only combinations of proteins containing both functional type 1 cohesin and dockerin domains should result in binary scaffold-enzyme complexes. TBD-agar plates as well as agar plates containing 5-bromo-4-chloro-3-indolyl β-D-glucopyranoside (X-gluc) (Fluka) at a concentration of 0.5 mM, were spotted with protein elution fractions for detection of nuclease (NucA and

NucA-coh1) and β-glucuronidase (UidA and UidA-dock1), respectively (Fig. 21). Plates were incubated for 1 hr at 37°C and analyzed for pink (NucA activity) or blue (UidA activity) color formation.

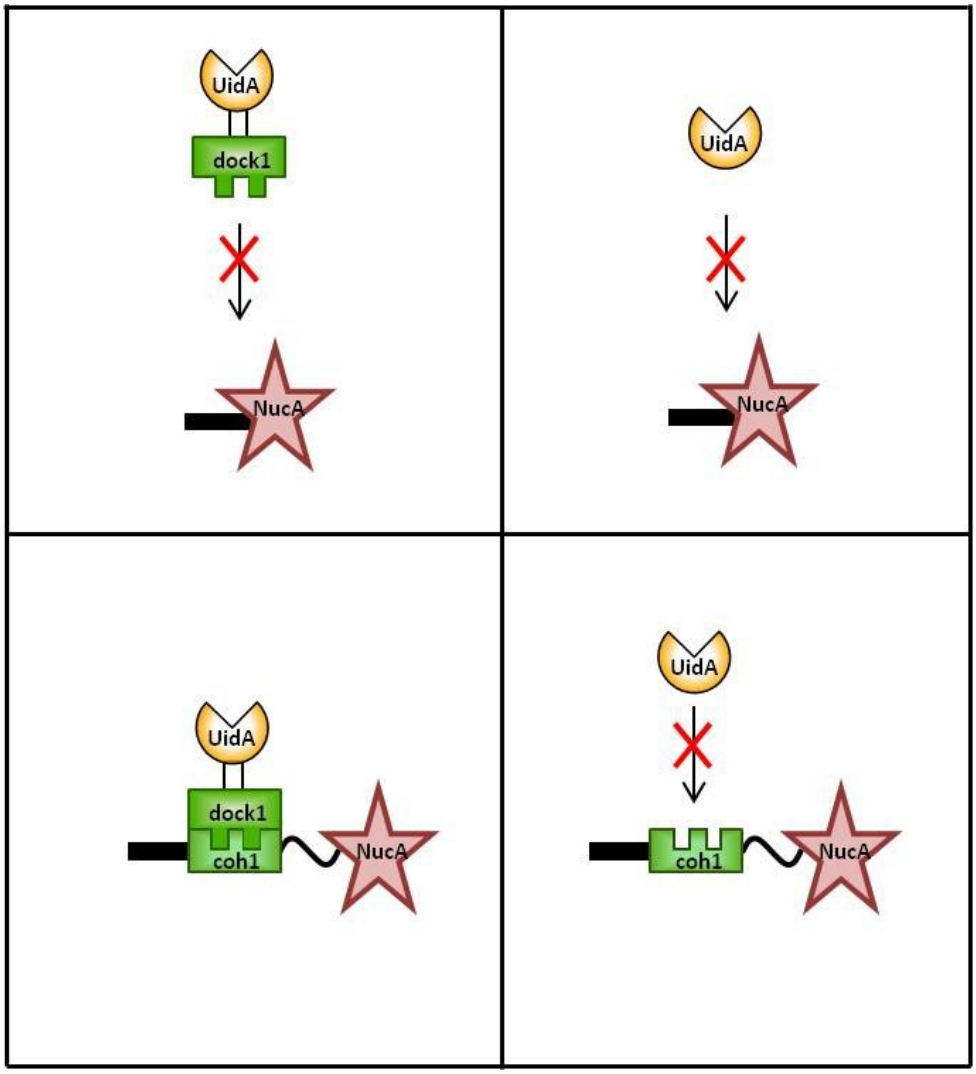

Figure 20. Graphic depiction of outcomes of the *in vitro* cohesin-dockerin binding assay. Assembly of the NucA-coh1 / UidA-dock1 complex is dependent on the interactions between the type 1 cohesin and dockerin partners. Elution fractions containing assembled complexes should therefore demonstrate both nuclease and β-glucuronidase activity.

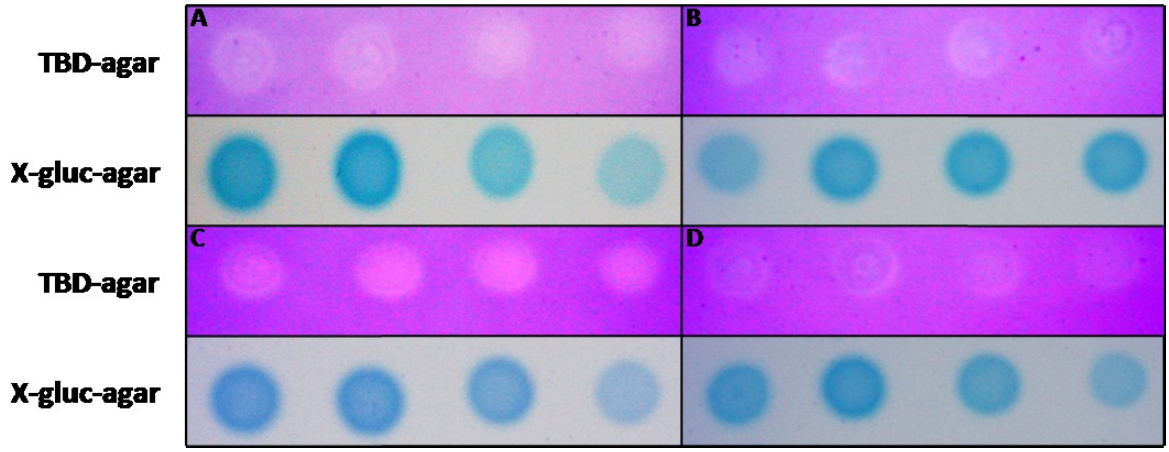

Figure 21. *In vitro* binding assay results. Elution fractions were spotted on both TBD-agar for nuclease detection and agar containing 5-bromo-4-chloro-3-indolyl β-D-glucopyranoside (X-gluc) for β-glucuronidase detection. Binding assays combined either **(A)** NucA-coh1 and UidA, **(B)** NucA and UidA, **(C)** NucA-coh1 and UidA-dock1, **(D)** NucA and UidA-dock1. Co-elution of components indicative of successful complex formation is characterized by dual enzymatic activity (C).

As illustrated in Figure 21, only elution fractions containing the assembled NucA-coh1-cwa / UidA-dock1 complex exhibited both nuclease and β-glucuronidase activity (Fig. 21 C). Protein fractions eluted from columns loaded with UidA showed only β-glucuronidase activity and no nuclease activity due to the lack of a dock1 domain on UidA (Fig. 21 A, B). Similarly, elution fractions combining NucA-cwa and UidA-dock1 also showed only β-glucuronidase activity suggesting lack of association between the two proteins due to the absence of coh1 (Fig. 21 D). Detection of nuclease activity in the wash steps of all control experiments confirmed that nuclease was present in the *L. lactis* extract, and that complex formation was dependant on the presence of both a functional dockerin and cohesin domain (data not shown).

Appendix B. Engineering fluorescent probes to target type 1 and type 2 cohesins on chimeric scaffolds

Non-enzyme proteins may be useful as probes for verifying the number of cohesins present on a surface-displayed scaffold. The variability in excitation and emission profiles of fluorescent proteins makes them ideal candidates for this task. Therefore, I generated a library of fluorescent probes as dockerin fusions (Fig. 22). A total of four fluorescent proteins, GFPuv, mCherry, mBanana, and mOrange (Shaner, et al., 2004), were fused with either a dock1 or dock2 domain. To achieve this, the genes encoding each protein without the stop codon were PCR amplified and inserted into the NheI/NotI restriction sites of either pET28(a), pETdock1, or pETdock2.

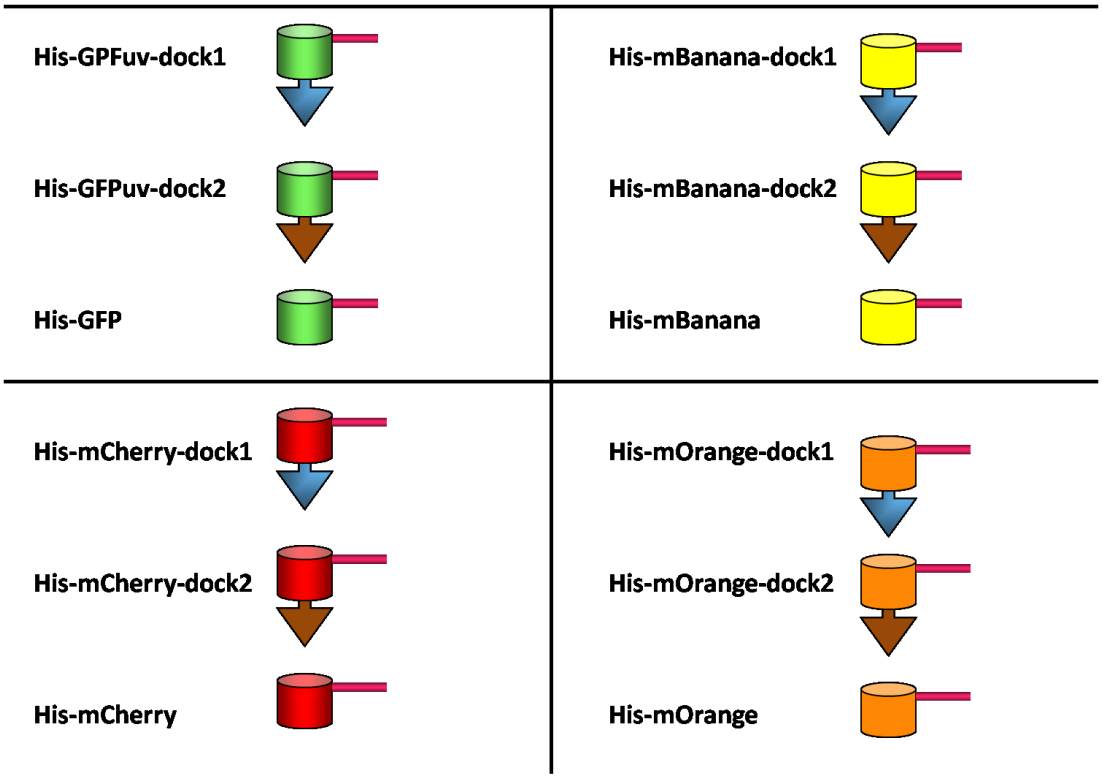

Figure 22. Library of Fluorescent probe-dockerin fusions. Fluorescent probes contain N-terminal 6xHis tags, as well as C-terminal dock1 or dock2 domains. Controls lack appropriate dockerin domains. Cylinders represent fluorescent proteins, large arrows represent C-terminal dockerins and pink bars represent N-terminal the 6XHis tag.

All vectors were established in *E. coli* BL21(DE3) by electroporation. Cultures were grown to an OD_{600} of approximately 0.3 (see section 2), after which they were cooled to 18 °C, induced with 1mM IPTG, and grown overnight. Using two fluorescent proteins to simultaneously target scaffold-displaying cells would require them to have non-overlapping excitation/emission profiles in order to independently observe each. Based on the excitation/emission profiles of the four proteins, I decided to test for signal overlap using GFPuv and mCherry, which demonstrate the least amount of signal overlap (Fig. 23).

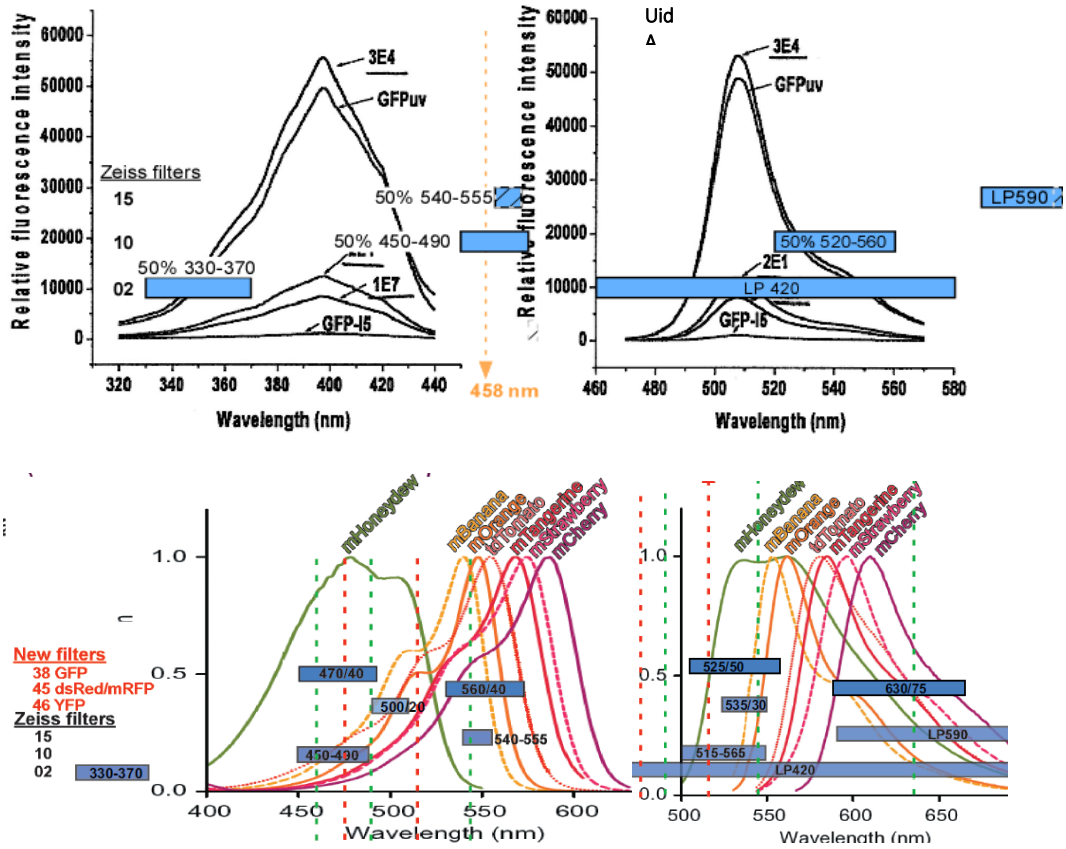

Figure 23. Emission and excitation spectra of fluorescent proteins and corresponding Zeiss Fluorescence Microscope filters. GFPuv is best detected using filter *02* since the indicated excitation (A) and emission (B) spectra offer detection of approximately 45% of the signal maximum. mCherry is best detected using filter *45* since the indicated excitation (C) and emission (D) spectra offer detection of approximately 85% of the signal maximum. (Courtesy Dr. Marc Champagne).

Cells of *E. coli* over-expressing GFPuv-dock1 and mCherry-dock1 were visualized using a *Zeiss* Fluorescence Microscope and corresponding filters (Fig. 24). These results suggest that the simultaneous targeting of both GFPuv and mCherry to chimeric scaffolds would result in their ability to be independently visualized using appropriate filters. This also suggests that quantification of proteins bound to cells could be achieved using alternate technologies such as a phosphoimager, knowing the calculated fluorescence intensity of a precise amount of protein.

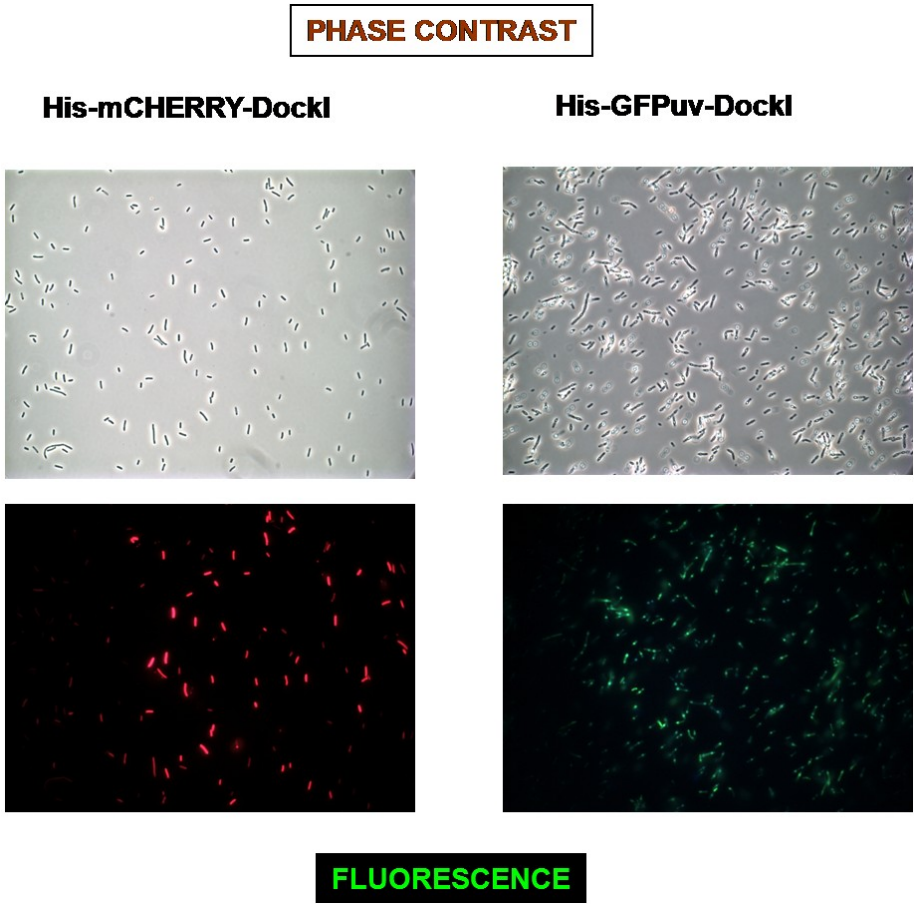

Figure 24. Detection of mCherrydock1 and GFPuvdock1 using fluorescence microscopy. *E. coli* BL21(DE3) cells over-expressing mCherrydock1 (left) and GFPuv (right) were visualized using a *Zeiss* Fluorescence Microscope. Filter *45* was used for mCherry detection, while filter *02* was used for GFPuv detection. When filters were switched, no signals were detected (data not shown).

Protein-dockerin fusions were purified using affinity chromatography as described in section 2. Visualization of each respective colour of fluorescent protein allowed simple and efficient detection of elution fractions containing the highest amount of over-expressed protein. SDS-PAGE was used to determine the homonogeneity of the purified probes. Interestingly, visualization of purified GFPuv-dock1 using SDS-PAGE revealed a band corresponding to the correct fusion, as well other bands (Fig. 25). I hypothesize that this was due to premature translational termination, since GFPuv was purified to homogeneity, suggesting that the dockerin domain was either inefficiently translated or subject to proteolytic degradation.

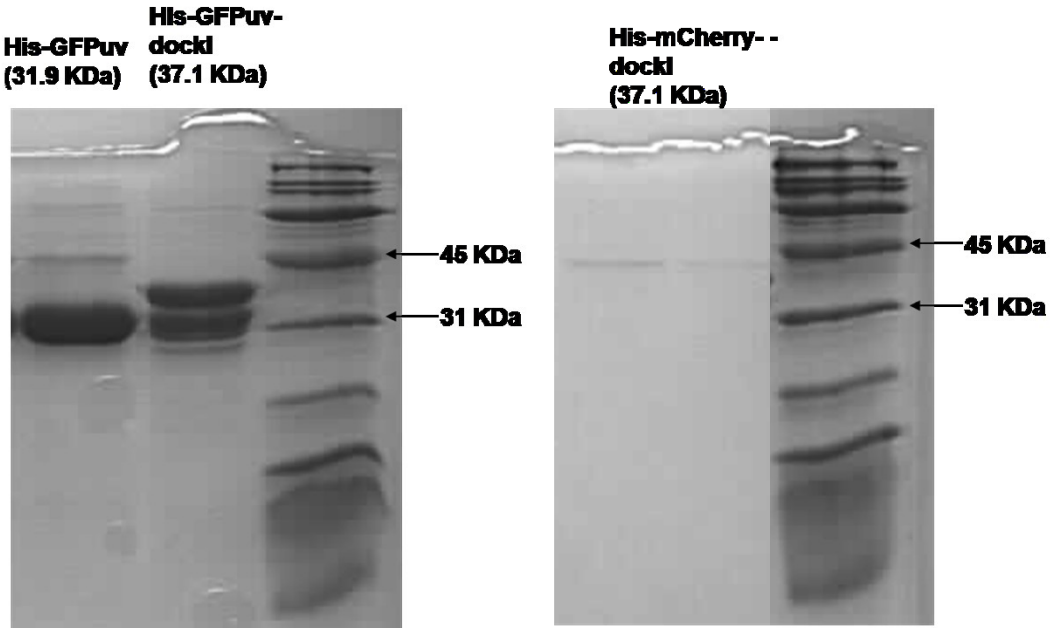

Figure 25. SDS-PAGE of elution fractions containing fluorescent probes. Elution fractions from the purification of His-tagged fluorescent probes were loaded onto a 12% acrylamide SDS-PAGE to visualize the homogeneity of the probes. Fractions containing His-GFPuv-dock1 contained two distinct bands, one of which probably corresponds to products subject to pre-mature translation termination, or C-terminal protease degradation.

Using the protocol for the *in vivo* binding assay described in section 2, I was successful in visualizing *L. lactis* cells displaying NucA-coh1 after incubation with GFPuv-dock1 (Fig. 26) and mCherry-dock1 (Fig. 27). This suggested that sufficient amounts of probe were able to associate with displayed NucA-coh1 for visualization, albeit with extended exposure times. Cells incubated with GFPuv-dock1 were visualized, however only a fraction of the total cells demonstrated the appropriate flurescence.

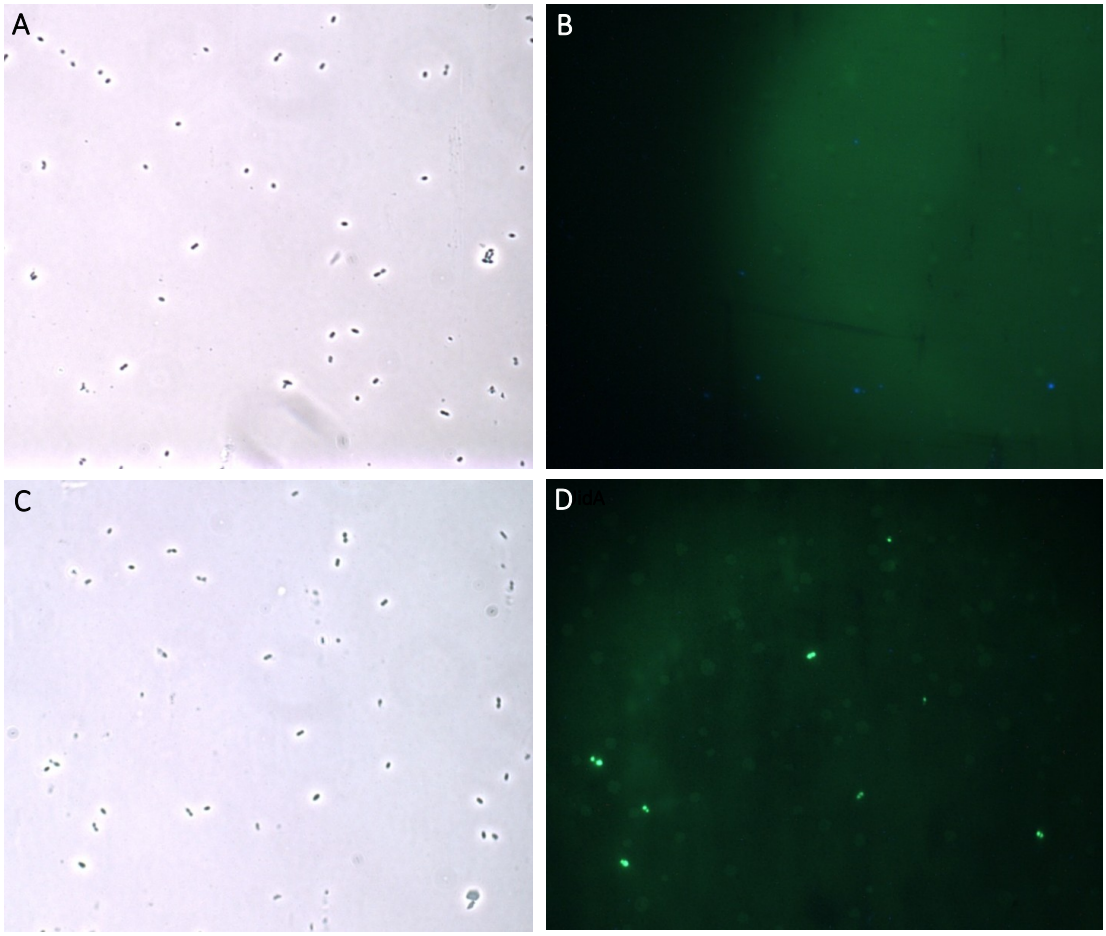

Figure 26. Docking GFPuv-dock1 on *L. lactis* cells displaying coh1-NucA. GFPuv (A, B) or GFPuv-dock1 (C, D) was targeted to cells displaying coh1-NucA and visualized using a Zeiss Fluorescence Microscope (1000X). Panels A and C demonstrate images using phase contrast, while B and D use fluorescence microscopy (filter *02*).

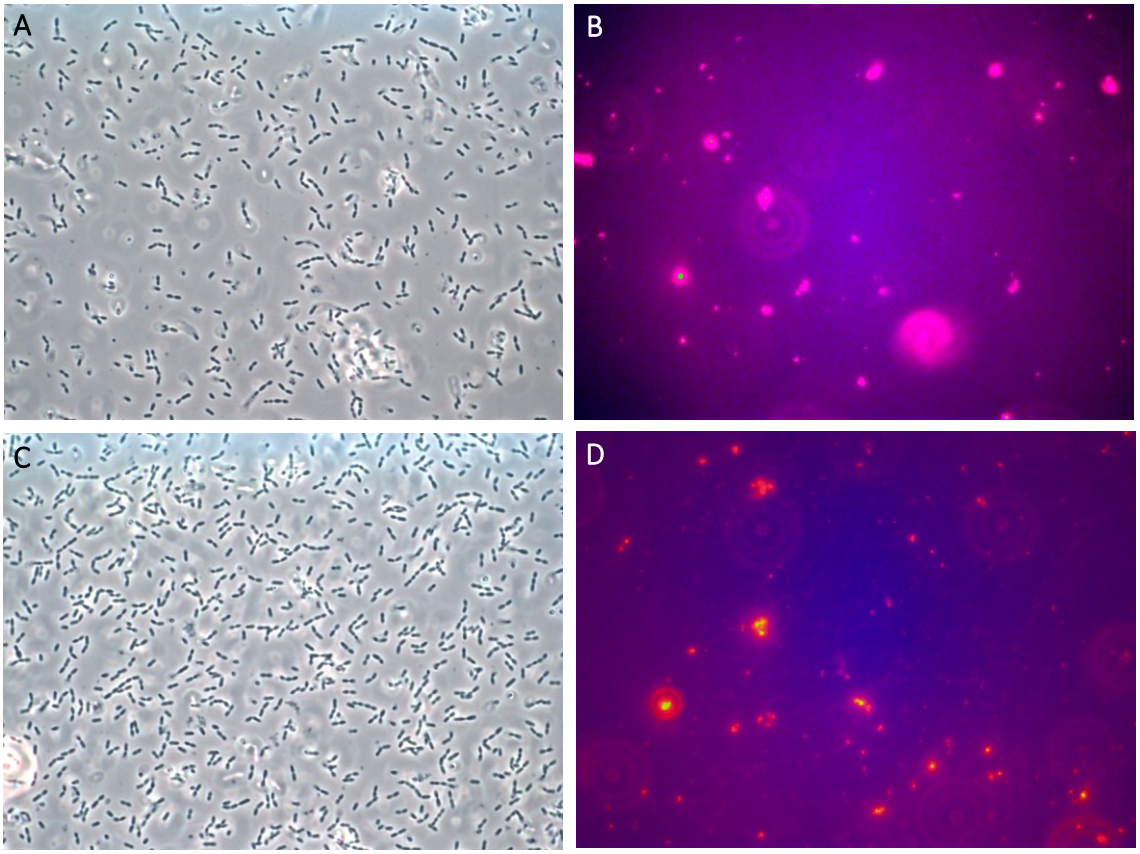

Figure 27. Docking mCherry-dock1 on *L. lactis* cells displaying coh1-NucA. Mcherry (A, B) or mCherry-dock1 (C, D) was targeted to cells displaying coh1-NucA and visualized using a Zeiss Fluorescence Microscope (1000X). Panels A and C demonstrate images using phase contrast, while B and D use fluorescence microscopy (filter *45*).

I also decided to test the possibility of docking GFPuv-dock1 and mCherry-dock1 on *C. Thermocellum* cellulosomes *in vivo*. GFPuv-dock1 was added to a growing culture of *C. thermocellum* in early-log phase. After 4 hrs, cultures were harvested and visualized using fluorescence microscopy (Fig. 28). GFP-dock1 successfully bound CipA on the bacterial cell surface and cells were visualized while actively growing on cellulose fibers (Fig28).

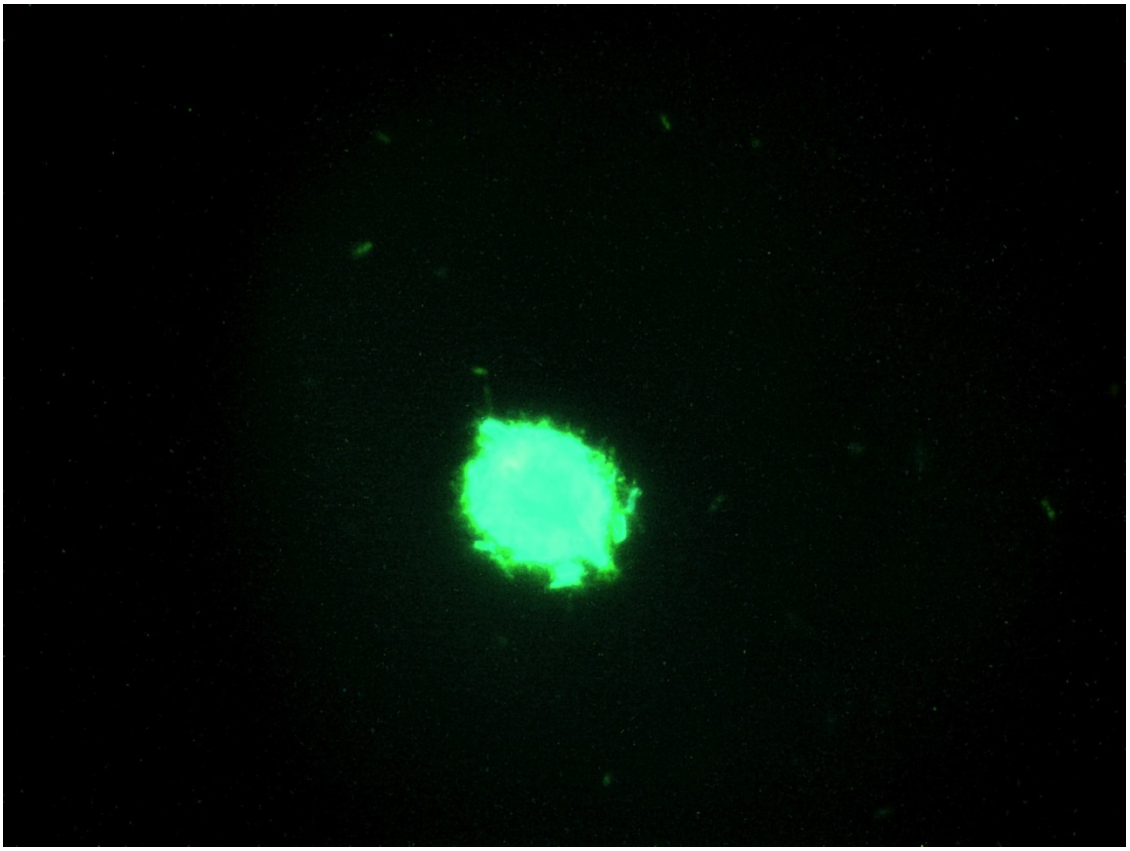

Figure 28. *In vivo* targeting GFPuv-dock1 to surface-displayed cellulosomes on *C. thermocellum*

GFPuv-dock1 and mCherry-dock1 were also targeted to the surface of *C. thermocellum* cells in similar fashion to binding assays described in section 2. Cells in early-log or mid-log phase were harvested, incubated with GFPuv-dock1 or mCherry-dock1 for 4 hrs, washed six times, and visualized using a Zeiss Fluorescence Microscope (Fig. 29).

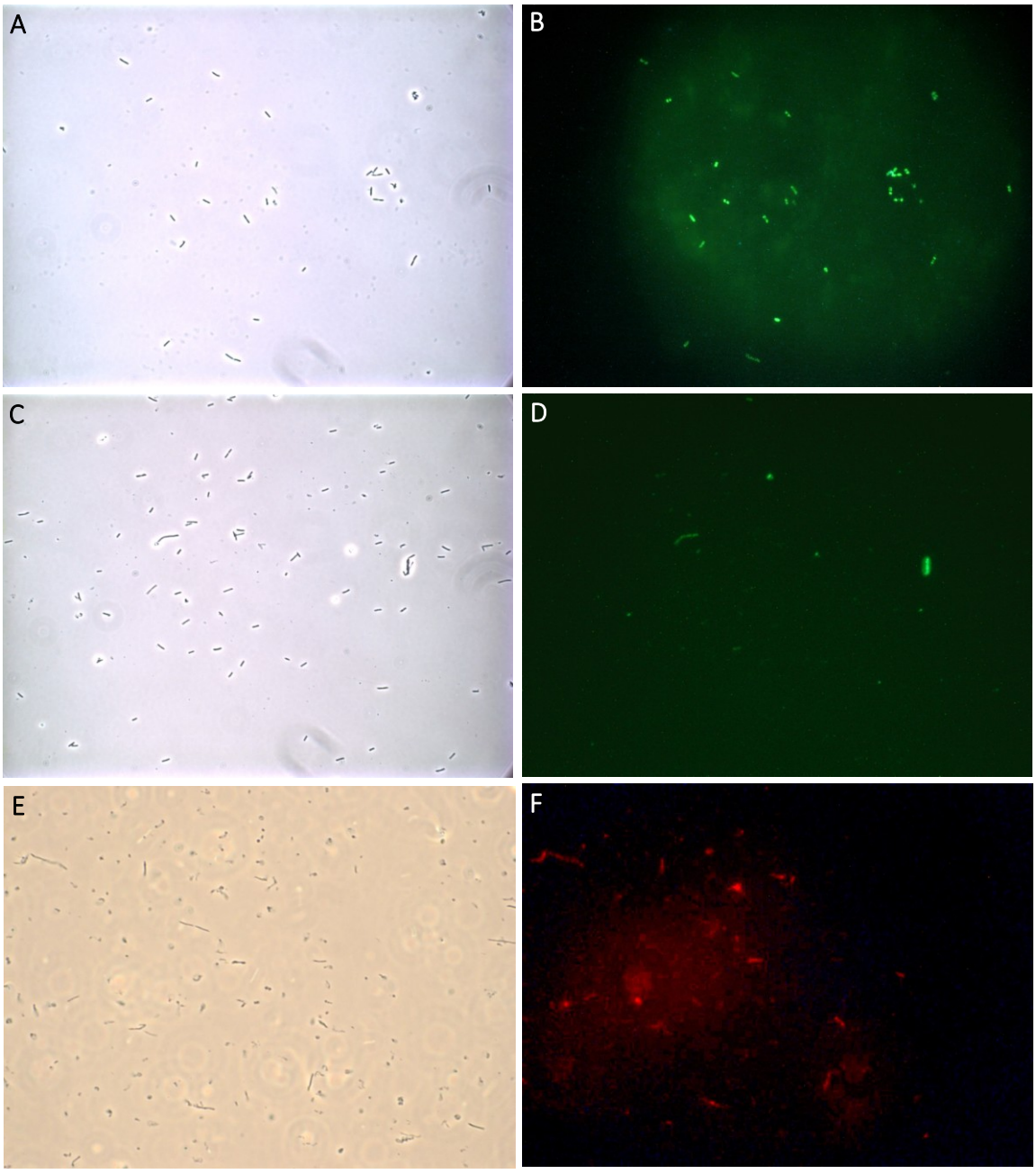

Figure 29. Targeting GFPuv-dock1 and mCherry-dock1 to harvested cells of *C. thermocellum*. Cells from actively growing cultures of C. thermocellum grown on cellobiose were harvested and incubated with GFPuv-dock1 (A, B, C, D) or mCherry-dock1 (E, F). Panels A and B demonstrate results of binding assay performed on cells harvested in early-log phase, while C, D, E, and F demonstrate results performed on cells harvested in mid-log phase. Panels A, C, and E involve microscopy conducted using phase-contrast, while panels B, D, and F involve fluorescence microscopy using appropriate filters (magnification 1000X).