

**DESIGNING MODULAR SYNTHETIC METABOLONS VIA dCAS9-GUIDED
ASSEMBLY AND PROTEIN NANOCAGES**

by

Emily Berckman

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry and Biochemistry

Spring 2020

© 2020 Emily Berckman
All Rights Reserved

**DESIGNING MODULAR SYNTHETIC METABOLONS VIA dCAS9-GUIDED
ASSEMBLY AND PROTEIN NANOCAGES**

by

Emily Berckman

Approved: _____
Brian J. Bahnson, Ph.D.
Chair of the Department of Chemistry and Biochemistry

Approved: _____
John A. Pelesko, Ph.D.
Dean of the College of Arts and Sciences

Approved: _____
Douglas J. Doren, Ph.D.
Interim Vice Provost for Graduate and Professional Education and
Dean of the Graduate College

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

Wilfred Chen, Ph.D.
Co-Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

Brian J. Bahnson, Ph.D.
Co-Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

John Koh, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

Colin Thorpe, Ph.D.
Member of dissertation committee

ACKNOWLEDGMENTS

As I look back on the years I have spent pursuing my dream of obtaining a Ph.D. in Chemistry and Biochemistry, I have made so many memories that I will cherish forever. Throughout my time in both the Chemistry and Biochemistry Department as well as the Chemical Engineering Department at University of Delaware, I have been blessed to have met so many wonderful and inspiring people.

First, I would like to acknowledge my advisor, Dr. Wilfred Chen. Often in graduate school, especially around recruiting, we as students are asked if we are happy with our advisor selection and I can honestly say that choosing the Chen lab was the single best decision in my academic career. Dr. Chen as a mentor is second to none, he is determined and demanding but his mentoring methods encourages his students to become the best, most well-rounded and creative biochemical engineers. One of the most inspiring characteristics Dr. Chen possesses is his overwhelming love for science. This love constantly ignites passionate, intellectual conversations often centered on hastily drawn cartoons on whatever paper is floating around. This climate of consistently innovative science and the level of faith Dr. Chen has in me to investigate my own creative solutions has benefited me and my academic career more than I could have imagined. Dr. Chen along with my thesis committee members, Dr. Brian Bahnson, Dr. John Koh and Dr. Colin Thorpe, were instrumental to the success of this thesis and my time here at the University of Delaware.

Next, I would like to acknowledge my past and present fellow group members. Dr. Qing Sun, Dr. Qi Chen, Dr. Maryam Raeeszadeh Sarmazdeh and Dr. Vincent

Price were excellent mentors in my early years. Dr. Rebecca Chen, Dr. Ka-Hei Siu and Dr. Heejae Kim were instrumental in generating many ideas and problem-solving solutions through brainstorming sessions (sometimes venting sessions), especially during my first year in lab. Dr. Andrew Swartz and Andrew Gaynor, two of my greatest friends in lab, made my time in graduate school unforgettable. We entered the Wilfred Chen lab together and through many failed experiments, frustrated walks around campus and much earned trivia nights we have grown as friends and scientists and I am truly grateful to have made such incredible lifelong friendships. I would also like to extend my gratitude toward my current lab mates: Emily Hartzell, Rachel Lieser, Alex Mitkas, Victoria Hunt, Daniel Yur, Hopen Yang, Antonio Goncalves and Anxhela Sinani. I hope that I have been a positive role model in your lives as you all have been in mine. There were very few days I was not excited to come into lab (even on weekends) to spend time with such an amazing group of young scientists.

Next, I would like to thank the many friends I have made here in Delaware. Not only have these wonderful people made Delaware truly feel like home, they offered much needed love and support through my Ph.D. journey. In particular, I would like to acknowledge Sarah Schulte, Jessica Szymanski and all of the incredible athletes at Ignite Fitness. Crossfit has been more than just a gym to me, it is a community of supportive people that push you to be the best version of yourself you can be. Regardless of how my day went in the lab, the moment I walked into the gym my day turned around.

Lastly, I would like to acknowledge my family and my soon to be in-laws. My family has always supported my dreams of obtaining a Ph.D. and entering the world of pharmaceuticals. From a young age my parents have instilled the value of a good

education and pushing myself to be the best person I can be. My brother, an incredible role model, was a driving force in choosing hard sciences as a career path and pushing me to get a doctorate, even if it was just so he has to refer to his younger sister as Doctor. My soon to be in-laws have been incredibly generous and supportive during my time in graduate school by always inquiring about my research and experiments, I am truly grateful to be entering into such a loving family. Most importantly, I want to thank Thomas Dewar, whom I will always love for supporting me and encouraging me to follow my dream the last 6 years. Not a day went by that Thomas did not show the immense love he has for me, whether it was listening to my crazy Chen lab stories and experiments, planning a bright future together or my personal favorite, making me dinner most nights, I could not have done this without him. I look forward to many more incredible years together.

TABLE OF CONTENTS

LIST OF TABLES	xi
LIST OF FIGURES	xii
ABSTRACT	xiv

Chapter

1	INTRODUCTION	1
1.1	Synthetic Metabolon Formation Inspired by Nature	3
1.1.1	Synthetic Metabolon via Protein Scaffolds	4
1.1.2	Synthetic Metabolon via Nucleic Acid Scaffolds	6
1.2	CRISPR/Cas Technology	8
1.2.1	Engineering of the CRISPR/Cas Technology	9
1.3	Cellulose Hydrolysis via Cellulosome Formation.....	10
1.3.1	Native Cellulosome Formation.....	11
1.3.2	Synthetic Cellulosome Formation	12
1.4	Orthogonal Post Translational Ligation Strategies.....	13
1.5	Self-Assembling Protein Nanocages	14
1.5.1	E2 Nanocage.....	15
1.5.2	Hepatitis B Virus (HBV) Nanocage	16
2	A DYNAMIC SYNTHETIC METABOLON VIA dCAS9 GUIDED ASSEMBLY.....	17
2.1	Abstract.....	17
2.2	Introduction	17
2.3	Materials and Methods	20
2.3.1	Materials	20
2.3.2	Strains and Cloning	20

2.3.3	Protein Expression.....	21
2.3.4	Protein Purification.....	22
2.3.5	RNA <i>In Vitro</i> Transcription and Purification.....	22
2.3.6	Single and Dual Beacon Assay.....	23
2.3.7	Electrophoretic Mobility Shift Assay.....	25
2.3.8	Triple Beacon Assay.....	25
2.3.9	Reconstitution of Split Nano Luciferase.....	26
2.3.10	Synthetic Cellulosome Assembly and Reducing Sugar Assay....	26
2.3.11	Conditional Beacon Assay.....	26
2.4	Results and Discussion.....	27
2.4.1	Single Binding of dCas9 Fusion Proteins.....	27
2.4.2	Dual Binding of Orthogonal dCas9 Fusion Proteins.....	31
2.4.3	Triple Binding of Orthogonal dCas9 Fusion Proteins.....	35
2.4.4	dCas9 Mediated Split Nano Luciferase Reconstitution.....	36
2.4.5	dCas9 Mediated Cellulosome Formation.....	38
2.4.6	Conditional dCas9 Mediated Cellulosome Formation.....	42
2.5	Conclusions.....	45
3	A MORE MODULAR SYNTHETIC METABOLON VIA dCAS9 GUIDED ASSEMBLY AND A POST TRANSLATIONAL LIGATION STRATEGY.....	46
3.1	Abstract.....	46
3.2	Introduction.....	46
3.3	Materials and Methods.....	48
3.3.1	Materials.....	48
3.3.2	Strains and Cloning.....	49
3.3.3	Protein Expression.....	49
3.3.3.1	Expression of dCas9-Tag Constructs.....	50
3.3.3.2	Purification of dCas9-Tag Constructs.....	50
3.3.3.3	Expression of Cellulosomal Enzyme-Catcher Constructs.....	51
3.3.3.4	Purification of Cellulosomal Enzyme-Catcher Constructs.....	51
3.3.4	RNA <i>In Vitro</i> Transcription and Purification.....	52
3.3.5	Post Translational Ligation Reactions.....	52
3.3.5.1	Tag/Catcher Lysate Reactions.....	52

3.3.5.2	Tag/Catcher Purified Protein Reactions	52
3.3.5.3	Tag/Catcher Reaction Time Course	53
3.3.6	Electrophoretic Mobility Shift Assay	53
3.3.7	Beacon Assay	53
3.3.8	Synthetic Cellulosome Assembly and Reducing Sugar Assay....	54
3.3.8.1	Step-wise Cellulosome Assembly and Reducing Sugar Assay	54
3.3.8.2	One Pot Cellulosome Assembly and Reducing Sugar Assay	55
3.4	Results and Discussion.....	55
3.4.1	Design of dCas9 proteins for orthogonal bioconjugation.....	55
3.4.2	Formation of dCas9-enzyme conjugates by post translational ligation reactions	57
3.4.2.1	Functionality and Orthogonality of the SpyCatcher and SnoopCatcher Reactions.....	58
3.4.2.2	Time Course of Post Translation Ligation Reaction	61
3.4.3	Stepwise Tag/Catcher Reaction for Mini Cellulosome Formation	63
3.4.4	One Pot Tag/Catcher Reaction for Mini Cellulosome Formation	66
3.5	Conclusions	69
4	A MODULAR SYNTHETIC METABOLON VIA SELF-ASSEMBLING PROTEIN NANOCAGES.....	71
4.1	Abstract.....	71
4.2	Introduction	71
4.3	Materials and Methods	73
4.3.1	Materials	73
4.3.2	Strains and Cloning	73
4.3.3	Protein Expression.....	73
4.3.4	Protein Purification.....	74
4.3.4.1	E2 Purification.....	74
4.3.4.2	HBV Purification.....	74

4.3.4.3	Purification of Cellulose Enzyme SpyCatcher Fusions.....	75
4.3.5	Protein Nanocage Decoration.....	75
4.3.6	Synthetic Cellulosome Assembly and Reducing Sugar Assay....	76
4.4	Results and Discussion	76
4.4.1	Artificial Cellulosome Formation via E2 Self-Assembly	76
4.4.2	Artificial Cellulosome Formation via HBV Self-Assembly	81
4.5	Conclusions	87
5	DISSERTAION CONCLUSIONS AND FUTURE WORK	88
5.1	Moving Towards a Fully Dynamic System.....	88
5.2	Expanding to Alternative Metabolic Pathways	91
	REFERNCES	91
Appendix		
A	LIST OF ABBREVIATIONS	105
B	BEACON SEQUENCES.....	107

LIST OF TABLES

Table 2.1 DNA sequences of single, dual and triple targets.	24
--	----

LIST OF FIGURES

Figure 1.1 Benefits of Metabolon Formation.	2
Figure 2.1 Formation of artificial cellulosome structure by dCas9-mediated assembly onto a double-stranded DNA template.	20
Figure 2.2 Functionality of individual dCas9 proteins using the Cas beacon assay.	28
Figure 2.3 EMSA analysis showing qualitative single binding.	30
Figure 2.4 Schematic of dual beacon.	31
Figure 2.5 Orthogonal dCas9 dual binding beacon assay.	33
Figure 2.6 EMSA analysis showing orthogonal and dual binding.	34
Figure 2.7 Optimal spacing for efficient dual binding.	35
Figure 2.8 Triple binding of three orthogonal dCas9 proteins.	36
Figure 2.9 Split Nluc reconstitution.	38
Figure 2.10 Binding efficiencies of SpdCas9-CBD and St1dCas9-CelA onto the dual beacon.	39
Figure 2.11 Assembly of two-component cellulosome using SpdCas9-CBD and ST1dCas9-CelA.	40
Figure 2.12 Comparison of cellulose hydrolysis by different synthetic cellulosomes.	41
Figure 2.13 Conditional assembly of dCas9 protein using a thgRNA.	42
Figure 2.14 Conditional Beacon Assay.	43
Figure 2.15 Conditional assembly of a two-component cellulosome.	44

Figure 3.1 Formation of artificial cellulosome structure by full-length chimeric proteins guided by engineered sgRNA onto a double-stranded DNA template.	48
Figure 3.2 More modular dCas9 proteins improve expression.	57
Figure 3.3 Purified proteins spontaneously form isopeptide bonds.	59
Figure 3.4 ST1dCas9-SnoopTag truncation is not capable of binding DNA	60
Figure 3.5 In Vivo Orthogonal Tag/Catcher Reactions.	61
Figure 3.6 Time Course Tag/Catcher Reactions.	63
Figure 3.7 Binding Efficiency of dCas9-Tag proteins pre and post ligation reactions in a stepwise fashion.	64
Figure 3.8 Step Wise Cellulosome Formation Enhanced Reducing Sugar Titers. 66	
Figure 3.9 SDS-PAGE of One Pot Reaction of both dCas9-Tag and Catcher-Enzyme fusions.	67
Figure 3.10 Binding Efficiency of dCas9-Tag proteins pre and post ligation reactions in one pot reaction.	68
Figure 3.11 One Pot Reaction Cellulosome Formation Enhanced Reducing Sugar Titers.	69
Figure 4.1 SDS-PAGE analysis of controlled decoration of CelA and CBD to the exterior of E2 nanocage.	78
Figure 4.2 TEM analysis of nanocage formation.	79
Figure 4.3 Decorated E2 nanocage increases cellulose hydrolysis.	80
Figure 4.4 SDS-PAGE analysis of controlled decoration of CelA and CBD to the exterior of the HBV nanocage.	83
Figure 4.5 TEM analysis of HBV nanocage formation	84
Figure 4.6 Decorated HBV nanocage increases cellulose hydrolysis.	86
Figure 5.1 Fully dynamic dCas9 scaffolding system.	90

ABSTRACT

In the past few decades, scientists have exploited the natural metabolic processes of microorganisms to achieve the synthesis of many products ranging from pharmaceuticals, and fuels to both bulk and fine chemicals. Metabolic engineering of microorganisms often exceeds the current ability of most organic synthetic chemists, but our existing techniques lack the ability to easily optimize cell productivity and product titer for expansion to a variety of enzymatic pathways. Our goal is to move away from traditional approaches of overexpressing rate limiting enzymes and competitive pathway deletion, towards a strategy in which a cell can sense the accumulation of an intermediate and elicit proximity control over enzyme organization. Throughout this work, several strategies were employed to mediate this enzyme organization, also known as a metabolon.

Nuclease- null Cas9 (dCas9) fusion proteins were combined with engineered single-guide RNA (sgRNA) along with a DNA template for the formation of the synthetic metabolon. The dCas9 protein was chosen as our scaffolding partner due to its high binding affinity as well as its innate ability to be directed to any DNA target with high specificity. Using DNA as the scaffolding domain also offers its own unique advantages like the ease of synthesis, its ability to form secondary and tertiary structures and its flexibility. The ability to utilize three orthogonal dCas9 fusion

proteins to bind to a single target was studied, as was the ability of our dCas9 fusion proteins to form a synthetic metabolon focused on cellulose hydrolysis. Furthermore, we employed an engineered toe-hold guide RNA (thgRNA) for conditional binding and subsequent metabolon formation.

To optimize the modularity of our dCas9 mediated metabolon and reduce expression issues associated with larger engineered proteins, we employed a post translation ligation strategy. We fused a small peptide tag (Tag) to the dCas9 protein and fused its binding partner (Catcher) to the pathway enzymes of interest. This tag/catcher system has been shown to form a spontaneous isopeptide bond in a wide range of temperatures, pHs and buffering conditions. We sought to utilize two orthogonal tag/catcher systems in order to maintain the site specificity offered by the dCas9 proteins.

Utilizing the post translational tag/catcher system not only optimized the dCas9 mediated scaffold but it opened up a pathway to study other scaffolding methods. Specifically using self- assembling protein nanocages for the spatio temporal control of pathway enzymes. We chose to use two protein nanocages, E2 and HBV, due to their differences in size and subunit composition. We decorated the exterior of the protein nanocages with different ratios of enzymes utilized in cellulose hydrolysis and studied the effects on reducing sugar titers.

When it comes to enhancing product titers from a variety of metabolic pathways there isn't a one option fits all. Here we have developed multiple ways to scaffold pathway enzymes, each offering their own advantages. We hypothesize, if

enzyme location is important to a particular pathway, the dCas9 mediated platform is second to none due to the site specificity offered by the Cas9:gRNA complex. Other pathways show increased product titers from the co-localization of enzymes and order of enzyme is less important, here one may wish to employ the nanocage mediated scaffolding system since it has fewer components.

Chapter 1

INTRODUCTION

Engineering cellular metabolism is an ever growing area of research due to its ability to accomplish highly specific and efficient synthesis of a variety of products such as pharmaceuticals, next generation biofuels and commodity chemicals.¹⁻³ Traditionally, researchers have focused on the overexpression of rate limiting enzymes, deletion of competitive pathways and resource management for the production of non-endogenous products.^{4,5} However, in nature, microorganisms have developed an alternative strategy to regulate complex metabolic reactions by eliciting proximity control over enzyme organization.⁶ This enzyme organization, also known as a metabolon, is a multienzyme complex that allows for the direct passage of a product from one enzymatic reaction to the next enzyme in a metabolic pathway.⁷⁻¹⁰ Metabolon formation offers several advantages over traditional metabolic engineering methods such as the ability to limit the diffusion of transient intermediates, facilitate the fast turnover of toxic intermediates, prevent crosstalk between competitive pathways as well as the ability to concentrate reactants which can drive unfavorable reactions.¹⁰

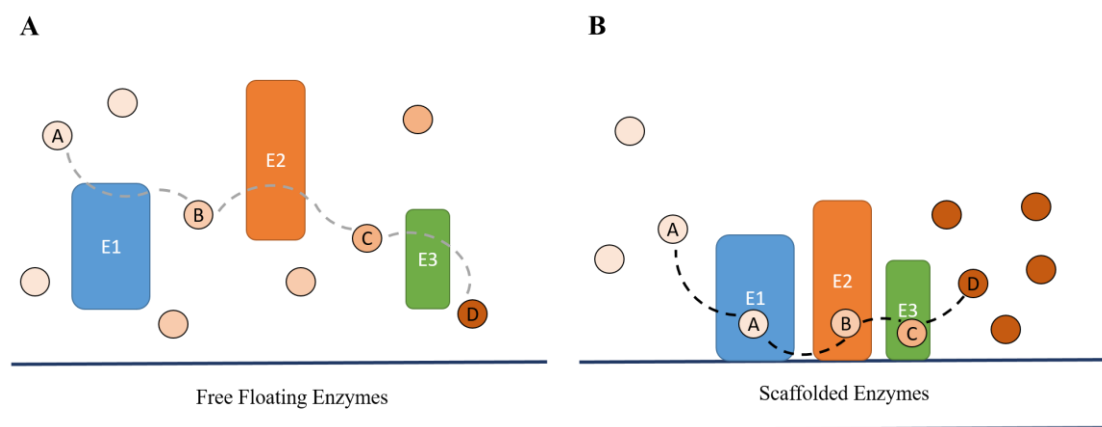


Figure 1.1 **Benefits of Metabolon Formation.** (A) Free floating pathway enzymes within the cell often suffer from loss of transient intermediates to side reactions, secretion, as well as induce toxicity to the host cells which limits products. (B) When pathway enzymes are scaffolded and brought into close proximity it offers several advantages such as the ability to limit the diffusion of transient intermediates, facilitate the fast turnover of toxic intermediates, prevent crosstalk between competitive pathways as well as the ability to concentrate reactants which can drive unfavorable reactions.

Many protein engineering groups have offered several synthetic metabolon platforms, each with unique characteristics. A few notable synthetic metabolon strategies include using proteins for scaffolding¹¹⁻¹³, while others focused on nucleic acid based strategies¹⁴⁻¹⁷, each applied to a wide range of enzymatic pathways. While these prior strategies had varied levels of success, they each suffer from some disadvantages. For instance, scaffolds made up by large fusion proteins tend to suffer from expression issues such as improper folding and loss of activity. While nucleic acid strategies focused around HaloTag fusion proteins are extremely costly and cannot be used in vitro in a site-specific manner, and more cost effective mechanisms

like zinc finger proteins require complex engineering only to have subpar binding affinities. These inherent disadvantages lead to less efficient synthetic metabolon formation or limitations in application space in general.

In contrast, engineering a new alternative synthetic scaffolding system that harnesses all of the advantages, such as high binding affinity and ability to function in vivo and in vitro, while being highly modular in nature could be immensely helpful. One such way to achieve this is to utilize the newly discovered CRISPR/Cas9 system in which enzymes could be directed to a DNA target of interest with only minor engineering. An alternative strategy would be to exploit the inherent properties of self-assembling nanocages to co-localize pathway enzymes.

1.1 Synthetic Metabolon Formation Inspired by Nature

Synthetic chemists often turn to nature for inspiration for today's engineering challenges; this includes efforts towards increasing the production of cost-effective and environmentally friendly products via metabolic engineering and synthetic metabolon formation. In nature, optimal metabolic pathway performance often arises from enzyme organization into specific complexes. Two naturally occurring examples of enzyme organization include those involved in tryptophan synthesis and purine synthesis. Specifically, the last two enzymes involved in tryptophan synthesis form a complex in which the intermediate indole is channeled from the active site of one enzyme to the next.¹⁸ Purine synthesis is a particularly interesting case due to the dynamic nature of the metabolon, also known in this case as a purinosome. Endogenously, purine synthesis is catalyzed by six enzymes in the cytoplasm of

eukaryotic cells, and it has been shown that the six enzymes co-localize in response to cellular purine levels.¹⁹ These two examples represent an ever growing number of natural enzyme organization that have inspired the study of synthetic metabolon formation.²⁰ An extensive amount of research has gone into the engineering of synthetic scaffolds either by protein scaffolding methods or nucleic acid based strategies due to the modular nature of design.^{10,21}

1.1.1 Synthetic Metabolon via Protein Scaffolds

Initial attempts to scaffold metabolic pathways enzymes focused on the engineering of chimeric proteins via the direct fusion of pathway enzymes with short linkers. In 1985, the first successful direct fusion protein scaffold was reported by Mosbach et al. This group directly fused *E.coli* β -galactosidase (LacZ) to *E. coli* galactokinase (GalK) for the conversion of lactose to galactose-1-phosphate.²² After this initial success of direct protein-protein fusion, other groups expanded this concept to other metabolic pathways with varying success. For instance, one group fused the *GPD*₁ and *GPD*₂ from *S. cerevisiae* for the production of glycerol in *E. coli*,²³ while others focused on enzyme fusion for the synthesis of trehalose,²⁴⁻²⁷ NADH recycling,²⁸ and the hydrolysis of starch.²⁹ The aforementioned examples of chimeric protein engineering presented with a range of success from a 0.2 – 3.0-fold increase in expected product titer.³⁰

This was not the case for all direct fusion proteins, as one such attempt from Snell and co-workers resulted in reduced product titer. This group attempted to simplify the production of polyhydroxybutyrate in plants by directly fusing a thiolase and reductase with a 26 amino acid linker; however, they saw half of the polyhydroxybutyrate produced when compared to the individual enzymes.³¹ Due to

the possible loss of function when enzymes are directly fused and the limitations that this platform only allows for two enzymes to be fused, researchers have moved away from direct fusion and towards a more modular approach.

Fierobe and co-workers spearheaded the move away from chimeric enzymes by utilizing complementary protein modules and localized enzymes on a protein scaffold. Specifically, this group focused on the species-specific cohesin-dockerin interactions, where they genetically fused together cohesin modules from various species (the scaffold) while they fused complementary dockerin modules to the enzymes in the cellulose hydrolysis pathway (binding partner).³² This modular protein scaffold inspired others to use native protein interacting domains, such as the metazoan signaling proteins (SH3, PDZ and GDB) with their respective peptide ligands.¹¹ Dueber and co-workers produced a protein scaffold composed of three orthogonal signaling proteins fused together while the peptide ligands were fused to pathway enzymes to modulate enzyme organization as well as enzyme stoichiometry. They were able to achieve up to a 77-fold improvement in mevalonate production though this impressive fold increase was not seen with the same platform when glucaric acid production was accessed.³³ This study emphasized the need for modularity in the synthetic platform as well as the importance that enzyme stoichiometry plays in pathway optimization.^{11,33,34} While this method of utilizing protein modules with complementary binding partners offered advantages such as increased modularity, it suffers from a few dramatic drawbacks. For instance, there are a finite number of protein binding pairs with a high binding affinity, in fact the metazoan signaling proteins are in the low micromolar range which leads to about 50% full scaffold formation.³⁵⁻³⁷

Although success was obtained using protein as a scaffold, either by direct fusion or utilizing protein scaffolds and binding partners, there remains a major issue across both methods. It has been well documented that large fusion proteins often suffer from expression issues, insolubility, protein misfolding as well as degradation problems.³⁸ Such drawbacks limit the range of metabolic pathways that these platforms can be applied towards.

1.1.2 Synthetic Metabolon via Nucleic Acid Scaffolds

Employing nucleic acids as a means of scaffolding pathway enzymes has a few unique advantages since DNA is easy to produce, has a highly predictable and controllable structure that is stable even at longer lengths.⁹ These advantages over protein scaffolds lead to various nucleic acid-based enzyme co-localization strategies such as the chemically conjugated nucleic acid strands or zinc finger protein (ZFP) fusions.

Due to the base pairing properties of DNA it is easy to understand why metabolic engineers looked to chemically modify pathway enzymes with synthetic DNA that could target a DNA scaffold in a controllable manner. One such group chemically modified both glucose oxidase and horseradish peroxidase with DNA oligonucleotides. These DNA oligonucleotides directed the enzymes to a single-stranded DNA template in a sequence specific manner. They reported that the enzyme cascade could not be activated in the absence of the scaffolding DNA.³⁹ While direct chemical modification of DNA to proteins allowed for the formation of functional synthetic scaffolds, these modifications often cause loss of function of the enzyme.⁴⁰ To circumvent this shortcoming, several groups have exploited the self-labeling HaloTag, which is a 34 kDa monomeric mutant haloalkane dehalogenase that allows

for the covalent binding to synthetic chlorohexane modified DNA.⁴¹ They utilized this HaloTag technology to co-localize four enzymes used in cellulose hydrolysis onto a single strand DNA target, which allowed for a 2.3-fold increase in product titers when compared to the free floating enzymes.⁴² The increase in product titer suggests that using DNA as a scaffold is possible. However, chemical modification of DNA to pathways enzymes is a costly technique and not practical *in vivo* when looking to co-localize pathway enzymes in a sequence specific manner.

The focus shifted away from DNA chemical modification and towards proteins that can site specifically bind to DNA, such as ZFPs. ZFPs have modular structures that can be engineered to bind any unique DNA sequence (9-18 base pairs) with sub-micromolar affinity, to date there are over 700 documented ZFPs.^{43,44} Our group harnessed the specificity offered by ZFPs and created ZFP fusions with enzymes used in cellulose hydrolysis. When the enzymes were co-localized on double stranded DNA directed by orthogonal ZFPs they saw a 1.7-fold increase in reducing sugar production.⁴⁵ This ZFP scaffolding platform was also used *in vivo* for investigating the effects of enzyme co-localization towards multiple metabolic pathways including resveratrol, 1,2-propanediol as well as mevalonate, all of which were found to improve product titers by 3- to 5-fold.⁹ While double stranded DNA appears to be a superior choice for *in vivo* metabolic scaffolding due to its stability, modularity and ease of production, pairing the platform with ZFPs may not be ideal. ZFPs have binding affinities in the sub-micromolar range, leading to about 50% scaffolding efficiency, as reported.⁴⁵ ZFPs are also relatively difficult to engineer. In order to keep the ZFPs as small as possible, they tend to only recognize 9 base pairs which can lead to off target

binding to the host genome.⁹ These shortcomings leave room for improvement for a modular DNA scaffolding platform.

1.2 CRISPR/Cas Technology

In 2002, Jansen and coworkers reported the discovery of clustered regularly interspaced short palindromic repeats (CRISPR). They noted that these CRISPR loci were found only in Archaea and bacteria and with adjacent CRISPR-associate (Cas) genes.^{46,47} The function of the CRISPR/Cas system remained elusive until 2005 when it was discovered that the spacer sequences found in the gene loci had shared homology to extrachromosomal elements leading to the hypothesis that the major function was in fact an immune system of sorts.⁴⁸⁻⁵⁰ It was then that the endogenous functions and structure of the CRISPR/Cas system began to be reported. Natively, the Cas proteins work in conjunction to incorporate novel spacers from invading viruses, storing these genetic materials in the CRISPR loci,⁵¹ the stored sequences act as a memory for invasive DNA.⁵² These proteins, in combination with nucleic acids, form the nucleoproteic complexes which can perform a variety of functions essential to the CRISPR/Cas immune system.⁵³

Since the discovery of CRISPR/Cas9 and the elucidation of some of the key features, there has been an exponential growth in research to fully harness this new technological capability. The CRISPR/Cas systems are divided into five classifications based on signature genes associated with the loci.⁵⁴ The type II CRISPR/Cas system is the simplest among the classes due to the Cas9 protein, which is a multidomain protein that is guided to targeted DNA via mature tracrRNA:crRNA molecule.⁵²⁻⁵⁴ The mature tracrRNA:crRNA is composed of two RNA components, the CRISPR RNA (crRNA) that harbors the 20 nucleotide guide segment needed for DNA targeting, and

the trans-activating crRNA (tracrRNA) which allows for hybridization with the Cas9 protein. The tracrRNA:crRNA hybrid structure binds to the Cas9 protein and guides the complex to the DNA, where the multidomain protein cleaves double stranded DNA (dsDNA) via the HNH and RuvC nuclease domain.⁵²⁻⁵⁷ Cas9 binding requires an additional 2 to 6 base pair sequence, coined the protospacer adjacent motif (PAM), proximal to the target site.⁵⁸ The PAM sequence, as well as the tracrRNA:crRNA complex, confer specificity to individual Cas9 species, thus allowing for Cas9 proteins from various species to function orthogonally in the same cell.⁵⁹ Since the discovery that Cas9 proteins function orthogonally they have been used for gene regulation,⁶⁰ as well as fluorescent labeling of chromosomal loci.⁶¹ Endogenous CRISPR/Cas9 while impressive on its own, has become a superstar amongst protein engineers, simplifying the components and expanding to an array of applications.

1.2.1 Engineering of the CRISPR/Cas Technology

As mentioned previously, endogenous Cas9 proteins hybridize with the tracrRNA:crRNA complex and are guided to a dsDNA target that harbors a PAM sequence specific for the species. The binding of the Cas9:RNA complex causes double stranded breaks, cleaving the DNA via two nuclease domains. In recent years, the Cas9 protein has been engineered into a nuclease-null (dCas9) version that lacks the nuclease capabilities. Qi and coworkers introduced two point mutations in the RuvC and the HNH nuclease domain (D10A and H841A) that caused a loss of nuclease function, but retained the ability to hybridize to the tracrRNA:crRNA complex and subsequent DNA binding.⁶² This engineering feat opens a new avenue for directed protein binding utilizing modular RNA targeting. In fact, the Cas9 protein has been shown to tolerate a wide variety of direct fusions such as fluorescent

proteins,^{61,63} transcription activators,^{64–66} transcription repressors,^{65–67} antigen tags,^{68,69} and light inducible Cry2 proteins^{70,71} to name a few. The high binding affinity offered endogenously by Cas9 as well as its ability to tolerate a wide range of protein fusion of varying sizes, make it a wise choice as a potential scaffolding partner.

To simplify the system even further, Jinek and coworkers engineered a single chimeric RNA (sgRNA) to guide the dCas9 protein to its target. Specifically, this group joined the 3' end of crRNA to the 5' end of the tracrRNA with a flexible GAAA tetraloop.⁵⁶ The engineered sgRNA not only retained its endogenous functions, it reduced the CRISPR/Cas9 system from three components to two. Recently, our lab further engineered the sgRNA to obtain conditional control of Cas9 binding. They added a toehold-gated riboregulator to the 5' end of the sgRNA scaffold (thgRNA), this change in conformation disrupted the ability of the gRNA to bind to the Cas9 protein, subsequently mitigating the ability to bind to a DNA target. However, when a trigger strand with complementary sequence to the toehold riboregulator is introduced to the system, the thgRNA opens up into its active form and regains the ability to complex with Cas9 and bind DNA.⁷² This conditional control could be of great use in terms of dynamically controlling scaffolding formation for optimized product titers.

1.3 Cellulose Hydrolysis via Cellulosome Formation

Annually, approximately 10^{11} tons of plant biomass are hydrolyzed by a variety of microbes; in regard to energy released, this is equivalent to 640 billion barrels of crude oil.⁷³ When compared to food crops that are currently used for biofuels, cellulose has a few key advantages: it is the most abundant renewable source of carbon and energy in the world and it is extremely cost effective. Next generation

biofuels based on cellulose biomass will have a lower carbon footprint and with improvements in methods of production, has the potential to be a major fuel source in the future.⁷⁴

Cellulose in nature is synthesized by individual molecules of glucosyl residues which undergo self-assembly. Individual cellulose molecules then further self-assemble into larger units, protofibrils, which are packed even tighter into larger units called microfibrils, which in turn are assembled into cellulose fibers.⁷⁵ These tightly packed fibers are stiffened by intra- and inter-chain hydrogen bonds. In addition to hydrogen bonds, adjacent sheets are held together by weak van der Waals forces.⁷⁵ These highly ordered fibrils are extremely difficult to penetrate and highly rigid. The qualities of cellulose fibers make them extremely difficult to break down for industrial purposes. Due to the significant role cellulose hydrolysis can have on biofuel production, there is an increased need for highly efficient cellulose hydrolysis systems that are economical.

1.3.1 Native Cellulosome Formation

Nature has evolved a pathway for breaking down cellulose, first discovered by Bayer and Lamed in the 1980s.⁷⁶⁻⁷⁸ They characterized a co-localized enzyme structure, coined a cellulosome, in the anaerobic thermophilic bacterium *Clostridium thermocellum*. The cellulosome is a highly ordered structure maintained via protein-protein (nM range) interactions between a scaffoldin (containing cohesin modules) and dockerins at the c-terminus of cellulosomal enzymes.⁷³ A variety of cellulases, including endoglucanases, exoglucanases and β -glucosidase enzymes are required for the synergist breakdown of cellulose.⁷⁹ The entire cellulosome is directed to its

cellulose substrate by a cellulose binding module (CBM).⁸⁰ Natural cellulosomal formation exploits enzyme synergism due to spatial proximity while retaining the flexibility required to optimize cellulose hydrolysis.

1.3.2 Synthetic Cellulosome Formation

For the past few decades, metabolic engineers have studied native cellulosomes and tried to replicate the highly ordered structure for a more efficient, industrially relevant scaffolding system capable of increasing the rate and product titers of cellulose hydrolysis. In line with these goals, Fierobe and co-workers produced chimeric scaffoldins and dockerin containing hybrid enzymes and saw that the cellulosome chimeras exhibited improved enzyme synergy on a microcrystalline cellulose substrate.³² This index case of synthetic cellulosome formation ignited other synthetic strategies. Through these studies, and developments of alternative scaffolding systems, groups learned more about the properties needed for effective cellulose hydrolysis such as overall accessibility of reactive sites, proximity of enzymes,^{81,82} incorporation of multiple divergent enzymes,^{82,83} and the role scaffold flexibility plays in cellulose hydrolysis.^{85,86} While the aforementioned strategies all revolve on the cohesin-dockerin interactions, other researchers sought to use alternative strategies to co-localize cellulolytic enzymes such as chemically modified DNA and ZFPs.

Specifically, the Chen lab has studied the use of DNA as the scaffolding template due to its overall flexible structure which has been shown to be a key guideline in synthetic cellulosome formation. Initially, they utilized zinc finger protein guided assembly to form two component cellulosome structures. Researchers fused an endoglucanase, CelA, and the cellulose binding module (CBM) to two orthogonal

ZFPs. They showed the ability to organize the two enzymes onto a DNA scaffold in a site-specific manner causes an increase in cellulose hydrolysis by 1.6-fold when compared to the free floating enzymes.¹⁷ While this increase in product titers confirms the ability to use DNA as a scaffold, the ZFPs have a binding affinity around the 0.1 to 1 μ M range, leading to only 50% cellulosome formation. In an effort to increase the binding affinity of the cellulolytic enzymes to the DNA scaffold, they began to shift away from ZFPs and towards a strategy utilizing the self-labeling HaloTag. HaloTag is a monomeric mutant haloalkane dehydrogenase that covalently binds to a synthetic chlorohexane (CH) modified DNA molecule.⁴¹ They genetically modified cellulosomal enzymes with HaloTag which allowed for subsequent attachment of CH-modified oligonucleotides. These oligonucleotide-decorated enzymes can then site-specifically hybridize with single stranded DNA, forming a functional cellulosome capable of increasing cellulose hydrolysis by 2.3 fold when compared to the free floating components.¹⁶ The increase in fold enhancement observed when HaloTag is used can be attributed to the increase in binding affinity but the technology itself can only be used for *in vitro* site-specific organization of multiple enzymes, thus leaving room for improvement.

1.4 Orthogonal Post Translational Ligation Strategies

As mentioned previously, larger protein fusions suffer from issues of expression, truncation and aggregations.³⁸ One way to alleviate the aforementioned issues is to express proteins separately and utilize post translation ligation strategies to form the full-length chimeric protein of interest. Recently, a novel method for post translation ligation was developed based on an engineered fibronectin-binding protein from *Streptococcus pyogenes*, which contains a domain capable of forming a

spontaneous isopeptide bond between lysine and asparagine residues.⁸⁷ They split the domain into two fragments: a short 13 amino acid peptide termed SpyTag, and a 15 kDa protein binding partner coined SpyCatcher. This split protein domain was capable of direct fusions to other proteins and retained its ability to rapidly form an isopeptide bond. This reaction was tested at a range of temperatures, pH and buffering conditions, across a broad range this reaction occurred rapidly and to greater than 80% efficiency, both *in vivo* and *in vitro*.⁸⁷

In an effort to develop an orthogonal method for post translation ligation of proteins, a different protein domain from *Streptococcus pneumoniae* was engineered in a similar fashion. Veggiani and coworkers engineered the adhesin RrgA into a 13 amino acid peptide, SnoopTag and a 15 kDa binding partner, SnoopCatcher. This SnoopTag: SnoopCatcher pair forms a spontaneous isopeptide bond with greater than 99% efficiency across a variety of buffering conditions and direct fusion pairs. The isopeptide bond formed with the SpyTag:SpyCatcher pair is completely orthogonal to the isopeptide bond formed when using the SnoopTag:SnoopCatcher pair.⁸⁸ The development of an orthogonal post translational ligation strategy allows for efficient expression of larger proteins of interest that can covalently bind in a specific manner for a variety of applications.

1.5 Self-Assembling Protein Nanocages

It should be noted that not all enzymatic pathways are enhanced by the same scaffolding characteristics. For instance, the mevalonate pathway has shown that a stoichiometric ratio of 1:2:2 (AtoB:HMGS:HMGR) was the optimal scaffolding parameters showing a 77-fold increase in mevalonate production,¹¹ while previously discussed cellulosome examples improved based on just co-localization.^{16,17,81-83}

Variable parameters of synthetic scaffolds include stoichiometric ratio, site specificity, flexibility of the scaffold, 2D structure (linear vs globular) and overall scaffold size.

Synthetic cellulosomes based on spherical nanoparticles, both inorganic and organic have been utilized with varying fold increases between 2 and 7, for example the 0.11 μm ferromagnetic nanoparticle,⁸⁹ 20 nm cadmium selenide nanoparticle,⁹⁰ 20 nm polystyrene nanoparticle,⁹¹ 5-10 nm quantum dots,⁹² and approximately 15 nm rosettazyme protein complex.⁹³ While each strategy showed a different level of enhancement, cellulosomes scaffolded onto nanoparticles had the increased benefit of being active and stable at a wider range of temperatures and pHs due to the inherent properties of the nanoparticles. The drawback to the aforementioned strategies are that they each rely on chemical modification or labor intensive synthesis for cellulosomal formation. As protein engineers, the Chen lab wanted to explore the possibility of using protein nanocages for the self-assembly of synthetic nanocages focused on the well characterized E2 and HBV nanocages.

1.5.1 E2 Nanocage

The E2 nanocage derived from the pyruvate dehydrogenase complex of the thermophilic bacterium *Bacillus stearothermophilus* is another attractive protein nanocage because of its intrinsic stability under extreme conditions.⁹⁴⁻⁹⁶ This naturally occurring protein nanocage is composed of 60 monomers that self-assemble into a highly stable 24 nm protein nanocage.^{94,97-100} E2 has been utilized in previous research for drug conjugation and peptide display due to its surface exposed N terminus, making it extraordinarily simple to add functionalities to the surface of the E2 nanoparticle.¹⁰¹ Previously, E2 has been decorated with the endoglucanase CelA via sortase-A ligation.⁹⁴ The decorated E2 nanocage was visualized via transmission

electron microscopy (TEM) and showed fully assembled capsids with an average of 33nm.⁹⁴ While this method showed the ability to utilize the E2 nanocage for cellulolytic enzyme decoration, there was no enzymatic study performed and the ligation method left a large amount of the E2 nanocage unreacted. Unreacted E2 monomers still retain the ability to self-assemble with decorated monomers leading to non-optimized scaffolds.

1.5.2 Hepatitis B Virus (HBV) Nanocage

Another protein nanoparticle of interest is the Hepatitis B virus (HBV) capsids due to its large size, made up of 240 monomer subunits (31 nm in diameter),¹⁰² the high yield expression in *E. coli* and its ability to tolerate genetic insertions of proteins to the interior and the exterior of the nanocage.¹⁰³ Though the HBV nanocage has been used to enhance detection of a protein marker in ELISA based assays,¹⁰⁴ the targeted delivery of siRNA by decorating the exterior with an integrin-binding peptide, and the interior with a p19- RNA binding protein,¹⁰⁵ HBV has not been studied as a scaffolding platform for cellulases. Due to its large size, stability at a range of buffering conditions and its 240-subunit makeup it makes an ideal self-assembling nanocage to study the effects of synthetic cellulosome formation.

Chapter 2

A DYNAMIC SYNTHETIC METABOLON VIA dCAS9 GUIDED ASSEMBLY

2.1 Abstract

Here we reported a new strategy to construct synthetic metabolons using dCas9-guided assembly. Three orthogonal dCas9 proteins were exploited to guide the independent and site-specific assembly of their fusion partners onto a single DNA scaffold. This new platform was applied towards the construction of a two-component cellulosome. Because of the superior binding affinity, the resulting structures exhibited both improved assembly and reducing sugar production. Conditional enzyme assembly was made possible by utilizing toehold-gated sgRNA (thgRNA), which blocks cellulosome formation until the spacer region is unblocked by a RNA trigger. This platform is highly modular owing to the ease of target synthesis, combinations of possible Cas9-fusion arrangements, and expansion to other metabolic pathways.

This Chapter is a reprint with permissions from the Royal Society of Chemistry. Emily Berckman and Wilfred Chen, Exploiting dCas9 fusion proteins for dynamic assembly of synthetic metabolons, *Chemical Communications*, 55, 8219-8222, 2019.

2.2 Introduction

For the past few decades, enzymes have been exploited for the production of a variety of products, ranging from both bulk and fine chemicals to active

pharmaceutical agents and next generation biofuels.^{106,107} In nature, enzymes are frequently clustered together to coordinate and optimize their collective action via a process called substrate channeling.^{21,108} Rather than random clustering, a protein scaffold is typically used to organize multiple enzyme partners in close proximity to facilitate their concerted actions.¹⁰⁹

One way to mimic this native enzyme assembly is through the design of synthetic metabolons,¹¹⁰ which are artificial multi-enzyme complexes that spatially organize enzymatic pathways in close proximity.⁷⁻⁹ These synthetic metabolons offer many advantages such as limiting the diffusion of transient inter-mediate, facilitating the fast turnover of toxic intermediates, preventing crosstalk between competitive metabolic pathways, as well as having the ability to concentrate reactants which can drive unfavorable reactions.^{17,16,111,112} Most popular synthetic scaffolds for enhanced pathway efficiency are based on the modular metazoan signaling proteins, SH3, PDZ, and GBD, which specifically recruit pathway enzymes tagged with their cognate interaction peptide ligands.¹⁴ However, one major drawback is the relative low binding affinity (K_d in the range of 0.1 to 8 μM). As a result, a very large protein scaffold containing 1 GBD, 6 SH3, and 2 PDZ domains was required to provide the maximum product titer for glucaric acid synthesis.³³ However, it is often difficult to vary the required enzyme density since a new protein scaffold must be redesigned.^{14,33} Moreover, more complex protein scaffolds tend to be either truncated or misfolded.¹¹¹

Nucleic acids represent another attractive biological scaffold for enzyme assembly because of our ability to manipulate the base-pairing property to create defined structures.¹¹³ More importantly, DNA templates can be easily modulated for simple and quick optimization of enzyme stoichiometry, order, and spacing. Although

enzymes can be chemically conjugated with DNA oligos for site-specific hybridization,¹⁶ this often leads to activity loss and limits the strategy to in vitro applications.¹⁷ One alternative is based on enzyme fusions with zinc finger proteins (ZFPs) for site-specific binding onto a double-stranded DNA template.^{9,17} Artificial cellulosomes based on ZFPs have been reported with levels of enhancement in cellulose hydrolysis similar to those achieved using protein scaffolds.¹⁷ While ZFP fusions offer the ease of assembly, the relatively modest binding affinity ($K_d \sim 100$ nM) makes it difficult to control the actual enzyme binding ratio onto each individual scaffold.

Inspired by the ease of ZFPs to guide the formation of enzyme cascades, we present here a new and potent approach that enables specific and high-affinity binding to DNA scaffolds using the CRISPR/Cas9 system. The Cas9 protein is ideally suited for modular enzyme assembly as it naturally offers high-affinity, sequence-specific DNA binding (K_d values ~ 0.05 to 0.5 nM).^{19,20,114} While the native function of Cas9 is for RNA-guided DNA cleavage, nuclease-null Cas9 (dCas9) proteins have been generated while preserving the same high-affinity binding capability.¹¹⁵ Three orthogonal dCas9 protein pairs, each capable of using its own unique sgRNA for independent and specific targeting, were exploited as a new way to create synthetic enzyme cascades. Because of our past success of using DNA scaffolds to create artificial cellulosome structures,¹⁶ we used this platform as an example to demonstrate the effectiveness of our new dCas9-guided assembly strategy (Figure 2.1).

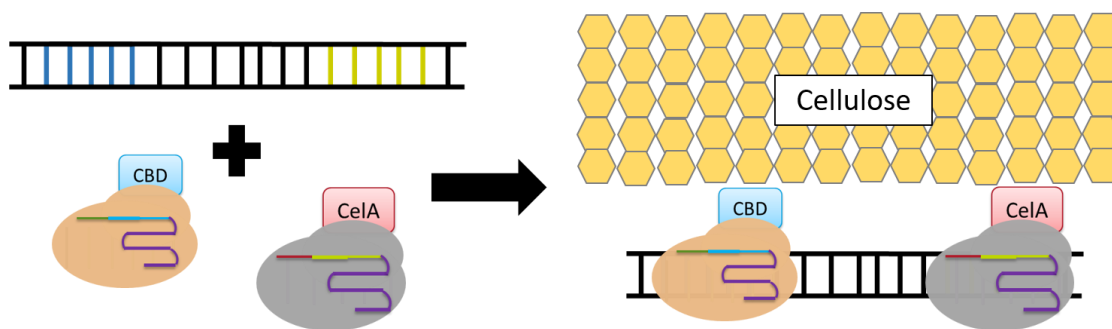


Figure 2.1 **Formation of artificial cellulosome structure by dCas9-mediated assembly onto a double-stranded DNA template.** The cellulose-binding domain (CBD) is used to recruit the endoglucanase CelA for cellulose hydrolysis.

2.3 Materials and Methods

2.3.1 Materials

All primers were purchased from IDT (Coralville, IA). All ingredients for culturing media were purchased from Fisher Scientific (Pittsburgh, PA), all ingredients for SDS-PAGE were purchased from BIO-RAD (Hercules, CA), all enzymes related to DNA manipulation and cloning were purchased from New England Biolabs (Ipswich, MA), and all other chemicals were purchased from Sigma Aldrich (St. Louis, MO).

2.3.2 Strains and Cloning

Escherichia coli NEB5 α (New England Biolabs, Ipswich, MA) [*fhuA2* Δ (*argF-lacZ*)U169 *phoA glnV44* Φ 80 Δ (*lacZ*)M15 *gyrA96 recA1 relA1 endA1 thi-1 hsdR17*] was used as the host for all genetic manipulations and plasmid maintenance. *E. coli* strain BLR (DE3) (EMD Millipore, Madison, WI) [*F*⁻ *ompT hsdS_B(r_B⁻ m_B⁻) gal*

dcm (DE3) Δ(srl-recA)306::Tn10 (Tet^R)] was used as the production host for all proteins.

The nuclease-null Cas9 proteins *S. pyogenes* (SpdCas9), *S. thermophilus* (ST1dCas9) and *S. aureus* (SadCas9) were purchased from Addgene. The dCas9 protein sequences were PCR amplified, double digested, and inserted into pET24a using SpeI and bamHI. Fluorescent protein markers were added into the vectors using the restriction sites bamHI and MfeI for the following final constructs: pET24a SpdCas9-mCherry-6xhis, pET24a SadCas9-mCherry-6xhis and St1dCas9-Lgbit-6xhis. The fluorescent proteins were swapped out for cellulose binding module (CBD) and the endoglucanase CelA5 using bamHI and MfeI to form pET24a SpdCas9-CBD-6xhis and pET24a St1dCas9-Cel5A-6xhis.

2.3.3 Protein Expression

All constructs were transformed into *E. coli* BLR [F- ompT hsdSB (r-B m-B) gal dcm(DE3) Δ(srl-recA)306::Tn10(TetR); Novagen, Madison, WI] cells for protein expression. Expression cultures were grown in Luria-Bertani (LB) media (10.0 g/L tryptone, 5.0 g/L yeast extract, 10.0 g/L NaCl) supplemented with 100 μg/mL ampicillin. Overnight cultures were used to inoculate 30 mL of fresh LB to an initial OD₆₀₀ of ~0.05 and grow to mid-exponential phase (OD₆₀₀ ~0.75) at 37°C. Protein expression of SpdCas9 fusion proteins and SadCas9 fusion proteins were induced with 200 μM isopropyl-β-thiogalactopyranoside (IPTG) and incubated at 20°C for ~16hrs. Expression of ST1dCas9 protein fusions was optimal with induction of 100 μM IPTG and incubation at 37°C for 5 hours.

Induced cultures were pelleted using centrifugation at 3000xg for 10 minutes at 4°C, and the pellet was suspended in 1x TBS (Tris Buffered Saline) to an OD₆₀₀ =

30. The cell suspension was sonicated to release protein and clarified using centrifugation at 15,000xg for 15 min at 4°C. Soluble proteins were immediately purified using affinity chromatography.

2.3.4 Protein Purification

Soluble proteins were mixed with imidazole to a concentration of 5 mM. Meanwhile, gravity columns were packed with HisPur™ Ni-NTA Resin (ThermoFisher Scientific, Waltham, MA). The column was used as instructed using TBS + 10 mM imidazole as the equilibration buffer, TBS + 25 mM imidazole as the wash buffer, and TBS + 250 mM imidazole as the elution buffer. Eluates were dialyzed against TBS overnight. The cell lysates and the purified proteins were loaded onto a 7.5% SDS-PAGE gel and stained with coomassie blue for analysis. When relevant, protein expression was also monitored via fluorescence protein markers.

2.3.5 RNA *In Vitro* Transcription and Purification

All RNAs (sgRNA and thgRNA) were transcribed in vitro via HiScribe T7 Quick High Yield RNA Synthesis Kit (New England BioLabs, Inc., Ipswich, MA, USA) and purified via phenol-chloroform extraction and ethanol precipitation. The purity, quality and quantity of RNA was analyzing by spectrophotometry using a NanoDrop 2000 UV-Vis spectrophotometer (Thermo Fisher Scientific, Inc., Waltham, MA, USA). Then subsequently using a denaturing urea PAGE electrophoresis with a 6% polyacrylamide gel containing 8 M urea.

2.3.6 Single and Dual Beacon Assay

All DNA probes were prepared using unmodified and fluorophore or quencher labeled oligonucleotides synthesized and purified by the commercial vendor Integrated DNA Technologies (Coralville, IA, USA). Target Cas9 beacon DNA shown in Table 2.1 were created by mixing equimolar amounts of complementary DNA strands at a final concentration of 1 μ M in a beacon hybridization buffer (40 mM Tris, pH 7.9, 100 mM NaCl); heating for 5 minutes at 90°C then cooled to 12°C at a rate of 0.1 /s using a S1000 Thermal Cycler (Bio-Rad Laboratories, INC., Hercules, CA, USA). All beacon binding assays were carried out in beacon binding buffer (20 mM Tris-HCl, 120 mM NaCl, 5% v/v glycerol, 0.1 mM DTT, 1 mM MgCl₂, 0.02% v/v Tween 20, pH 7.9) and fluorescence measured using a Synergy H4 Hybrid microplate reader (BioTek Instruments, Inc., Winooski, VT, USA). Final assay mixtures contained 1 nM hybridized beacon target DNA, 10 nM dCas9 protein(s), and 10 nM sgRNA(s). Excitation and emission wavelengths were dictated by the fluorophore in use for each target (A: FAM, ex: 498nm, em:520 nm; B: Cy5 ex: 648nm, em: 668nm).

Sp Single Beacon	5' Iowa Black FQ ataGGTATCACATGACTAAACGAAGGtct 3' FAM tatCCATAGTGTACTGATTTGCTTCCaga
St1 Single Beacon	5' Iowa Black FQ tatATGGAATGGAATGGAATGGATTGGAAaga 3' FAM ataTACCTTACCTTACCTTACCTAACCTTtct
Sa Single Beacon	5' Iowa Black FQ ataGAAGTGTCCGGTGACAGGAAGAGGGGAtct 3' FAM tctCTTCACAGCCACTGTCCTTCTCCCCTaga
Sp/St1 Dual Beacon	5' Iowa Black FQ ataGGTATCACATGACTAAACGAAGGtctagagacta 3' FAM tatCCATAGTGTACTGATTTGCTTCCagatctctgat cgaccgatcattacTTCCAA TCCATTCCATTCCATTCCAT tat 3' Cy5 gctggctagtaatg AAGGTT AGGTAAGGTAAGGTAAGGTAata 5' Iowa Black RQ
Sp/Sa Dual Beacon	5' Iowa Black FQ ataGGTATCACATGACTAAACGAAGGtctagagacta 3' FAM tctCCATAGTGTACTGATTTGCTTCCagatctctgat cgaccgatcattacACCCCT CTTCCTGTACCCGACACTTC tat 3' Cy5 gctggctagtaatg TGGGGAGA AAGGACAGTGGCTGTGAAGata 5' Iowa Black RQ
Sa/St1 Dual Beacon	5' Iowa Black FQ ataGAAGTGTCCGGTGACAGGAAGAGGGGAtctagaga 3' FAM tatCTTCACAGCCACTGTCCTTCTCCCAagatctct ctacgaccgatcattacTTCCAA TCCATTCCATTCCATTCCAT tat 3' Cy5 gatgctggctagtaatg AAGGTT AGGTAAGGTAAGGTAAGGTAata 5' Iowa Black RQ
Triple Beacon	3' FAM ataGGTATCACATGACTAAACGAAGGtctagagactacgaccgatcatta 5' tatCCATAGTGTACTGATTTGCTTCCagatctctgatgctggctagtaat cACCCCT CTTCCTGTACCCGACACTTC ttagataactaactacgaccgat g TGGGGAGA AAGGACAGTGGCTGTGAAGaatctattgattgatgctggcta caATGGAATGGAATGGAATGGATTGGAAata 3' gt TACCTTACCTTACCTTACCTAACCTT tat 5'

Table 2.1 DNA sequences of single, dual and triple targets.

2.3.7 Electrophoretic Mobility Shift Assay

Single, dual and triple beacon assays were qualitatively verified using electrophoretic mobility shift experiments carried out in 4.5% non-denaturing acrylamide gels. Final assay mixtures contained 10 nM hybridized beacon target, 100 nM dCas9 protein(s) and 100 nM sgRNA(s). The fully hybridized beacon DNA only, as well as the beacon DNA bound by the dCas9;sgRNA complex were loaded and run for 90 minutes at a constant 90V. The gel was then imaged using an Amersham™ Typhoon™ Biomolecular Imager (GE Healthcare Bio-Sciences Corporation, Marlborough, MA).

2.3.8 Triple Beacon Assay

All DNA probes were prepared using unmodified and fluorophore labeled oligonucleotides synthesized and purified by the commercial vendor Integrated DNA Technologies (Coralville, IA, USA). Target Cas9 beacon DNA shown in Table 2.1 were created by mixing equimolar amounts of complementary DNA strands at a final concentration of 1 μ M in a beacon hybridization buffer (40 mM Tris, pH 7.9, 100 mM NaCl); heating for 5 minutes at 90°C then cooled to 12°C at a rate of 0.1 /s using a S1000 Thermal Cycler (Bio-Rad Laboratories, INC., Hercules, CA, USA). All beacon binding assays were carried out in beacon binding buffer (20 mM Tris-HCl, 120 mM NaCl, 5% v/v glycerol, 0.1 mM DTT, 1 mM MgCl₂, 0.02% v/v Tween 20, pH 7.9). Final assay mixtures contained 10 nM hybridized beacon target, 100 nM dCas9 protein(s) and 100 nM sgRNA(s). Assay mixtures were loaded and run on a 4.5% non-denaturing acrylamide gels and fluorescence measured using an Amersham™ Typhoon™ Biomolecular Imager (GE Healthcare Bio-Sciences Corporation,

Marlborough, MA). The triple target was labeled with a FAM fluorophore and with excitation and emission wavelengths at ex: 498nm and em: 520 nm.

2.3.9 Reconstitution of Split Nano Luciferase

SpdCas9-Smbit-6xhis and St1dCas9-Lgbit-6xhis were expressed and purified via the above methods. Dcas9-protein fusions were directed to a synthetic DNA target using sgRNA and subsequent reconstitution of full length nano luciferase was monitored via luminescence measurement. Luminescence assays were performed according to NanoGlo[®] vendor's instructions (Promega Corporation, Madison, WI, USA) using a Synergy[™] H4 Hybrid microplate reader.

2.3.10 Synthetic Cellulosome Assembly and Reducing Sugar Assay

Phosphoric acid-swollen cellulose (PASC) was prepared as described previously.¹¹⁶ Final assay mixtures contained 20 nM dCas9 proteins, 20 nM sgRNA(s) with or without 20 nM target DNA in the presence of PASC. Synthetic cellulosome assembly was assayed at 30°C in 20 nM Tris-HCl buffer pH 6.0 with shaking. Samples were collected periodically over a 40 hour time period and immediately mixed with 0.25 mL DNS reagent (10 g/L dinitrosalicylic acid, 10 g/L sodium hydroxide, 2 g/L phenol, 0.5 g/L sodium sulfite) and incubated at 95°C for 10 minutes. After incubation, 0.5 mL of 40% Rochelle salts were added to fix the color before measuring the absorbance using a spectrophotometer at 575 nm.

2.3.11 Conditional Beacon Assay

Conditional beacon assay and cellulosme assembly were run as described above with the exception of the sgRNA being replaced with a thgRNA. Toehold mediated strand displacement reaction was performed as described previously.⁷²

Final assay mixtures contained 10-20 nM beacon DNA target, 10-20 nM dCas9 protein(s) and 10-20 nM sgRNA or thgRNA with or without 100-200 nM trigger strand.

2.4 Results and Discussion

With the overall goal of creating a dCas9-mediated synthetic metabolon capable of cellulosome formation three dCas9 orthologs from *Streptococcus pyogenes* (SpdCas9), *Streptococcus thermophiles* (St1dCas9), and *Streptococcus aureus* (SadCas9) were engineered. The following experiments focus on the localization of dCas9 fusion proteins in a single, dual and triplicate manner onto a synthetic DNA scaffold. Successful co localization was then used as a basis for enzyme reconstitution, synthetic cellulosome formation as well as conditional control of metabolon formation.

2.4.1 Single Binding of dCas9 Fusion Proteins

To demonstrate the feasibility of our new dCas9-guided assembly strategy, three dCas9 orthologs from *Streptococcus pyogenes* (SpdCas9), *Streptococcus thermophiles* (St1dCas9), and *Streptococcus aureus* (SadCas9) were used. These dCas9 proteins require not only different PAM sequences for DNA targeting but also different scaffold domains in their sgRNAs for dCas9 binding.⁵⁹ The first step toward achieving dCas9-guided assembly is to demonstrate that these dCas9 orthologs can bind specifically onto DNA targets in a single, dual, and triplicate manner. A Cas beacon assay, developed previously to determine SpdCas9 binding to their corresponding DNA targets, was adapted for this purpose (Figure 2.2A).¹¹⁷ Binding of the dCas9:sgRNA complex unwinds the DNA strands, causing displacement of the

quencher oligo 2, which can be monitored by an increase in fluorescence (Figure 2.2A).

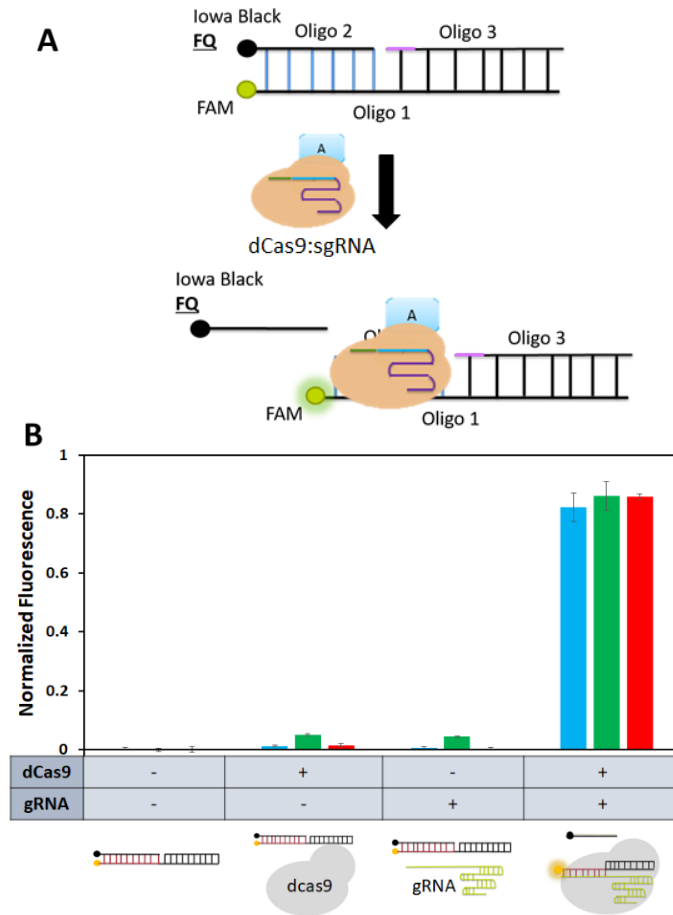


Figure 2.2 **Functionality of individual dCas9 proteins using the Cas beacon assay.** (A) Binding of a dCas9 fusion protein to a Cas beacon target and the subsequent displacement of the quencher strand. (B) Binding efficiencies of three different dCas9 fusion proteins to their corresponding Cas beacon targets. In all cases, dCas9 binding resulted in a significant increase in fluorescence. Blue: SpdCas9, green: ST1dCas9, and red: SadCas9.

We first investigated the binding of individual dCas9 proteins to their unique DNA targets. All assays were performed using purified dCas9 proteins and *in vitro* transcribed sgRNAs (Figure 2B). For each dCas9:sgRNA pair, no detectable increase in fluorescence was observed when either dCas9 or sgRNA was added. In contrast, a significant increase in fluorescence was observed when both were added, indicative of dCas9 binding (Figure 2.2B). Binding of each dCas9 protein onto the corresponding beacon was further validated using the electrophoretic mobility shift assay (Figure 2.3), and a slower mobility band was detected in all cases.

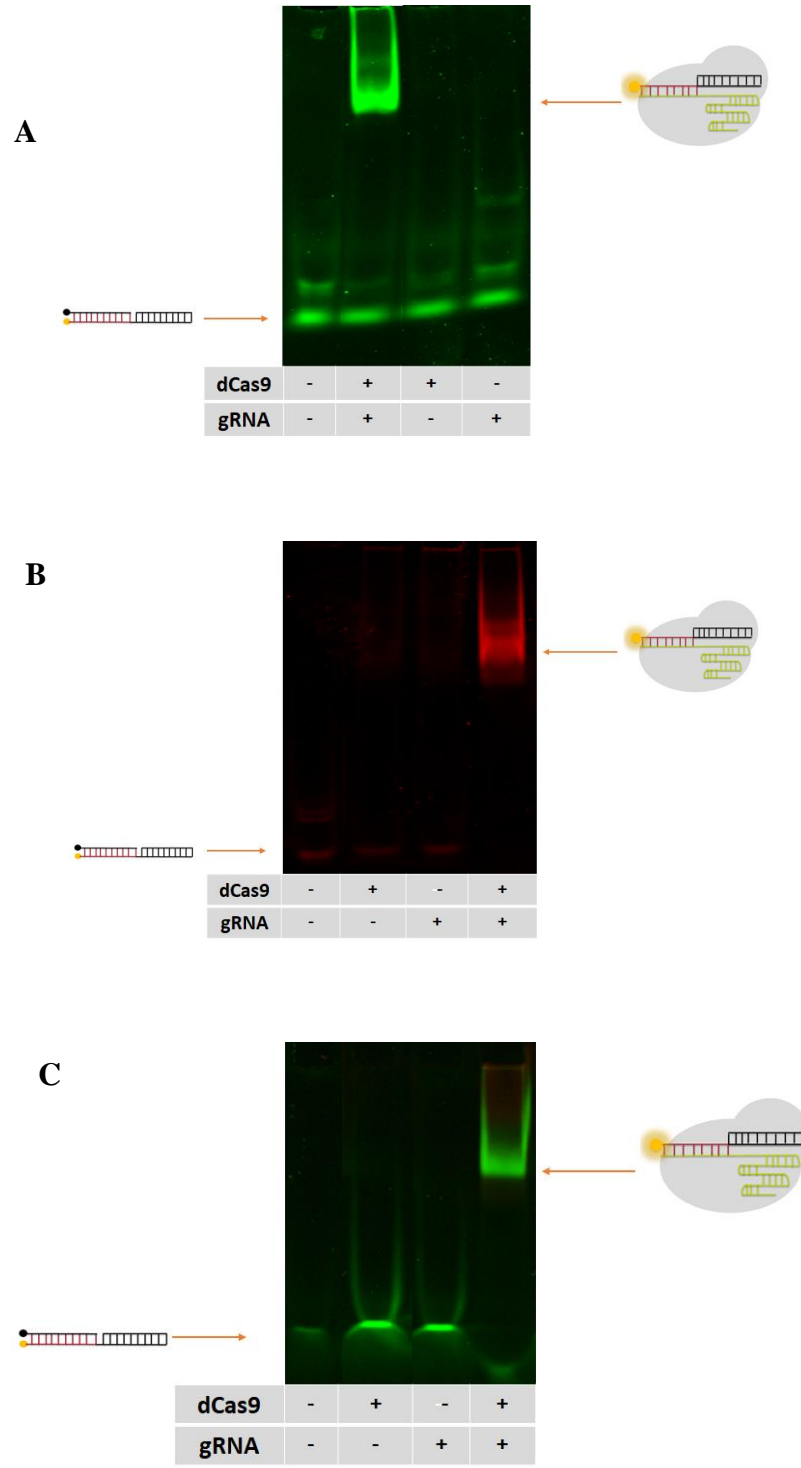


Figure 2.3 EMSA analysis showing qualitative single binding. (A) SpdCas9;SpgRNA (B) St1dCas9;St1gRNA and (C) SadCas9;SagRNA.

2.4.2 Dual Binding of Orthogonal dCas9 Fusion Proteins

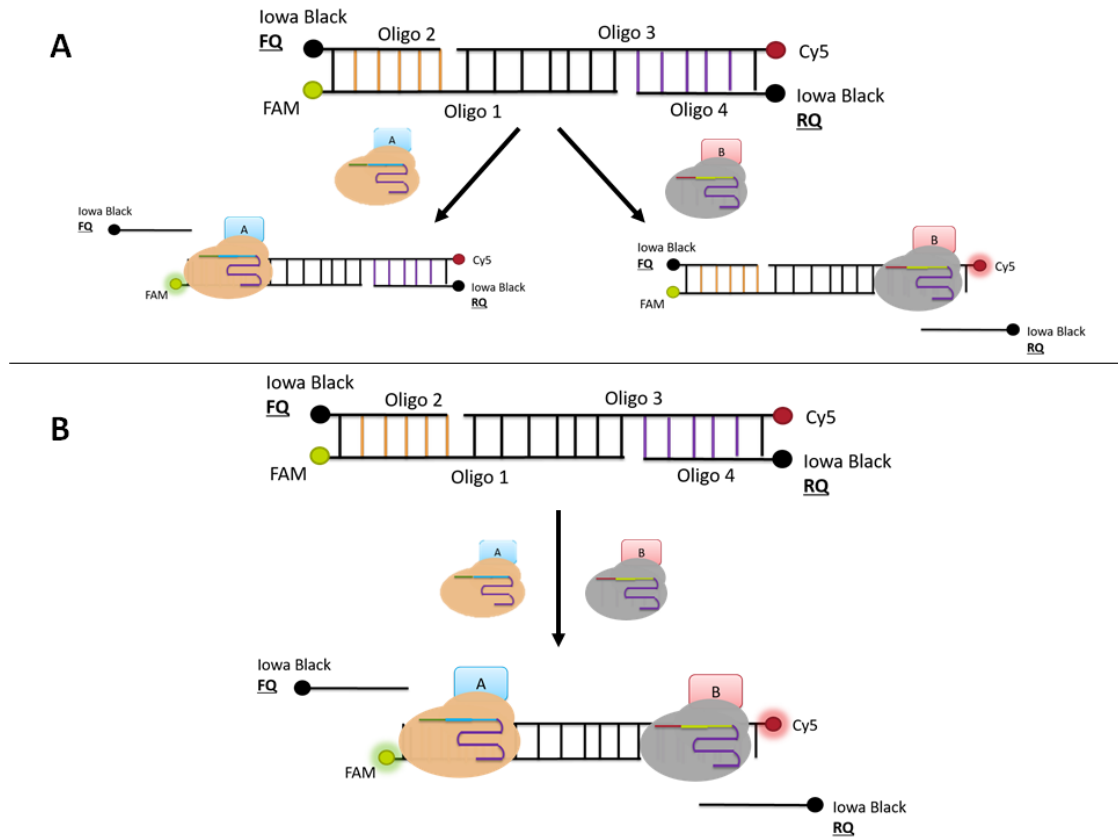


Figure 2.4 **Schematic of dual beacon.** (A) Diagram of single binding of orthogonal dCas9 to the dual beacon. (B) Diagram of dual binding of orthogonal dCas9 to the dual beacon.

Having confirmed the binding ability of all three dCas9 proteins, we next explored their orthogonality to their respective targets. To do so, we expanded upon the single beacon format and developed a new dual beacon approach. Two different fluorophores (FAM and Cy5) were used to detect simultaneous binding of two dCas9 fusions onto the same dual beacon (Figure 2.4 and Figure 2.5A). Independent of the

dCas9 combinations tested, a significant increase in the respective fluorescent signal was observed only when the corresponding dCas9A:sgRNA complex was added (Figure 3B). These results support correct displacement of the quencher strand, indicative of the orthogonality of the three dCas9:sgRNA pairs. More importantly, addition of both dCas9:sgRNA complexes did not hinder binding of either dCas9 protein as a similar increase in fluorescent signal was detected (Figure 2.5 B-D).

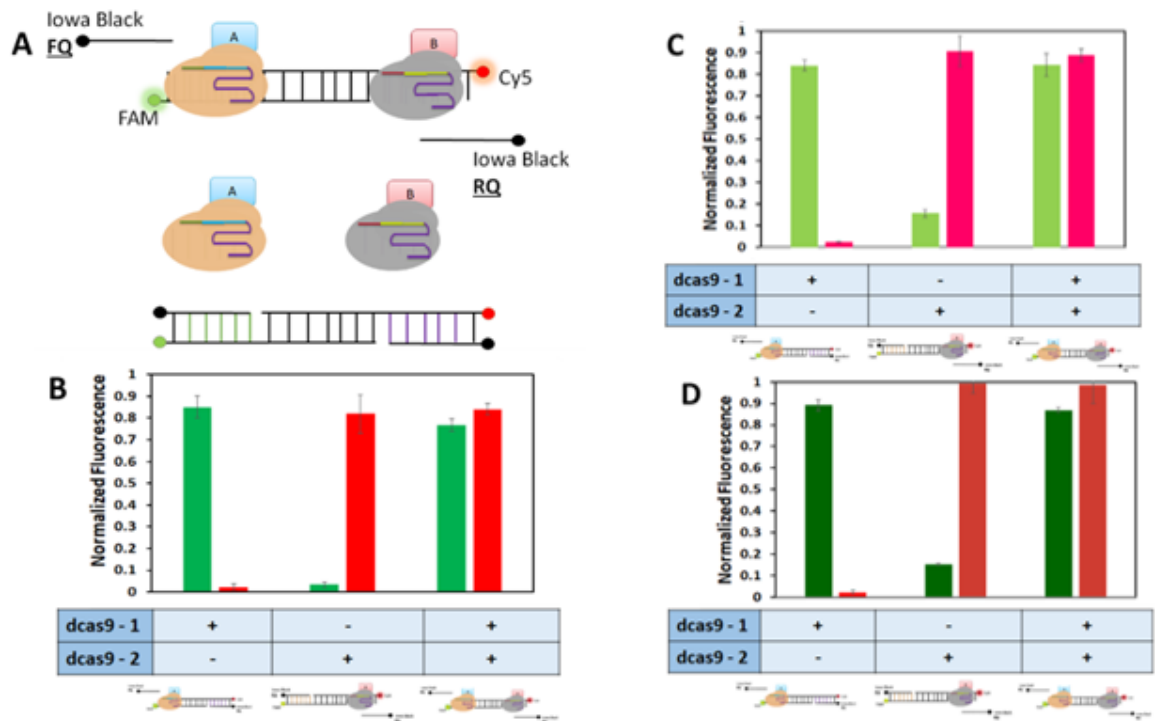


Figure 2.5 **Orthogonal dCas9 dual binding beacon assay.** (A) Schematic of the dual beacon assay. (A) Fully hybridized dual beacon has only background fluorescence. Addition of dCas9A/sgRNA complex results in displacement of the quencher oligo 2 and an increase in FAM fluorescence. While incubation with dCas9B/sgRNA complex displaces only oligo 4 resulting in Cy5 fluorescence. When both dCas9 fusion proteins and the cognate sgRNAs are present, both quencher strands are displaced causing an increase in both FAM and Cy5 fluorescence. (B-D) Binding efficiencies of three dCas9 fusion protein combinations to their corresponding dual beacon targets. (B) SpdCas9/St1dCas9, (C) SpdCas9/SadCa9Green, (D) SadCas9/St1dCas9.

These findings were further supported using the EMSA assay, in which single and dual binding were clearly demonstrated (Figure 2.6).

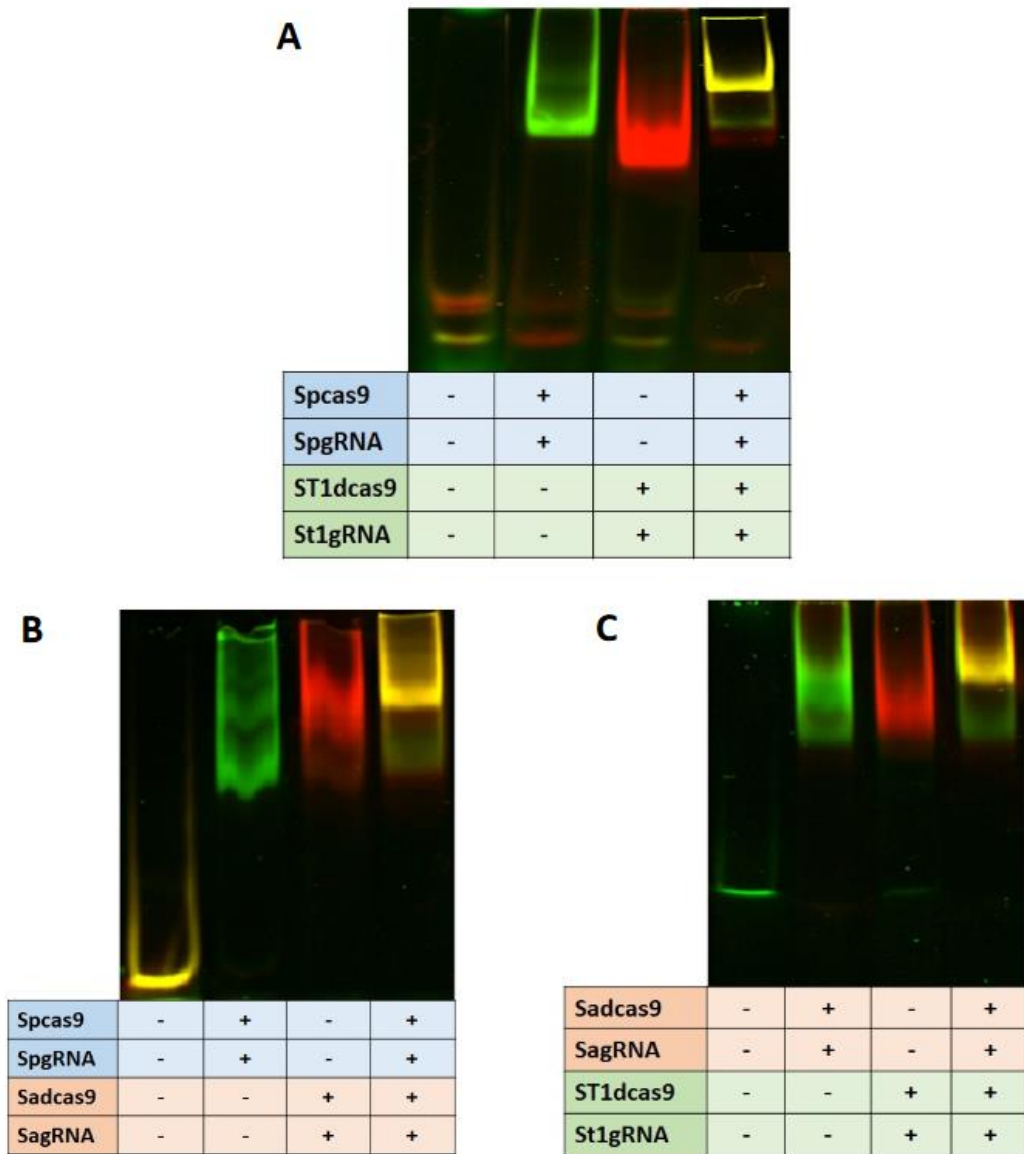


Figure 2.6 **EMSA analysis showing orthogonal and dual binding.** (A) SpdCas9 and ST1dCas9, (B) SpdCas9 and SadCas9, and (C) SadCas9 and ST1dCas9 to their respective dual beacon. Displacement of both quenchers by dual dCas9 binding resulted in the detection of a yellow fluorescent protein – DNA complex band.

It should be noted that a 25 basepair (bp) spacing between the two binding sites is necessary to minimize the steric effect of dualdCas9 binding as a much smaller increase in fluorescence was observed using a shorter 15 bp spacer (Figure 2.7).

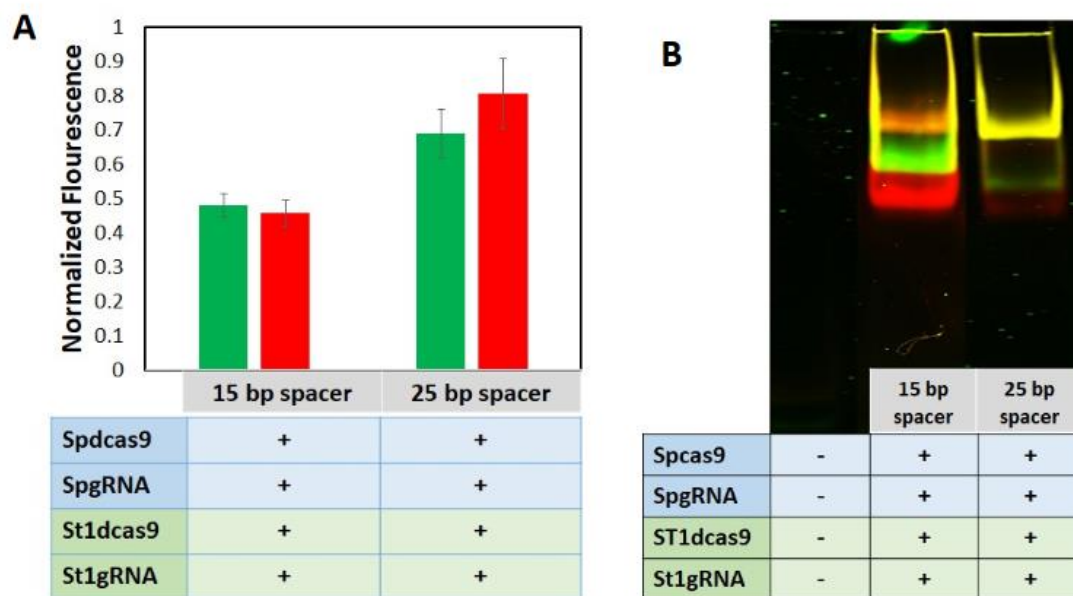


Figure 2.7 **Optimal spacing for efficient dual binding.** (A) Comparison of SpdCas9 and ST1dCas9 binding onto dual beacons containing either a 15-bp or 25-bp spacer between the two binding sites. (B) EMSA analysis showing the less efficient binding of SpdCas9 and ST1dCas9 onto a dual beacon with a 15-bp spacer as compared to a dual beacon containing a 25-bp spacer.

2.4.3 Triple Binding of Orthogonal dCas9 Fusion Proteins

To demonstrate that all three dCas9 proteins can co-localize simultaneously onto a single DNA scaffold, a new template harboring three binding sites was designed. One of the two oligos is fused to a FAM dye for tracking by EMSA (Figure 2.8). Consistent with the beacon results, a gradual decrease in the complex mobility

was observed when either one, two, or three dCas9 proteins was bound. No unbound DNA was detected indicating 100% binding in all three cases. This is significantly better than the roughly 50% assembly using the ZFP-guided approach.¹⁷ These binding results confirm our ability to recruit multiple dCas9 proteins simply by modulating the number of orthogonal binding sites on a single DNA scaffold.

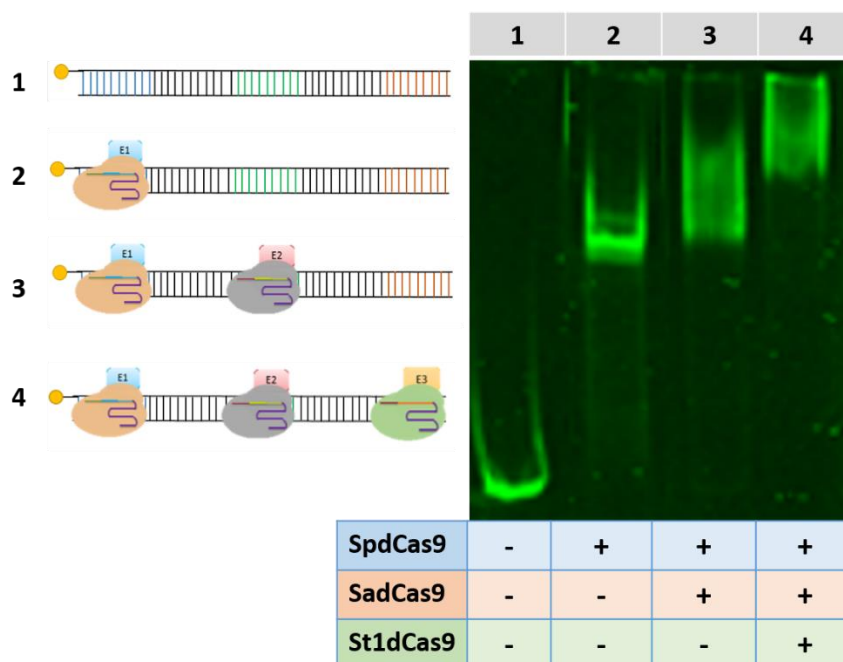


Figure 2.8 **Triple binding of three orthogonal dCas9 proteins.** The sequential binding of all three dCas9 fusion proteins to a FAM-labeled DNA scaffold. Binding of additional dCas9 fusion proteins results in the formation of DNA-protein complexes with decreased mobility.

2.4.4 dCas9 Mediated Split Nano Luciferase Reconstitution

As an initial test to ensure the spacing required for binding of multiple dCas9 fusion proteins to a single target is close enough to encourage enzyme synergy we tested the ability of a split protein to be reconstituted via our synthetic scaffold. Nano

luciferase (Nluc) is an engineered small subunit (19 kDa) from the deep sea shrimp *Oplophorus gracilirostris* luciferase.¹¹⁸ Nluc has been further engineered into two fragments, a 159 amino acid fragment large bit (LgBit) and a 11 amino acid fragment small bit (SmBit), for the use as a complementation reporter dictated by the association or colocalization of their fusion partners.¹¹⁹ We utilized this reporter by fusing the larger fragment to our SpdCas9 (SpdCas9-LgBit-6xhis) and the smaller fragment to ST1dCas9 (ST1dCas9-SmBit-6xhis) and monitored luminescence when one or both fragments were bound to a synthetic DNA target. Only in the presence of both dCas9 fusion proteins and the synthetic DNA target do we obtain a high level of luminescence indicating enzyme reconstitute directed by the co-localization of the orthogonal dCas9 fusion proteins (Figure 2.9). Not only does the reconstitution of split Nluc confirm the ability of our orthogonal dCas9 fusions to bind to our synthetic DNA target but it indicates that the spacing required for dual binding is close enough for enzyme interactions to take place.

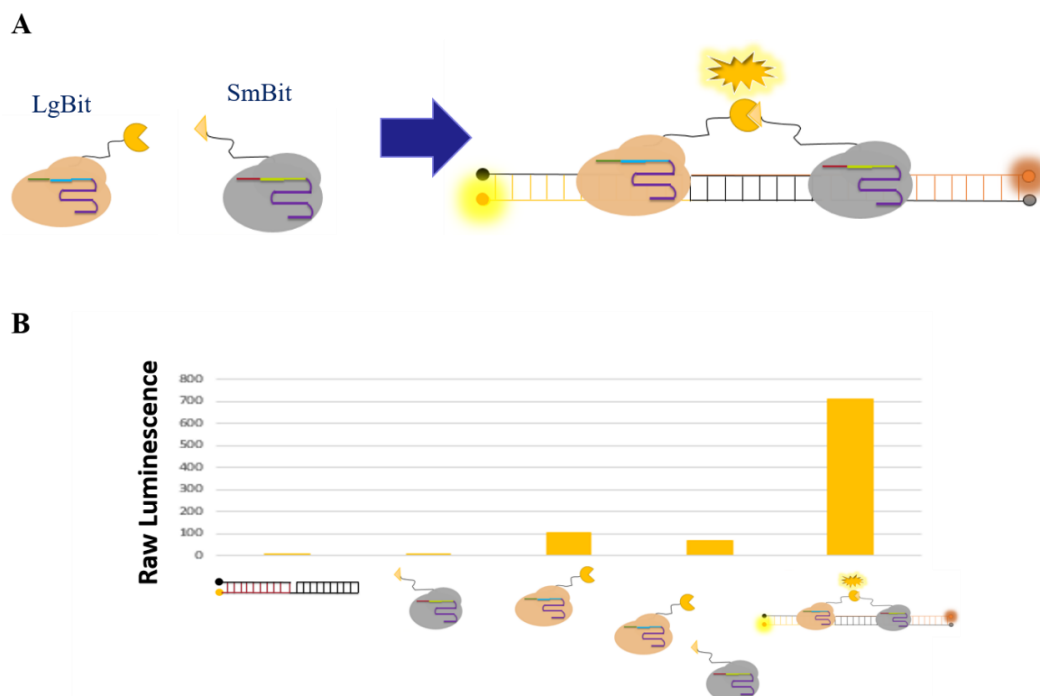


Figure 2.9 **Split Nluc reconstitution.** (A) Schematic of split nano luciferase reconstitution dictated by the co-localization of orthogonal dCas9 fusion proteins. (B) Individual components, DNA only, St1dCas9-SmBit, SpdCas9-LgBit and the two enzymes fusions without DNA show background levels of luminescence. Only when all the components are present do we have reconstitution of split nano luciferase.

2.4.5 dCas9 Mediated Cellulosome Formation

After confirming our ability to co-localize multiple dCas9 proteins onto a single DNA scaffold, we next explored this strategy to construct artificial cellulosomes. A two-component mini-cellulosome structure was created by fusing CBD to the C-terminus of SpdCas9 and the endoglucanase CelA to the C-terminus of St1dCas9. The ability to co-localize both dCas9 fusions was first tested using the Cas beacon assay. Addition of CBD or CelA had no impact on dCas9 binding as a similar

increase in fluorescence was detected (Figure 2.10). The two dCas9 fusions were then incubated with a DNA scaffold to generate a CBD-CelA cellulosome structure. As depicted in Fig. 2.11, a 2.6-fold enhancement in cellulose hydrolysis was observed as compared with unassembled dCas9 fusions. This level of improvement is even better than that obtained using assembly by direct DNA hybridization,¹⁶ again highlighting the benefit of the sub-nanomolar binding affinity of dCas9 proteins in optimizing co-localization of the two cellulosomal components.

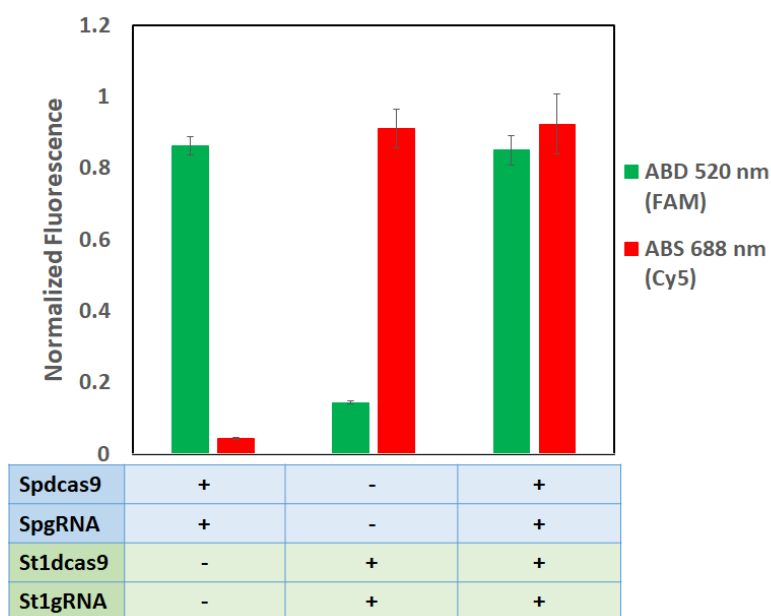


Figure 2.10 Binding efficiencies of SpdCas9-CBD and St1dCas9-CelA onto the dual beacon. Left bars show binding of SpdCas-CBD to the FAM portion of the dual beacon. Middle bars show single binding of ST1dCas9-CelA onto the dual beacon resulting in Cy5 fluorescence only. Right bars show dual binding of both SpdCas9-CBD and ST1dcas-CelA indicated by an increase in both FAM and Cy5 fluorescence.

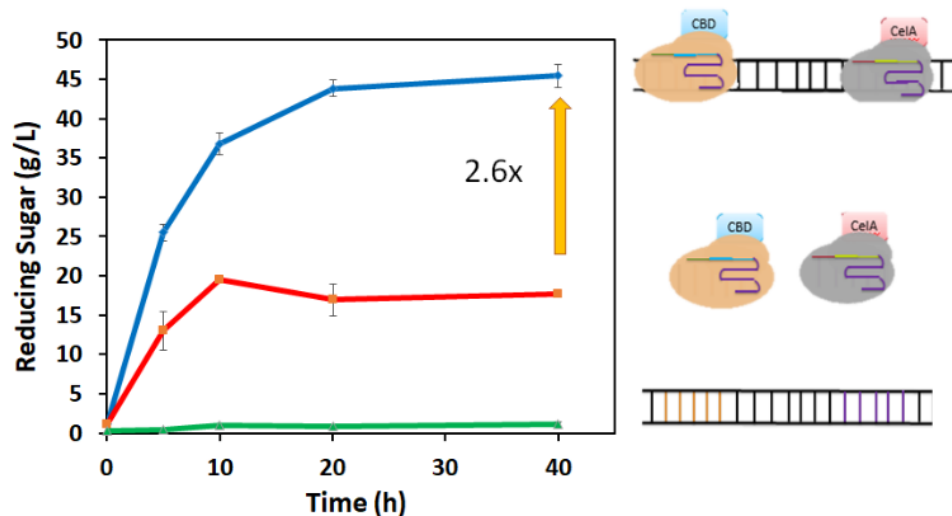


Figure 2.11 **Assembly of two-component cellulosome using SpdCas9-CBD and ST1dCas9-CelA.** The production of reducing sugars from DNA alone, free enzyme fusions, and the assembled cellulosome were compared.

Since the efficiency of many enzyme cascades can be optimized by tuning enzyme stoichiometry and order, another benefit of our dCas9-guided assembly is the ease of reconfiguring the assembled proteins simply by changing the DNA scaffold. A new DNA scaffold was first designed to change the order from CBD-CelA to CelA-CBD (Figure 2.12). A similar level of enhancement in cellulose hydrolysis was obtained since both cellulosomes are equally efficient in bringing CelA in proximity to the substrate. A more complex cellulosome structure containing three CBDs and CelAs was also created by changing the number of binding sites on the DNA scaffold (Figure 2.12). Although the amount of proteins used was the same as before, complete assembly was achieved even with 1/3 the amount of DNA due to the tight binding of dCas9s. This feature is particularly attractive for assembling *in vivo* enzyme cascades in *E. coli* as increasing numbers of binding sites can be inserted into plasmid DNA to

enhance co-localization.^{9,16,17} The feasibility for both *in vitro* and *in vivo* enzyme assemblies while providing sub-nanomolar affinity is perhaps the most distinguishing feature of our new approach.

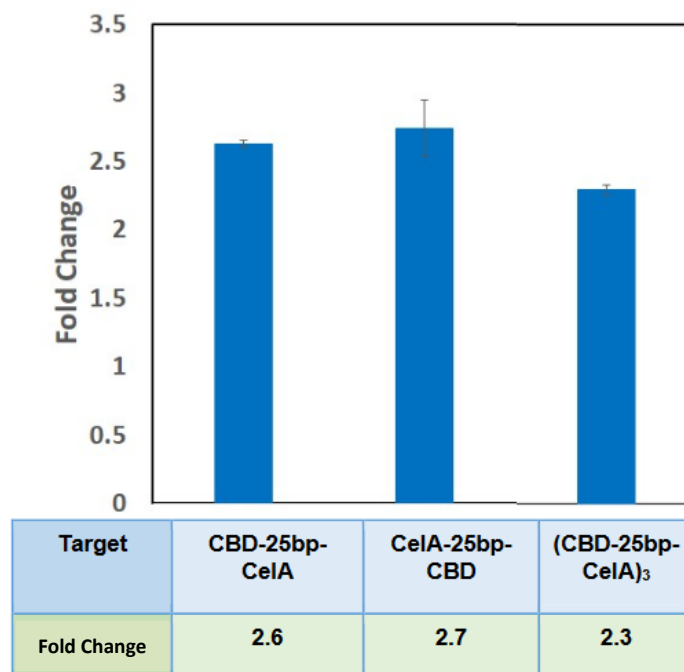


Figure 2.12 **Comparison of cellulose hydrolysis by different synthetic cellulosomes.** Different DNA templates were used to assemble SpdCas9-CBD and ST1dCas9-CelA.

2.4.6 Conditional dCas9 Mediated Cellulosome Formation

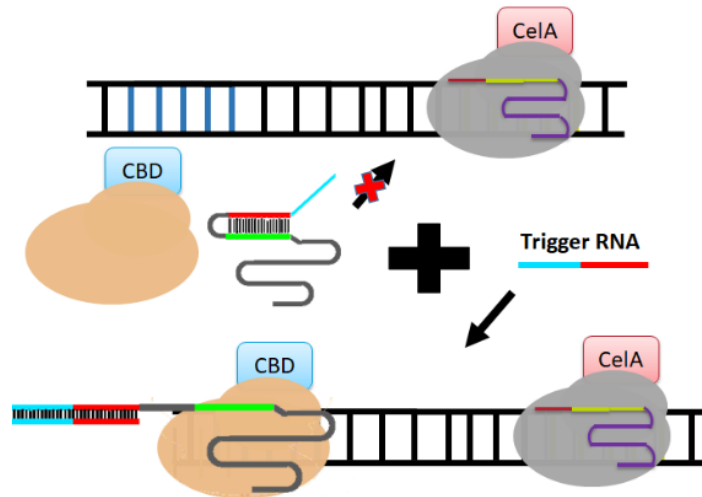


Figure 2.13 **Conditional assembly of dCas9 protein using a thgRNA.** Binding of SpdCas9-CBD is activated by the use of a RNA trigger to unblock the spacer regions of thgRNA by toehold-mediated strand displacement.

In plants, many native metabolons are dynamic in nature to provide adaptive changes in cellular responses.¹²⁰ While protein scaffolds have succeeded in improving enzymatic functions, they are restricted for dynamic modulation.¹²¹ Using a technology newly developed for conditional binding of dCas9,⁷² we next demonstrated the ability to provide dynamic modulation of cellulosome assembly. A toehold-gated sgRNA (thgRNA),¹²¹ with the spacer region initially blocked by a hairpin, is used. Addition of a RNA trigger unblocks the thgRNA by toehold-mediated strand displacement and reactivates the spacer for dCas9 binding to the DNA target (Figure 2.13).

To demonstrate dynamic cellulosome assembly, we designed a new thgRNA to modulate the conditional binding of SpdCas9-CBD. While the SpdCas9-CBD:sgRNA

complex restored full fluorescence to the beacon, incubation with the SpdCas9-CBD:thgRNA complex exhibited only minimal background, reaffirming the ability of thgRNA to inhibit SpdCas9 binding. Full activation was achieved with the addition of a RNA trigger to a level similar to that with the unblocked sgRNA (Figure 2.14). These data support the ability to gain conditional control over dCas9 binding using a simple RNA trigger.

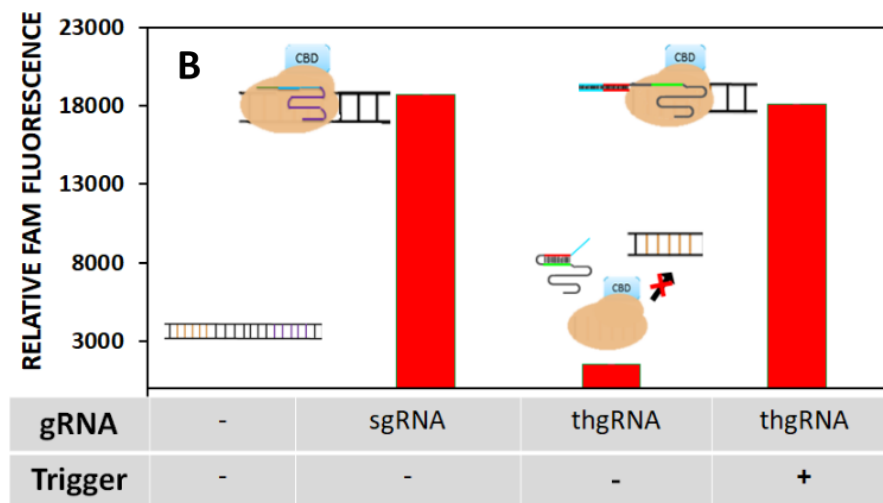


Figure 2.14 **Conditional Beacon Assay**. Conditional binding of SpdCas9-CBD using the Cas beacon assay. Binding was performed using either sgRNA or thgRNA. Activation was observed with the thgRNA only when a corresponding RNA trigger was added.

We next assessed the ability to control cellulosome assembly. In this design, St1dCas9-CelA is always bound to the DNA scaffold while SpCas9-CBD binding is dynamically modulated (Fig. 2.15). Simply by replacing the regular sgRNA with the thgRNA version, the level of reducing sugars reduced back to that of unassembled enzymes, consistent with the lack of binding by SpdCas9-CBD. Addition of the trigger

RNA restored not only SpdCas9-CBD binding but also the enhancement in cellulose hydrolysis. The ability to control enzyme clustering dynamically using a simple RNA input provides a powerful platform to modulate the efficiency of many enzyme cascade reactions. The strategy is particularly attractive for dynamic control of *in vivo* enzyme assembly, as even a full-length mRNA has been shown to be capable of activating SpdCas9 binding.¹²¹

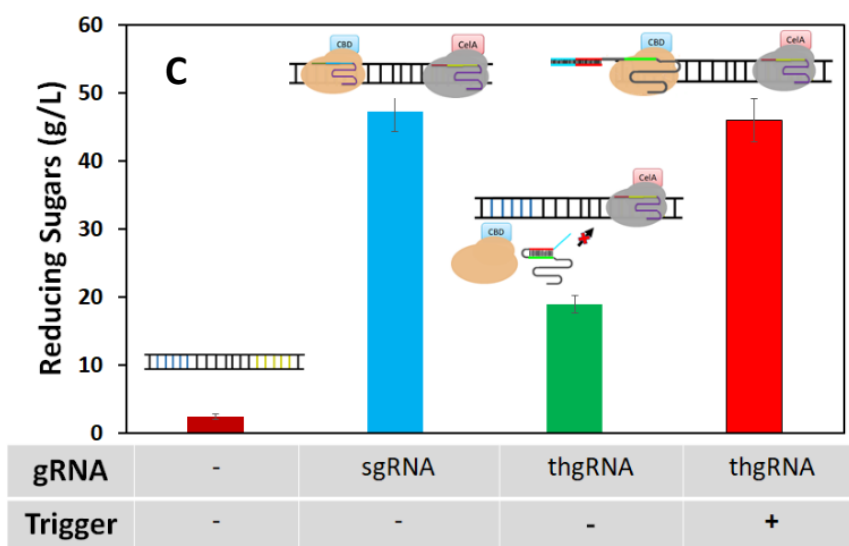


Figure 2.15 **Conditional assembly of a two-component cellulosome.** Cellulose hydrolysis by conditional cellulosome formation. The ability to enhance cellulose hydrolysis by activating SpdCas9-CBD/thgRNA binding by toehold-mediated strand displacement. Cellulosomes assembled using sgRNA and the activated thgRNA produced similar reducing sugar titers while the inactivated thgRNA produced less reducing sugars comparable to the unassembled enzymes.

2.5 Conclusions

In summary, we presented here a new approach to create enzyme cascades using dCas9-mediated assembly. By using orthogonal dCas9 pairs, site-specific assembly can be achieved. The high binding affinity afforded by dCas9 proteins enables easy control over enzyme stoichiometry, order, and spacing without sacrificing co-localization efficiency. The feasibility to modulate enzyme assembly dynamically using a simple RNA input should open up many *in vivo* opportunities for conditional activation of metabolic pathways.

Chapter 3

A MORE MODULAR SYNTHETIC METABOLON VIA dCAS9 GUIDED ASSEMBLY AND A POST TRANSLATIONAL LIGATION STRATEGY

3.1 Abstract

Synthetic scaffolding platforms have been a highly researched area for over a decade. Recently, the Wilfred Chen lab reported a new strategy to construct synthetic metabolons using dCas9-guided assembly. This dCas9 metabolon platform was applied towards construction of a two-component cellulosome that enhanced product titers by 2.6-fold. While this platform is modular owing to the ease of target synthesis, combinations of possible dCas9 fusion arrangements and the possibility to expand to other metabolic pathways, we strive to make the strategy even more modular. To increase the modularity and move towards an optimized system for a wide range of metabolic pathways, we employed an orthogonal post translation ligation strategy. This additional point of modularity not only increases the likelihood that other enzymatic pathways could easily be applied to this platform, it also increases the flexibility of the scaffold which has been shown to be important for many enzymatic pathways.

3.2 Introduction

Enzyme clustering and co localization has been shown in nature,^{18,19,21,108} as well as in synthetic forms²²⁻³⁴ to enhance metabolic pathway titers of a variety of products. This enzyme clustering, or metabolon formation, is a multienzyme complex

that allows for the direct passage of a product from one enzymatic reaction to the next enzyme in a metabolic pathway.⁷⁻¹⁰ This enzyme organization facilitates increased product titers due to its ability to limit the diffusion of transient toxic intermediates, prevents crosstalk between competitive pathways as well as concentrates rate limiting reactants which can drive unfavorable reactions.¹⁰

Our lab has recently reported a new dynamic scaffolding platform that utilizes the high binding-affinity dCas9 proteins to guide the assembly of synthetic metabolons.¹²² This dCas9-mediated assembly approach was used to demonstrate the formation of a synthetic cellulosome by fusing the endoglucanase CelA and a cellulose binding domain (CBD) to the C-terminus of two orthogonal dCas9 proteins. When scaffolded on the DNA target, a 2.6-fold higher level of reducing sugars was produced as compared to the free enzymes. While this synthetic platform offers high specificity, high binding affinity and modularity, it is often difficult to generate larger dCas9 fusion proteins due to their insolubility issues. To this end, a new modular platform is needed to decorate each orthogonal dCas9 protein with the enzyme of interest for site-specific assembly.

It is well documented that larger fusion proteins suffer from low expression rates, insolubility as well as protein misfolding and unwanted aggregation.³⁸ This is particularly problematic for dCas9 proteins due to their inherent large size (130-160 kDa). To circumvent this issue, we sought to express the dCas9 proteins and pathway enzymes separately, while directing them to associate together post translationally using the recently engineered, orthogonal Tag/Catcher systems from *Streptococcus pyogenes*⁸⁶ and *Streptococcus pneumoniae*.⁸⁷ These split proteins have been shown to form a stable isopeptide bond between a 13 amino acid tag and the ~15 kDa catcher

protein under a wide range of conditions, both *in vivo* and *in vitro*.^{86,87} Because of the orthogonality of the catcher systems, a different catcher protein can be fused to each pathway enzyme to direct its specific conjugation to the target dCas9 protein carrying the corresponding peptide tag. The full-length chimeric proteins can be guided by our engineered sgRNAs for site-specific assembly onto the DNA scaffold (Figure 3.1). The increased modularity of this new assembly framework was used to demonstrate the site-specific, co-localization of two cellulosomal components and the resulting impact on cellulose hydrolysis.

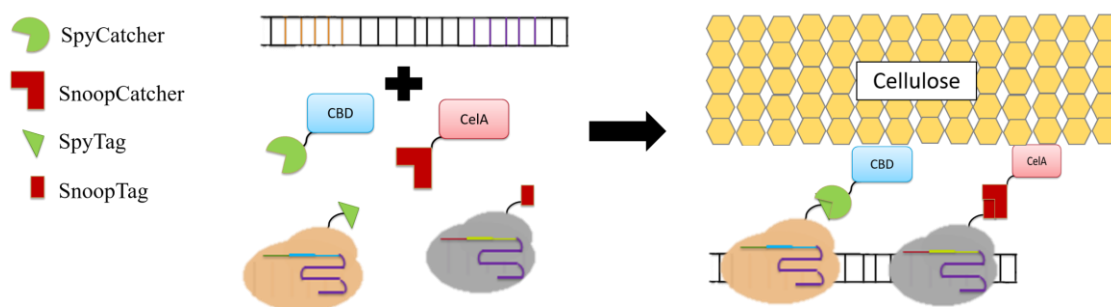


Figure 3.1 Formation of artificial cellulosome structure by full-length chimeric proteins guided by engineered sgRNA onto a double-stranded DNA template. Each protein component is expressed individually and form an isopeptide bond post translation to form the full-length chimeric protein. This chimeric dCas9 protein is guided to the DNA target by sgRNA which is capable of co-localization and cellulosome formation.

3.3 Materials and Methods

3.3.1 Materials

All primers were purchased from IDT (Coralville, IA). All ingredients for culturing media were purchased from Fisher Scientific (Pittsburgh, PA), all

ingredients for SDS-PAGE were purchased from BIO-RAD (Hercules, CA), all enzymes related to DNA manipulation and cloning were purchased from New England Biolabs (Ipswich, MA), and all other chemicals were purchased from Sigma Aldrich (St. Louis, MO).

3.3.2 Strains and Cloning

Escherichia coli NEB5 α (New England Biolabs, Ipswich, MA) [*fhuA2* Δ (*argF-lacZ*)*U169 phoA glnV44 Φ 80 Δ (lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*] was used as the host for all genetic manipulations and plasmid maintenance. *E. coli* strain BLR (DE3) (EMD Millipore, Madison, WI) [*F ompT hsdSB(r_B⁻ m_B⁻) gal dcm (DE3) Δ (*srl-recA*)306::*Tn10 (Tet^R)*] was used as the production host for all proteins.*

The nuclease-null Cas9 proteins *S. pyogenes* (SpdCas9), *S. thermophilus* (ST1dCas9) and *S. aureus* (SadCas9) were purchased from Addgene. The dCas9 protein sequences were PCR amplified, double digested, and inserted into pET24a using SpeI and bamHI. SpyTag and SnoopTag peptides were added into the vectors using the restriction sites bamHI and MfeI for the following constructs: pET24a SpdCas9-SpyTag-6xhis and St1dCas9-SnoopTag-6xhis. For the pathways enzymes, bamHI and XhoI restriction sites were used to add the catcher proteins to the vectors for the following constructions: pET24a StrepII-CBD-Spycatcher and pET24a SnoopCatcher-CelA-Flagtag-StrepII.

3.3.3 Protein Expression

All constructs were transformed into *E. coli* BLR [*F- ompT hsdSB (r-B m-B) gal dcm(*DE3) Δ (*srl-recA*)306::*Tn10(TetR)*; Novagen, Madison, WI] cells for protein

expression. Expression cultures were grown in Luria-Bertani (LB) media (10.0 g/L tryptone, 5.0 g/L yeast extract, 10.0 g/L NaCl) supplemented with 100 µg/mL ampicillin (dCas9-Tag fusions) or kanamycin (enzyme-catcher fusions).

3.3.3.1 Expression of dCas9-Tag Constructs

Overnight cultures were used to inoculate 30 mL of fresh LB to an initial OD₆₀₀ of ~0.05 and grow to mid-exponential phase (OD₆₀₀ ~0.75) at 37°C. Protein expression of SpdCas9-SpyTag-6xhis was induced with 200 µM isopropyl-β-thiogalactopyranoside (IPTG) and incubated at 20°C for ~16hrs. Expression of ST1dCas9-SnoopTag-6xhis protein fusion was optimal with induction of 100 µM IPTG and incubation at 37°C for 5 hours.

Induced cultures were pelleted using centrifugation at 3000xg for 10 minutes at 4°C, and the pellet was suspended in 1x TBS (Tris Buffered Saline) to an OD₆₀₀ = 30. The cell suspension was sonicated to release protein and clarified using centrifugation at 15,000xg for 15 min at 4°C. Soluble proteins were immediately purified using affinity chromatography.

3.3.3.2 Purification of dCas9-Tag Constructs

Soluble proteins were mixed with imidazole to a concentration of 5 mM. Meanwhile, gravity columns were packed with HisPur™ Ni-NTA Resin (ThermoFisher Scientific, Waltham, MA). The column was used as instructed using TBS + 10 mM imidazole as the equilibration buffer, TBS + 25 mM imidazole as the wash buffer, and TBS + 250 mM imidazole as the elution buffer. Eluates were dialyzed against TBS overnight. The cell lysates and the purified proteins were loaded onto a 7.5% SDS-PAGE gel and stained with coomassie blue for analysis.

3.3.3.3 Expression of Cellulosomal Enzyme-Catcher Constructs

Overnight cultures were used to inoculate 30 mL of fresh LB to an initial OD₆₀₀ of ~0.05 and grow to mid-exponential phase (OD₆₀₀ ~0.75) at 37°C. Protein expression of both StrepII-CBD-SpyCatcher and SnoopCatcher-CelA-FlagTag-StrepII proteins were induced with 200 μM isopropyl-β-thiogalactopyranoside (IPTG) and incubated at 20°C for ~16 hours.

Induced cultures were pelleted using centrifugation at 3000xg for 10 minutes at 4°C, and the pellet was suspended in 1x TBS (Tris Buffered Saline) to an OD₆₀₀ = 30. The cell suspension was sonicated to release protein and clarified using centrifugation at 15,000xg for 15 min at 4°C. Soluble proteins were immediately purified using affinity chromatography.

3.3.3.4 Purification of Cellulosomal Enzyme-Catcher Constructs

Soluble StrepII-CBD-Spycatcher proteins were added to gravity columns packed with Strep-Tactin® (IBA, Germany) The column was used as instructed washing the column with the proprietary wash buffer (buffer W) and eluting with 2.5 mM desthiobiotin (Buffer E). Eluates were dialyzed against TBS overnight. The cell lysates and the purified proteins were loaded onto a 7.5% SDS-PAGE gel and stained with coomassie blue for analysis.

StrepTag purification interfered with the SnoopCatcher-SnoopTag reaction post purification. Due to this conflict, SnoopCatcher-CelA-FlagTag-StrepII was purified by incubating soluble lysate in a 60°C water bath for 15 minutes followed by centrifugation at 15,000xg for 15 min at 4°C to separate purified SnoopCatcher-CelA-FlagTag-StrepII proteins from aggregated proteins. The cell lysates and the purified

proteins were loaded onto a 12% SDS-PAGE gel and stained with coomassie blue for analysis.

3.3.4 RNA *In Vitro* Transcription and Purification

All RNAs were transcribed *in vitro* via HiScribe T7 Quick High Yield RNA Synthesis Kit (New England BioLabs, Inc., Ipswich, MA, USA) and purified via phenol-chloroform extraction and ethanol precipitation. The purity, quality and quantity of RNA was analyzing by spectrophotometry using a NanoDrop 2000 UV-Vis spectrophotometer (Thermo Fisher Scientific, Inc., Waltham, MA, USA). Then subsequently using a denaturing urea PAGE electrophoresis with a 6% polyacrylamide gel containing 8 M urea.

3.3.5 Post Translational Ligation Reactions

3.3.5.1 Tag/Catcher Lysate Reactions

Soluble lysates were mixed together at a 1:1 ratio, SpdCas9-SpyTag-6xhis and StrepII-CBD-SpyCatcher as well as ST1dCas9-SnoopTag-6xhis and SnoopCatcher-CelA-FlagTag-StrepII. The improper Tag/Catcher pairs were also mixed to confirm orthogonality. Reactions were incubated overnight at 4°C with rotation. Overnight reaction concentrations were quantified by measuring A595 after incubation with Bradford reagent and were loaded onto a 7.5% SDS-PAGE gel and stained with coomassie blue for analysis.

3.3.5.2 Tag/Catcher Purified Protein Reactions

Purified proteins were mixed together at a total concentration of 24 μ M in a 1:1 ratio, SpdCas9-SpyTag-6xhis and StrepII-CBD-SpyCatcher as well as ST1dCas9-

SnoopTag-6xhis and SnoopCatcher-CelA-FlagTag-StrepII. Reactions were incubated overnight at 4°C with rotation. Overnight reaction concentrations were quantified by measuring A595 after incubation with Bradford reagent and were loaded onto a 7.5% SDS-PAGE gel and stained with coomassie blue for analysis.

3.3.5.3 Tag/Catcher Reaction Time Course

Purified proteins were mixed together at a total concentration of 24 μ M in a 1:1 ratio, SpdCas9-SpyTag-6xhis and StrepII-CBD-SpyCatcher as well as ST1dCas9-SnoopTag-6xhis and SnoopCatcher-CelA-FlagTag-StrepII. Reactions were incubated at room temperature and samples were taken at time 0, .5, 1 and 2 hours. Samples were quenched by addition of SDS-PAGE loading dye and boiling. Samples were loaded onto a 7.5% SDS-PAGE gel and stained with coomassie blue for analysis.

3.3.6 Electrophoretic Mobility Shift Assay

Binding of ST1dCas9-SnoopTag with a large truncation was qualitatively verified using electrophoretic mobility shift experiments carried out in 4.5% non-denaturing acrylamide gels. Final assay mixtures contained 10 nM hybridized beacon target, 100 nM dCas9 protein and 100 nM sgRNA. The fully hybridized beacon DNA only, as well as the beacon DNA bound by the dCas9:sgRNA complex were loaded and run for 90 minutes at a constant 90V. The gel was then imaged using an Amersham™ Typhoon™ Biomolecular Imager (GE Healthcare Bio-Sciences Corporation, Marlborough, MA).

3.3.7 Beacon Assay

All DNA probes were prepared using unmodified and fluorophore or quencher labeled oligonucleotides synthesized and purified by the commercial vendor Integrated

DNA Technologies (Coralville, IA, USA). Target Cas9 beacon DNA shown in Table 2.1 were created by mixing equimolar amounts of complementary DNA strands at a final concentration of 1 μ M in a beacon hybridization buffer (40 mM Tris, pH 7.9, 100 mM NaCl); heating for 5 minutes at 90°C then cooled to 12°C at a rate of 0.1 /s using a S1000 Thermal Cycler (Bio-Rad Laboratories, INC., Hercules, CA, USA). All beacon binding assays were carried out in beacon binding buffer (20 mM Tris-HCl, 120 mM NaCl, 5% v/v glycerol, 0.1 mM DTT, 1 mM MgCl₂, 0.02% v/v Tween 20, pH 7.9) and fluorescence measured using a Synergy H4 Hybrid microplate reader (BioTek Instruments, Inc., Winooski, VT, USA). Final assay mixtures contained 1 nM hybridized beacon target DNA, 10 nM dCas9 protein(s), and 10 nM sgRNA(s). Excitation and emission wavelengths were dictated by the fluorophore in use for each target (A: FAM, ex: 498nm, em:520 nm; B: Cy5 ex: 648nm, em: 668nm).

3.3.8 Synthetic Cellulosome Assembly and Reducing Sugar Assay

Phosphoric acid-swollen cellulose (PASC) was prepared as described previously.¹¹⁶

3.3.8.1 Step-wise Cellulosome Assembly and Reducing Sugar Assay

Tag/Catcher reactions were performed in separate tubes at total concentrations of 24 μ M in a 1:1 ratio, SpdCas9-SpyTag-6xhis and StrepII-CBD-SpyCatcher as well as ST1dCas9-SnoopTag-6xhis and SnoopCatcher-CelA-FlagTag-StrepII. Reactions were incubated overnight at 4°C with rotation or at room temperature for 1-2 hours. Full length chimeric proteins were then diluted added together to final concentrations of 20 nM each and added to 20 nM sgRNA(s) with or without 20 nM target DNA in the presence of PASC. Synthetic cellulosome assembly was assayed at 30°C in 20 nM

Tris-HCl buffer pH 6.0 with shaking. Samples were collected periodically over a 40 hour time period and immediately mixed with 0.25 mL DNS reagent (10 g/L dinitrosalicylic acid, 10 g/L sodium hydroxide, 2 g/L phenol, 0.5 g/L sodium sulfite) and incubated at 95°C for 10 minutes. After incubation, 0.5 mL of 40% Rochelle salts were added to fix the color before measuring the absorbance using a spectrophotometer at 575 nm.

3.3.8.2 One Pot Cellulosome Assembly and Reducing Sugar Assay

Tag/Catcher reactions were performed in the same test tube each at a total concentration of 24 μ M in a 1:1 ratio, SpdCas9-SpyTag-6xhis and StrepII-CBD-SpyCatcher as well as St1dCas9-SnoopTag-6xhis and SnoopCatcher-CelA-FlagTag-StrepII. Reactions were incubated at room temperature for 1-2 hours and diluted to final concentrations of 20 nM each and added to 20 nM sgRNA(s) with or without 20 nM target DNA in the presence of PASC. Synthetic cellulosome assembly was assayed at 30°C in 20 nM Tris-HCl buffer pH 6.0 with shaking. Samples were collected periodically over a 40-hour time period and immediately mixed with 0.25 mL DNS reagent (10 g/L dinitrosalicylic acid, 10 g/L sodium hydroxide, 2 g/L phenol, 0.5 g/L sodium sulfite) and incubated at 95°C for 10 minutes. After incubation, 0.5 mL of 40% Rochelle salts were added to fix the color before measuring the absorbance using a spectrophotometer at 575 nm.

3.4 Results and Discussion.

3.4.1 Design of dCas9 proteins for orthogonal bioconjugation

To create a modular approach to decorate dCas9 proteins with any desirable enzyme of interest, we further engineered SpdCas9 and St1dCas9 by inserting either a

SpyTag or SnoopTag to the C-terminus for orthogonal conjugation. This strategy not only minimizes improper folding, aggregation and truncations because of direct expression of larger dCas9 fusion proteins,¹²⁴ it also offers a more modular synthetic scaffolding platform capable of being utilized for endless metabolic pathways with little optimization.

To investigate whether insertion of the smaller SpyTag and SnoopTag to the dCas9 proteins could improve the overall protein yield, we first compared expression of SpdCas9-SpyTag and St1dCas9-SnoopTag to their corresponding, larger dCas9 fusions (SpdCas9-CBD and ST1dCas9-SmBit). While full-length expression was similar for both SpdCas9 proteins, SpdCas9-CBD had a larger fraction of truncation products as well as more misfolded insoluble proteins (Fig. 3.2A). In contrast, swapping out SmBit with a SnoopTag increased expression by almost two-fold, again highlighting the benefit of using the smaller Tag fusions with dCas9 (Figure 3.2B).

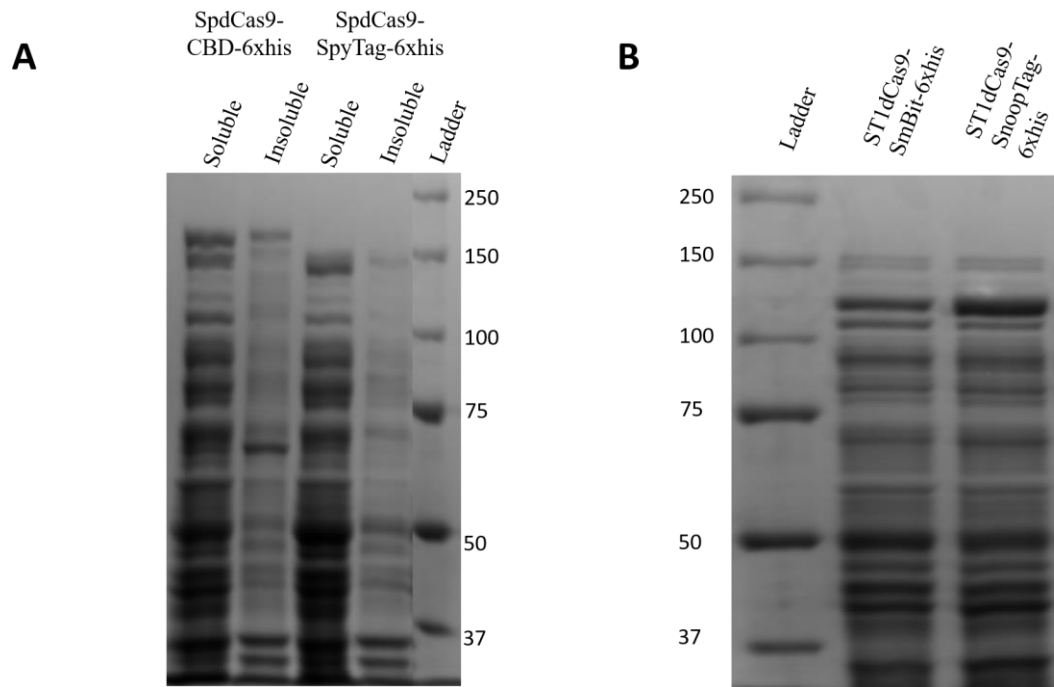


Figure 3.2 More modular dCas9 proteins improve expression. (A) SpdCas9-CBD-6xhis suffers from low expression, large truncation as well as proteins being lost in the insoluble form. Switching out the large CBD protein for the small peptide SpyTag increased protein expression by 4.8% and relieved the loss of protein to the insoluble fraction. (B) ST1dCas9-SmBit-6xhis showed relatively low rates of expression while directly fusing SnoopTag to ST1dCas9 caused a 12.3% increase in protein production.

3.4.2 Formation of dCas9-enzyme conjugates by post translational ligation reactions

To ensure that the fused Spytag and SnoopTag could retain their full functionality, we investigated the ability of our dCas9-Tag fusions to react with the appropriate catcher-enzyme fusions either with purified proteins or directly in cell lysates. We also investigated the rate at which the reaction occurs in order for this platform to be useful for in vivo metabolon formation.

3.4.2.1 Functionality and Orthogonality of the SpyCatcher and SnoopCatcher Reactions

We first tested the ability of purified SpdCas9-SpyTag and ST1dCas9-SnoopTag to form an isopeptide bond with their corresponding protein partners. Both dCas9 proteins were purified using the C-terminus his6 tag and mixed with the corresponding Catcher fusion partner (either CBD-SpyCatcher or SnoopCatcher-CelA) in a 1:1 ratio, Figure 3.3. After overnight conjugation, the reaction products were analyzed by gel electrophoresis. For the SpyTag/SpyCatcher pair, a clear product band around 192 kDa was detected supporting that SpdCas9-SpyTag and CBD-SpyCatcher are capable of reacting post purification. This reaction reached nearly 100% completion with almost no SpdCas9-SpyTag leftover. Similarly, ligation was observed for the SnoopTag/SnoopCatcher pair. However, the conjugation efficiency was lower, as reported in literature, with only 70% of the purified ST1dCas9-SnoopTag reacted.⁸⁸

It should be noted that several other product bands were also observed for the SnoopTag/SnoopCatcher pair due to ligation with purified ST1dCas9-SnoopTag truncated at the N-terminus. To ensure these truncation products within the dCas9 region are not capable of DNA binding, we performed a Cas beacon binding assay. As shown in Figure 3.4 there is only one product band formed indicating only the full-length portion of the protein retains the ability to bind DNA while the truncation does not, thus the N-terminal truncation will not hinder scaffolding formation.

We next tested the ligation reactions directly in cell lysate. As shown in Figure 3.5, a larger product band was observed when the appropriate components were mixed in an equal 1:1 ratio. When components from different species were mixed together, no product band was detected in either case (Figure 3.5). This result indicates that

these Tag/Catcher systems are highly specific and can be utilized for orthogonal conjugation of pathway enzymes to different dCas9 proteins. Since both Catcher reactions can be executed in complex cell lysates, we are confident that these orthogonal reactions can be employed even for orthogonal enzyme conjugation *in vivo*.

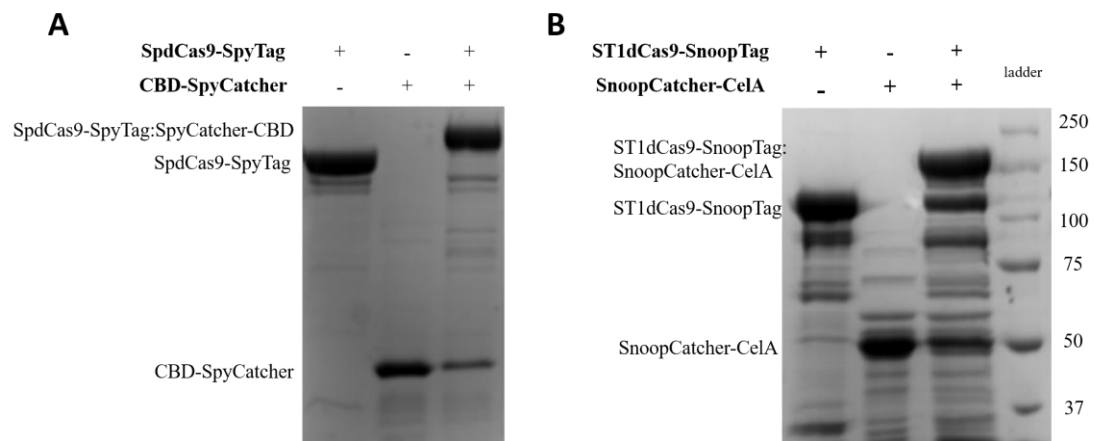


Figure 3.3 **Purified proteins spontaneously form isopeptide** bonds. (A) purified SpdCas9-SpyTag and CBD-SpyCatcher forms a product band at 192.5 kDa with near 100% efficiency. (B) St1dCas9-SnoopTag and SnoopCatcher-CelA form a product band at 188.3 kDa post protein purification. The efficiency of SnoopTag/Catcher to form a covalent bond is less than that of SpyTag/SpyCatcher which has been reported before in literature.

DNA	+	+
ST1dCas9-SnoopTag	-	+
St1gRNA	-	+



Figure 3.4 **ST1dCas9-SnoopTag truncation is not capable of binding DNA** EMSA analysis of ST1dCas9-SnoopTag:St1gRNA binding to our synthetic DNA beacon shows only one product band indicating the N-terminal truncation does not retain binding capabilities.

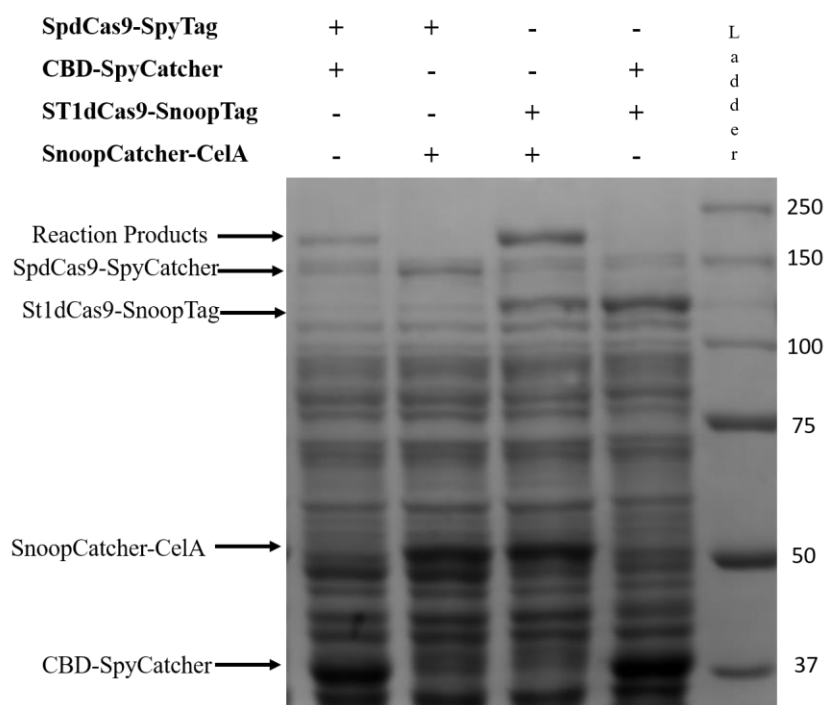


Figure 3.5 **In Vivo Orthogonal Tag/Catcher Reactions.** SpdCas9-SpyTag is capable of forming a covalent bond with CBD-SpyCatcher in cell lysate as indicated by the product formation at 192.5 kDa. ST1dCas9-SnoopTag is capable of forming a spontaneous isopeptide bond with SnoopCatcher-CelA in cell lysate, forming a product band of 188.3 kDa. When the components from different species are mixed, there is no product formation indicating the pairs are orthogonal.

3.4.2.2 Time Course of Post Translation Ligation Reaction

To evaluate whether this modular scaffolding platform can be utilized for rapid assembly, the reaction kinetics of the Tag/Catcher systems were investigated by mixing purified proteins together and allowing them to react at room temperature for

various times before quenching the reaction by boiling in SDS buffer. As shown in Figure 3.6A, the SpdCas9-SpyTag and CBD-SpyCatcher reaction began rapidly with a product band detected immediately after mixing. The reaction went to near 100% completion by 30 min, in line with previous studies.⁸⁶ Although the SnoopCatcher system is slightly less efficient, requiring about 2 h to achieve 70% ligation, both reactions are sufficiently fast to provide quick enzyme conjugation for *in vivo* cascading (Figure 3.6B).

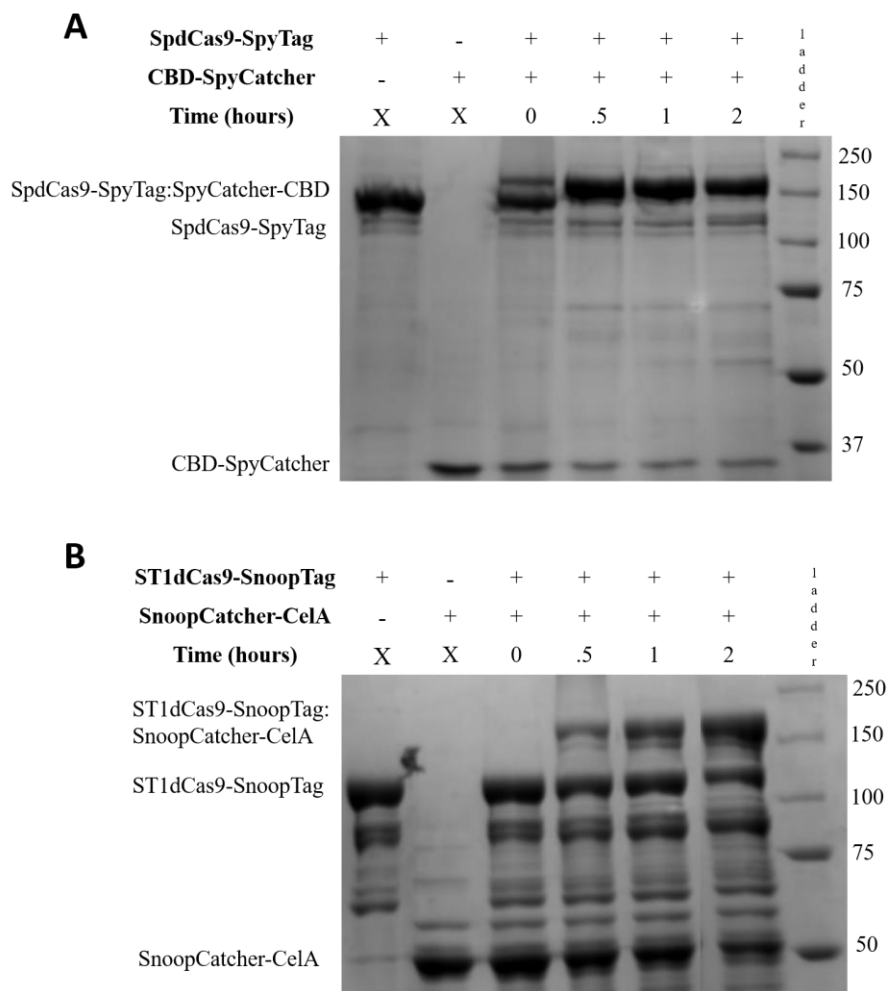


Figure 3.6 **Time Course Tag/Catcher Reactions.** (A) SpdCas9-SpyTag began reacting with CBD-SpyCatcher immediately upon mixing and was completely reacted by 30 minutes. (B) St1dCas9-SnoopTag was completely reacted with SnoopCatcher-CelA at 2 hours post mixing.

3.4.3 Stepwise Tag/Catcher Reaction for Mini Cellulosome Formation

We next exploited the orthogonal Catcher reactions for the dCas9-guided assembly of a two-component cellulosome in either a stepwise or one pot fashion. For the stepwise assembly, we first conjugated CBD-SpyCatcher to SpdCas9-SpyTag and SnoopCatcher-CelA to ST1dCas9-SnoopTag in two separate reactions. The ligated

products were mixed at a 1:1 ratio before incubating with their cognate sgRNAs and target DNA. As shown in Figure 3.7, the binding ability of both SpdCas9-SpyTag and ST1dCas9-SnoopTag was not hindered by reaction with the appropriate catcher protein fusion. Both the conjugated and unconjugated dCas9 proteins caused a similar level of increase in fluorescence upon binding to the dual Cas beacon. These dual-binding results are better than those previously reported using the direct dCas9 fusion proteins possibly due to the overall increase in protein quality achieved.

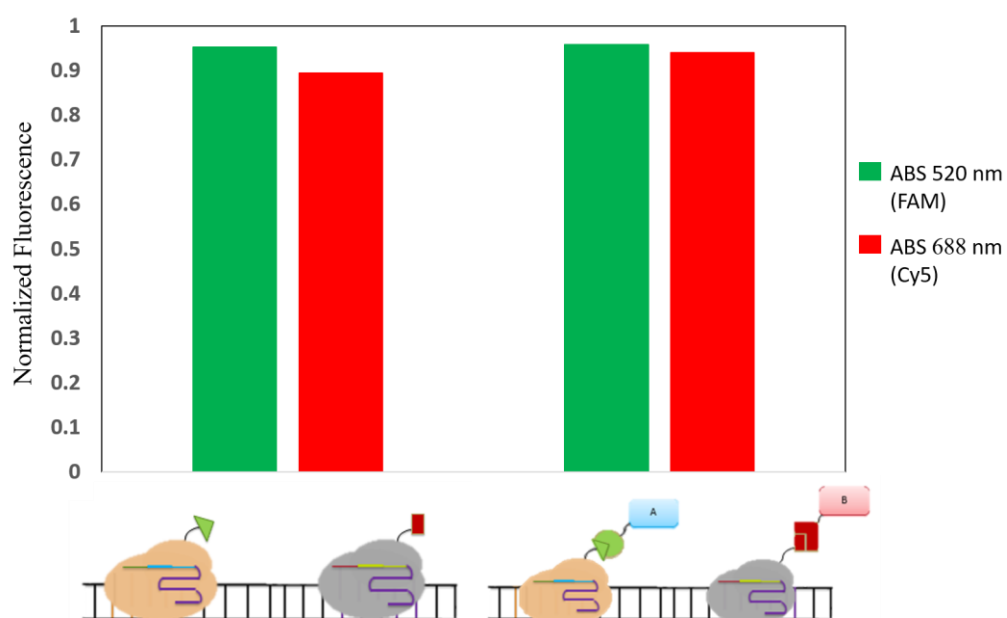


Figure 3.7 Binding Efficiency of dCas9-Tag proteins pre and post ligation reactions in a stepwise fashion. Left bars indicate SpdCas9-SpyTag and ST1dCas9-SnoopTag are capable of binding the dual beacon with high efficiency. Right bars show similar levels of dual dCas9-Tag binding after the ligation reactions have occurred in separate tubes.

Having confirmed that the dCas9-Tag fusions are capable of binding in a dual fashion to a DNA target even after ligation, we next tested whether the resulting cellulosome assembled in this stepwise fashion could achieve similar increase in cellulose hydrolysis. The assembled enzyme-DNA complex was mixed with PACS and the overall reducing sugar level was compared with the free enzymes. As shown in Figure 3.8, DNA alone, as well as DNA bound by SpdCas9-SpyTag and St1dCas9-SnoopTag only, showed background levels of reducing sugar production. While free enzymes (either direct fusions or ligated dCas9-enzyme conjugates) produced a higher level of reducing sugar, the overall PACS hydrolysis was improved by 2.8-fold when the dCas9-enzyme conjugates are co-localized onto a DNA scaffold. This is slightly better than using the direct dCas9 fusion enzymes possibly due to the improved protein quality as well as the extended enzyme accessibility granted by the Tag/Catcher system.

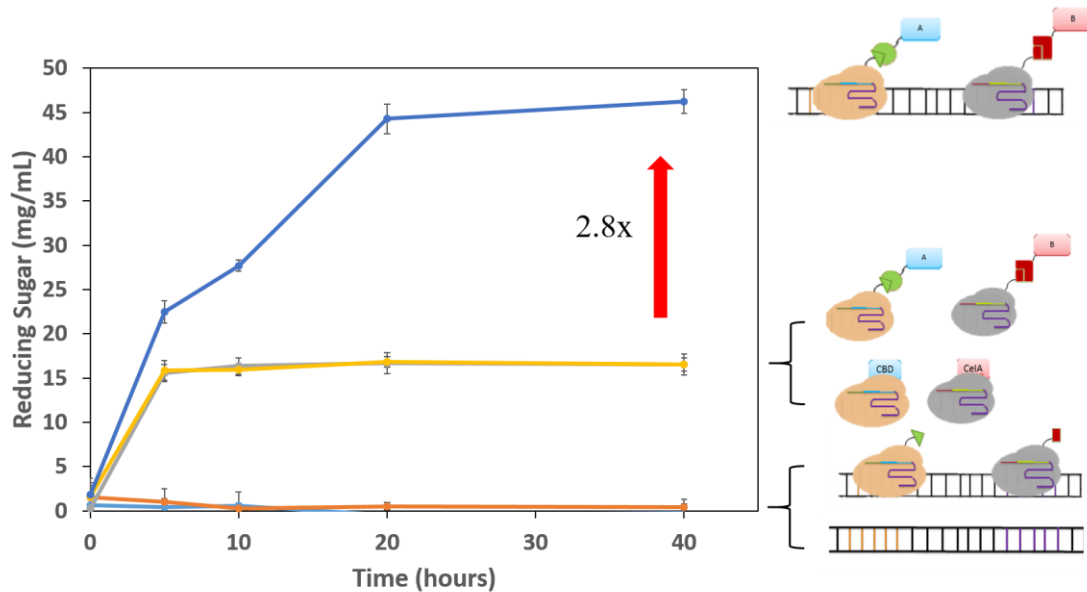


Figure 3.8 Step Wise Cellulosome Formation Enhanced Reducing Sugar Titters. DNA only (light blue) as well as SpdCas9-SpyTag and ST1dCas9-SnoopTag co-localized on the DNA (orange) show background levels of reducing sugar production. Free floating pathways enzymes in direct fusion form (green) as well as in post ligation form of SpdCas9-SpyTag:CBD-SpyCatcher and ST1dCas9-SnoopTag:SnoopCatcher-CelA show some reducing sugar production. However when SpdCas9-SpyTag:CBD-SpyCatcher and ST1dCas9-SnoopTag:SnoopCatcher-CelA (dark blue) are co-localized on the DNA target we see a 2.8-fold increase in reducing sugar production.

3.4.4 One Pot Tag/Catcher Reaction for Mini Cellulosome Formation

While the stepwise procedure is useful for *in vitro* enzyme assembly, the orthogonality of the Catcher/Tag pairs allows the one-pot conjugation and DNA binding for *in vivo* metabolon formation. To assess this feasibility, we mixed the dCas9 proteins, Catcher-tagged proteins, and sgRNAs all together. After two hours, reaction mixture was analyzed on a SDS-PAGE gel. From Figure 3.9, both conjugation reactions reached near completion, with a small fraction of ST1dCas9 left as previously shown.

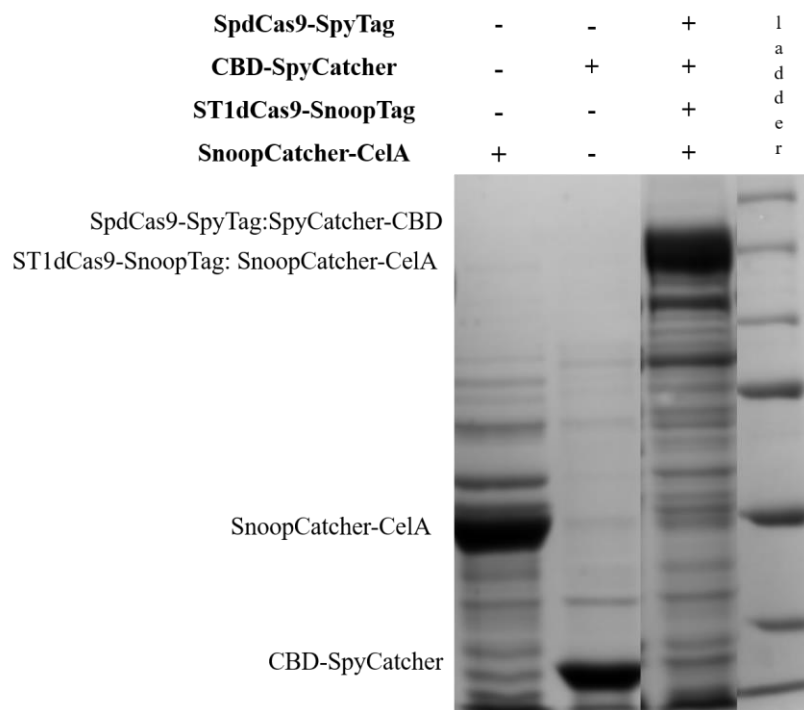


Figure 3.9 **SDS-PAGE of One Pot Reaction of both dCas9-Tag and Catcher-Enzyme fusions.** Both dCas9-Tag proteins, SpdCas9-SpyTag and ST1dCas9-SnoopTag, reaction with their respective Catcher fusions, CBD-SpyCatcher and SnoopCatcher-CelA to form product bands of 192.5 and 188.3 kDa respectively.

Having confirmed the ability for one-pot ligation via SDS-PAGE, we assessed the DNA-binding capability of the ligated products using the dual beacon assay. Similar to stepwise ligation, both conjugated dCas9 complexes were capable of co-localizing onto the same DNA target as indicated by an increase in both FAM and Cy5 fluorescence upon binding (Figure 3.10). Due to efficient binding, we next prepared a synthetic cellulosome structure using the same one-pot reaction and investigated the improvement in reducing sugar titers. While free enzymes showed higher levels of reducing sugar produced than DNA alone, it is not until the dCas9 conjugates were co-

localized onto the DNA scaffold that a much higher product titer was observed. In fact, the same 2.8-fold increase in reducing sugar titer (Figure 3.11) was detected using the one-pot assembly approach when compared to stepwise assembly, indicating this platform could be used for both *in vivo* applications as well as *in vitro* applications.

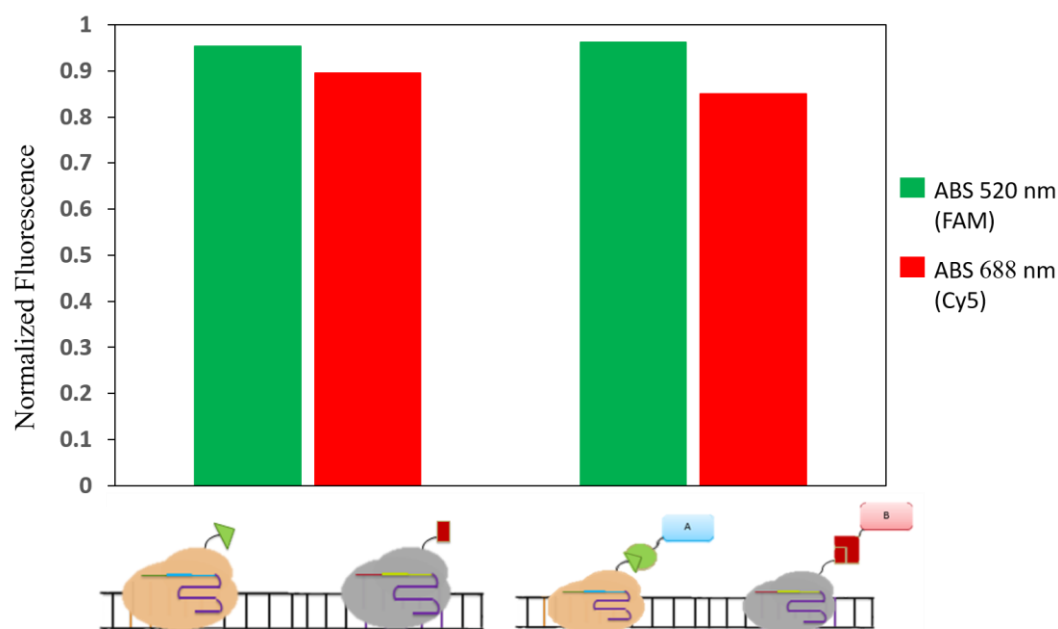


Figure 3.10 **Binding Efficiency of dCas9-Tag proteins pre and post ligation reactions in one pot reaction.** Left bars indicate SpdCas9-SpyTag and ST1dCas9-SnoopTag are capable of binding the dual beacon with high efficiency. Right bars show similar levels of dual dCas9-Tag binding after the ligation reactions have occurred in a single tube.

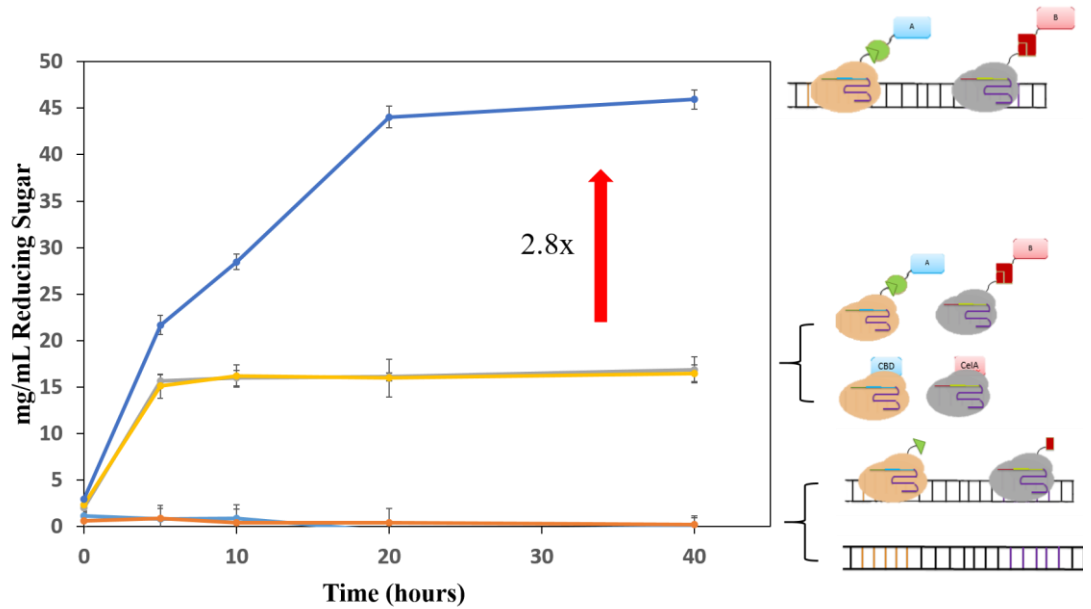


Figure 3.11 One Pot Reaction Cellulosome Formation Enhanced Reducing Sugar Titers. DNA only (light blue) as well as SpdCas9-SpyTag and ST1dCas9-SnoopTag co-localized on the DNA (orange) show background levels of reducing sugar production. Free floating pathway enzymes in direct fusion form (green) as well as in post ligation form of SpdCas9-SpyTag: CBD-SpyCatcher and ST1dCas9-SnoopTag: SnoopCatcher-CelA show some reducing sugar production. However when SpdCas9-SpyTag: CBD-SpyCatcher and ST1dCas9-SnoopTag: SnoopCatcher-CelA (dark blue) are colocalized on the DNA target we see a 2.8-fold increase in reducing sugar production.

3.5 Conclusions

In conclusion, we demonstrated here a modular approach to create enzyme cascades using dCas9-mediated assembly and an orthogonal post-translation ligation strategy. By expressing the dCas9 proteins separately from the pathway enzymes and ligating them together with orthogonal Catcher reactions, we were able to increase protein expression and the overall enzyme assembly. This platform retains the high binding affinity afforded by dCas9 proteins for easy control over enzyme

stoichiometry, order, and spacing without sacrificing co-localization efficiency and adds an additional layer of modularity and flexibility that has further increased the overall scaffolding benefits as shown by the improved fold enhancement of reducing sugar production. The ability to execute this platform in either a stepwise or one-pot fashion allows flexibility of applications for both *in vivo* and *in vitro* assembly of a wide array of enzyme cascades with minimal engineering.

Chapter 4

A MODULAR SYNTHETIC METABOLON VIA SELF-ASSEMBLING PROTEIN NANOCAGES

4.1 Abstract

The wide variety of enzymatic pathways that can benefit from enzyme scaffolding is astronomical however it is unlikely that one synthetic scaffolding system would optimize all pathways identically. Some metabolic pathways, such as mevalonate, show the best results when enzyme location and stoichiometry are directed while others, such as cellulosomal enzymes, have shown optimized values with randomized co-localization. Here we discuss the co-localization of the endoglucanase CelA and cellulose binding domain, CBD, onto the self-assembling nanocages, E2 and Hepatitis B Virus (HBV) due to their differences in size and molecular makeup. While both protein nanocages enhanced reducing sugar production *in vitro*, the increased rigidity offered by the nanocages may not be optimal for cellulose hydrolysis.

4.2 Introduction

For decades, researchers have investigated alternative methods to co-localize metabolic enzymes using a synthetic scaffold to enhance the production of a variety of products.¹⁻³ Some approaches focused on protein-based scaffolds,²²⁻³⁷ while others utilized nucleic acid-based,³⁹⁻⁴⁵ and nanoparticle-based scaffolding systems.⁸⁹⁻⁹³ Although majority of these scaffolding attempts have shown improved product titers,

the array of success varies from enhancements of 77-fold,¹¹ to less than 1-fold.^{89–93} Even when the protein-scaffolding platform with the 77-fold increase in mevalonate production,¹¹ was applied to the production of glucaric acid, only a 5-fold improvement was observed, illustrating the need for multiple scaffolding techniques for each target metabolic pathway.³³

Protein nanocages have gained attention as a promising scaffold because of their ability to self-assemble into nanostructures with defined sizes and shapes. They are ideal scaffolds for protein assembly due to their high stability under a wide range of conditions including pH, temperature and salt concentrations. Although the exterior of nanocages can be modified with a variety of enzymes using chemical conjugations, these methods are costly and can cause unwanted damage to the nanocage or enzymes themselves. Our group has recently reported the use of a benign post-translation ligation strategy based on the SpyTag/SpyCatcher pair to conjugate proteins onto nanocages of interest.⁸⁷ Because multiple functional proteins can be simultaneously decorated onto the same protein nanocage, this strategy is ideally suited to generate complex enzyme cascades in a highly modular fashion.

In this chapter, we further exploited this robust ligation strategy to decorate nanocages into functional cellulosome structures for enhanced cellulose hydrolysis. We chose to focus on two well-characterized nanocages: the 25 nm E2 nanocage derived from the pyruvate dehydrogenase complex of *Bacillus stearothermophilus* and the larger 34 nm nanocage derived from the hepatitis B virus (HBV) capsid. These two nanocages were chosen not only due to their differences in size but also their unique ability to provide either 60 or 240 total surface decorations, allowing us to investigate the effects of both size and decoration density on enhancement.

4.3 Materials and Methods

4.3.1 Materials

All primers were purchased from IDT (Coralville, IA). All ingredients for culturing media were purchased from Fisher Scientific (Pittsburgh, PA), all ingredients for SDS-PAGE were purchased from BIO-RAD (Hercules, CA), all enzymes related to DNA manipulation and cloning were purchased from New England Biolabs (Ipswich, MA), and all other chemicals were purchased from Sigma Aldrich (St. Louis, MO).

4.3.2 Strains and Cloning

Escherichia coli NEB5 α (New England Biolabs, Ipswich, MA) [*fhuA2* Δ (*argF-lacZ*)U169 *phoA glnV44* Φ 80 Δ (*lacZ*)M15 *gyrA96 recA1 relA1 endA1 thi-1 hsdR17*] was used as the host for all genetic manipulations and plasmid maintenance. *E. coli* strain BLR (DE3) (EMD Millipore, Madison, WI) [*F* *ompT hsdS_B(r_B⁻ m_B⁻) gal dcm* (DE3) Δ (*srl-recA*)306::*Tn10* (*Tet^R*)] was used as the production host for all proteins.

Pathways enzymes, CelA and CBD were directly fused to SpyCatcher using *bam*HI and *Xho*I restriction sites for the following constructs: pET24a StrepII-CBD-Spycatcher and pET24a StrepII-CelA-SpyCatcher. Nanocage plasmids, pET11a SpyTag-E2 (158) and pET24a HBV-SpyTag were gifts from Chen Lab group members Dr. Andrew Swartz and Daniel Yur respectively.

4.3.3 Protein Expression

All constructs were transformed into *E. coli* BLR [*F*- *ompT hsdS_B* (r-B m-B) *gal dcm*(DE3) Δ (*srl-recA*)306::*Tn10*(*Tet^R*); Novagen, Madison, WI] cells for protein

expression. Expression cultures were grown in Luria-Bertani (LB) media (10.0 g/L tryptone, 5.0 g/L yeast extract, 10.0 g/L NaCl) supplemented with 100 $\mu\text{g}/\text{mL}$ ampicillin (nanocages) or kanamycin (pathway enzymes). Overnight cultures were used to inoculate 30 mL of fresh LB to an initial OD_{600} of ~ 0.05 and grow to mid-exponential phase ($\text{OD}_{600} \sim 0.75$) at 37°C . Protein expression of all fusion proteins were induced with 200 μM isopropyl- β -thiogalactopyranoside (IPTG) and incubated at 20°C for ~ 16 hours.

Induced cultures were pelleted using centrifugation at $3000\times g$ for 10 minutes at 4°C , and the pellet was suspended in 1x PBS (Phosphate Buffered Saline) to an $\text{OD}_{600} = 30$. The cell suspension was sonicated to release protein and clarified using centrifugation at $15,000\times g$ for 15 minutes at 4°C . Soluble proteins were immediately purified using various techniques.

4.3.4 Protein Purification

4.3.4.1 E2 Purification

E2-SpyTag proteins were partially purified by incubating at 70°C for 10 min and centrifugation at $15,000\times g$ for 15 min to isolate the soluble proteins. The soluble E2 sample was filtered through a $0.8/0.2\mu\text{m}$ Supor Acrodisc syringe filter. E2 nanocage assembly was confirmed via transmission electron microscopy and SDS-PAGE gel analysis.

4.3.4.2 HBV Purification

To partially purify HBV-SpyTag proteins, the lysate was incubated at 65°C for 30 minutes at pH 8, followed immediately by centrifugation at max speed for 15 minutes at 4°C . The supernatant was removed and ammonium sulfate was added until

40% saturation was reached. The salt solution was incubated on ice for 30 minutes before being centrifuged again at max speed, for 15 minutes 4°C. The supernatant was again removed and the pellet was resuspended in 1x PBS. HBV nanocage assembly was confirmed via transmission electron microscopy and SDS-PAGE gel analysis.

4.3.4.3 Purification of Cellulose Enzyme SpyCatcher Fusions

Soluble StrepII-CBD-Spycatcher and StrepII-CelA-SpyCatcher proteins were added to gravity columns packed with Strep-Tactin® (IBA, Germany). The column was used as instructed washing the column with the proprietary wash buffer (buffer W) and eluting with 2.5 mM desthiobiotin (Buffer E). Eluates were dialyzed against 1x PBS overnight. The cell lysates and the purified proteins were loaded onto a 7.5% SDS-PAGE gel and stained with coomassie blue for analysis.

4.3.5 Protein Nanocage Decoration

Purified proteins were mixed together at a total concentration of 24 µM in different ratios: for the E2 nanocage we tested a 1:1 ratio as well as a 5:1 ratio of CelA:CBD decorated on exterior of the capsid. For the HBV nanocage we decorated the capsid in a 1:1, 3:1 and 11:1 CelA:CBD ratios. Concentrations of all components were quantified by measuring A595 after incubation with Bradford reagent and run on SDS-PAGE gels and stained with coomassie blue for analysis. Once the proper concentration were calculated the components were mixed and incubated overnight at 4°C with rotation. Overnight reaction and were loaded onto a 7.5% SDS-PAGE gel and stained with coomassie blue for analysis.

4.3.6 Synthetic Cellulosome Assembly and Reducing Sugar Assay

Phosphoric acid-swollen cellulose (PASC) was prepared as described previously.¹¹⁶ Final assay mixtures contained 20 nM decorated nanocage of various ratios in the presence of PASC. Synthetic cellulosome assembly was assayed at 30°C in 20 mM Tris-HCl buffer pH 6.0 with shaking. Samples were collected periodically over a 40-hour time period and immediately mixed with 0.25 mL DNS reagent (10 g/L dinitrosalicylic acid, 10 g/L sodium hydroxide, 2 g/L phenol, 0.5 g/L sodium sulfite) and incubated at 95°C for 10 minutes. After incubation, 0.5 mL of 40% Rochelle salts were added to fix the color before measuring the absorbance using a spectrophotometer at 575 nm.

4.4 Results and Discussion

The SpyCatcher/SpyTag bioconjugation method was exploited to enable the highly modular assembly of designer nanocages capable of co-localizing multiple enzymes for improved biocatalysis. The utility of this strategy was exploited to decorate two different protein nanocages with cellulosomal components to evaluate the effects of particle size and decoration density on synergistic cellulose hydrolysis.

4.4.1 Artificial Cellulosome Formation via E2 Self-Assembly

E2 self-assembles from 60 identical subunits into a pentagonal dodecahedron that is approximately 25 nm in diameter. Fusing the small SpyTag to the N-terminus of each monomer allows the ligation of up to 60 proteins onto the exterior surface of the E2 nanocage. By simply modulating the protein ratios for ligation, we generated E2-scaffolded cellulosomes decorated with 1:1 and a 5:1 ratio of CelA:CBD. The correct ligation ratios were confirmed by both SDS-PAGE (Figure 4.1A-B) quantification based on densitometry (Figure 4.1B). It is clear from these results we

have the ability to control the precise decoration ratio even in a one-pot ligation reaction.

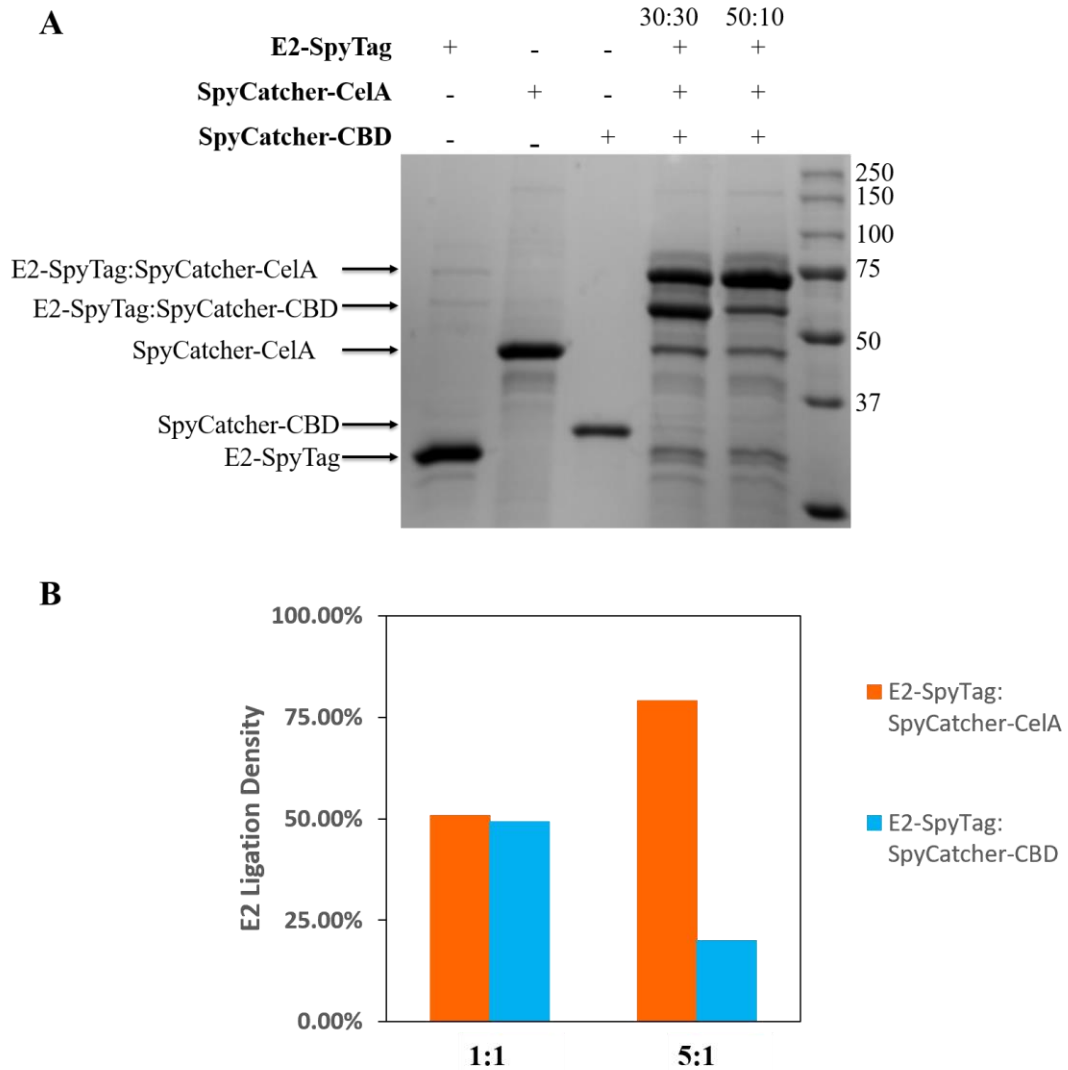


Figure 4.1 **SDS-PAGE analysis of controlled decoration of CelA and CBD to the exterior of E2 nanocage.** (A) Individual purified components are shown in the first three lanes of the SDS-PAGE gel, SpyTag-E2, SpyCatcher-CBD and SpyCatcher-CelA respectively. The components are then mixed at different ratios, 1:1 and 5:1 CelA:CBD, lanes 4 and 5 respectively. (B) Densitometry analysis of product bands in the 1:1 ratio as well as the 5:1 ratio.

To confirm that the decorated E2 nanocages retain their ability to self-assemble into dodecahedron nanostructures, we analyzed ligation products by TEM. The undecorated nanocage as well as nanocages in a 1:1 ratio and 5:1 CelA-CBD ratio were compared (Figure 4.2). The E2-SpyTag protein formed nanocages that are about 25 nm in diameter while the decorated nanocages, regardless of the enzyme ratio, are larger at ~33 nm in diameter, due to the outer layer of proteins on the surface. This is similar to what has been reported previously in our lab with homogenous decoration of CelA onto the surface.¹⁰¹

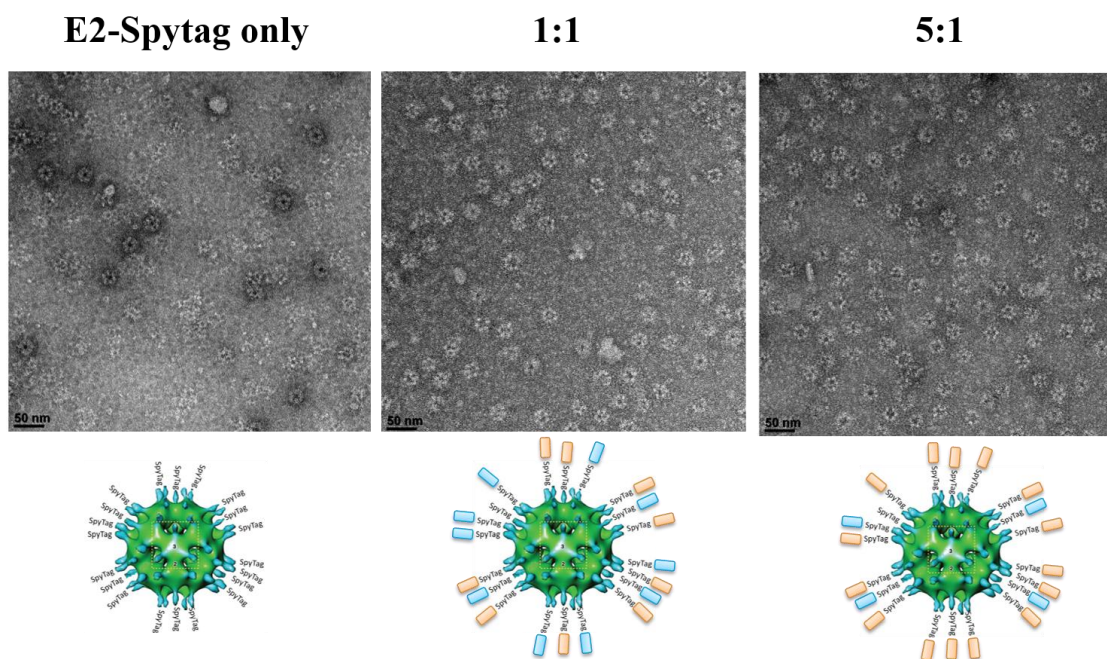


Figure 4.2 **TEM analysis of nanocage formation.** E2-SpyTag only shows nanocage formation with 30 nm diameter while E2-SpyTag decorated with CelA-SpyCatcher and CBD-SpyCatcher in a 1:1 ratio as well as in the 5:1 ratio have an increased size with an average diameter of 33nm.

Having confirmed the ability of our decorated E2-SpyTag proteins to self-assemble into intact nanocages, we employed this scaffolding system for the hydrolysis of cellulose. We compared the E2-SpyTag nanocages decorated with CelA and CBD at the different enzyme ratios against the unmodified E2 nanocage and free CelA and CBD in the ratios of 1:1 and 5:1. Independent of the ligation ratio, E2 decorated with both CelA and CBD produced roughly 2.2-fold more reducing sugars from cellulose.

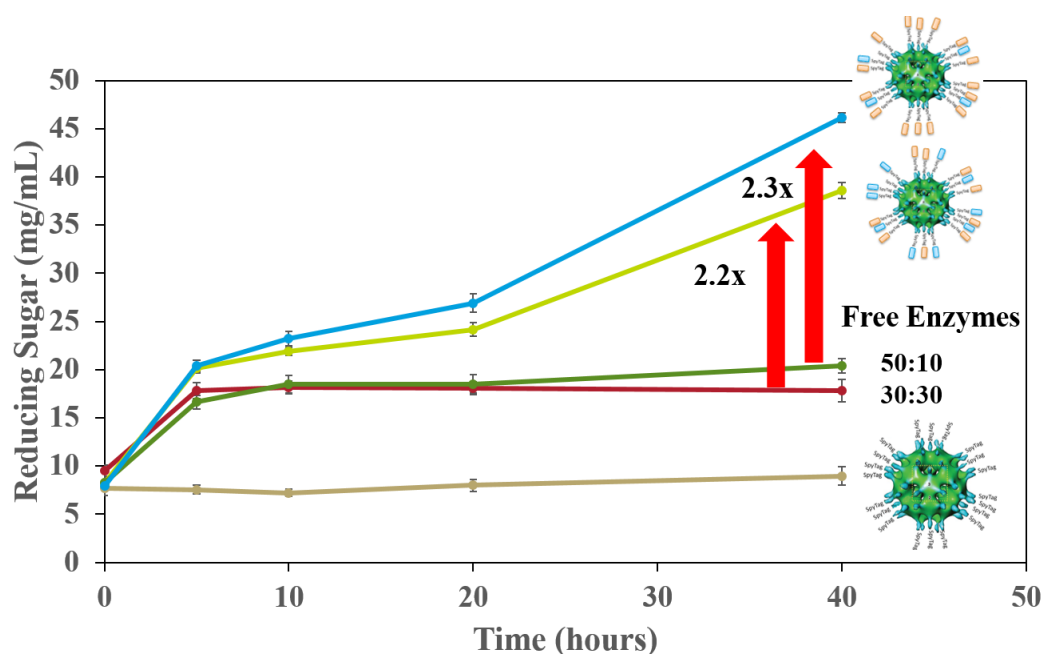


Figure 4.3 **Decorated E2 nanocage increases cellulose hydrolysis.** E2 nanocage was used as a scaffolding platform for the synthesis of mini cellulosomes. The E2-SpyTag nanocage alone showed background levels of reducing sugar production (brown). While the free floating enzymes in either the 1:1 ratio (red) or the 5:1 ratio (green) showed moderate levels of reducing sugar production. Scaffolding the CelA and CBD enzymes onto the E2-SpyTag nanocage caused a 2.2-fold increase in the 1:1 system (yellow) and a 2.3 fold increase in the 5:1 ratio (blue).

The level of enhancement is consistent with that observed when both CelA and CBD were displayed onto the yeast surface¹³⁰ but slightly lower than what we have previously reported with our dCas9-mediated DNA scaffolding platform. We suspect this is due to the higher flexibility afforded by the DNA scaffold, allowing it to wrap around the cellulose fiber in a 3-D context for improved hydrolysis. While the level of enhancement may not be optimal for cellulose hydrolysis, this study does confirm the feasibility to co-localize multiple enzymes onto a single E2 nanocage for a wide range of metabolic pathways of interest.

4.4.2 Artificial Cellulosome Formation via HBV Self-Assembly

To see whether enzyme density and nanocage size can impact the overall synergy, the larger 34nm HBV nanocage was used to assemble the CelA-CBD cellulosomes. Inserting the SpyTag into a surface loop of the HBV monomer provides up to 240 potential sites for ligation. Since the HBV monomer first dimerizes before polymerizing into the 34 nm icosahedral capsid, the surface loops are close to each other in the dimers and only one large protein can be ligated into each dimer. Using the maximum 120 total ligation sites, we chose to study a 1:1, 6:1 and an 11:1 ratio of CelA:CBD scaffolded onto the HBV nanocage for synthetic cellulosome formation. Purified HBV-SpyTag was mixed with varying concentrations of purified CelA-SpyCatcher and CBD-SpyCatcher, allowed to react with mixing, and ran on a 7.5% SDS-PAGE gel for analysis (Figure 4.4 A-B). As expected, roughly 50% of the HBV monomer was reacted and all added Catcher proteins were consumed. The ligation ratios were further quantified using densitometry (Figure 4.4B). From the quantification, we achieved a decorated ratio of 1.2:1, 5.8:1, and 11.2:1, which are

very close to the intended ligation ratios. It is clear by controlling enzyme concentration we can control the relative ratio of enzymes on the HBV nanocage.

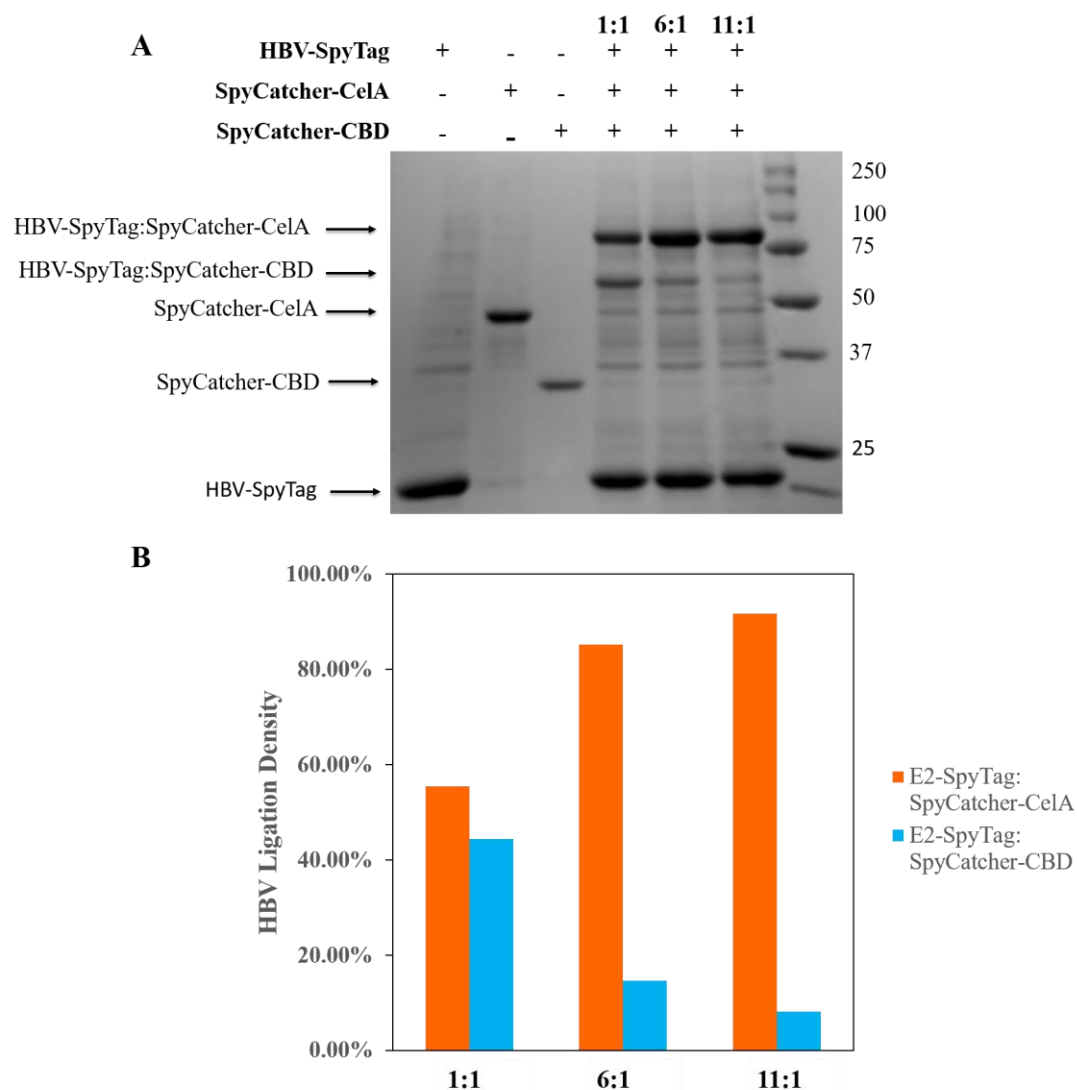


Figure 4.4 **SDS-PAGE analysis of controlled decoration of CelA and CBD to the exterior of the HBV nanocage.** (A) Individual purified components are shown in the first three lanes of the SDS-PAGE gel, SpyTag-HBV, SpyCatcher-CBD and SpyCatcher-CelA respectively. The components are then mixed at different ratios, 1:1, 6:1 and 11:1 CelA:CBD, lanes 4,5 and 6 respectively. (B) Densitometry analysis of product bands in the 1:1 6:1 ratio as well as the 11:1 ratio.

To confirm the decorated HBV nanocages retain the ability to self-assemble into highly ordered nanocages, we analyzed the structures by TEM. The undecorated nanocage as well as the nanocage in the 1:1, 6:1 and 11:1 CelA:CBD ratio were compared (Figure 4.5). The HBV-SpyTag protein self-assembled into a nanocage that is about 34 nm in diameter while the decorated nanocage, regardless of the enzyme ratio, has an increased size of approximately 39 nm in diameter.

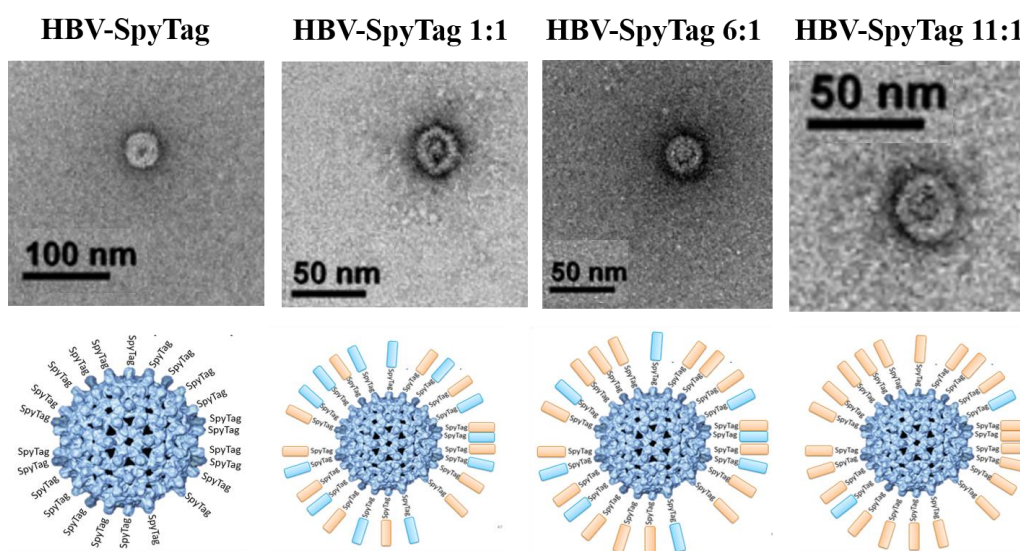


Figure 4.5 TEM analysis of HBV nanocage formation HBV-SpyTag without any enzymes on the exterior (far left) shows native nanocage formation with a 31 nm diameter. HB-SpyTag decorated with CelA-SpyCatcher and CBD-SpyCatcher in the 1:1 (middle left), 6:1 (middle right) and 11:1 (right) has an increased diameter of approximately 34 nm.

Having confirmed that our decorated HBV-SpyTag proteins retain the ability to self-assemble into nanocages, we employed this scaffolding system for the hydrolysis of cellulose. To do this we compared the HBV-SpyTag nanocage alone, as

well as free CelA and CBD not scaffolded to the nanocage, in the same ratios of 1:1, 6:1 and 11:1 against HBV-SpyTag nanocages ligated with CelA and CBD at the same enzyme ratios. As shown in Figure 4.6, the HBV nanocage alone showed background levels of reducing sugar production while free enzymes showed higher levels similar to what we have seen before with non-scaffolded CelA and CBD. However, scaffolding CelA-SpyCatcher and CBD-Spycatcher onto the HBV nanocage showed a 1.8, 1.9 and 2.1-fold increase in reducing sugar production for the 1:1, 6:1 and 11:1 ratios respectively.

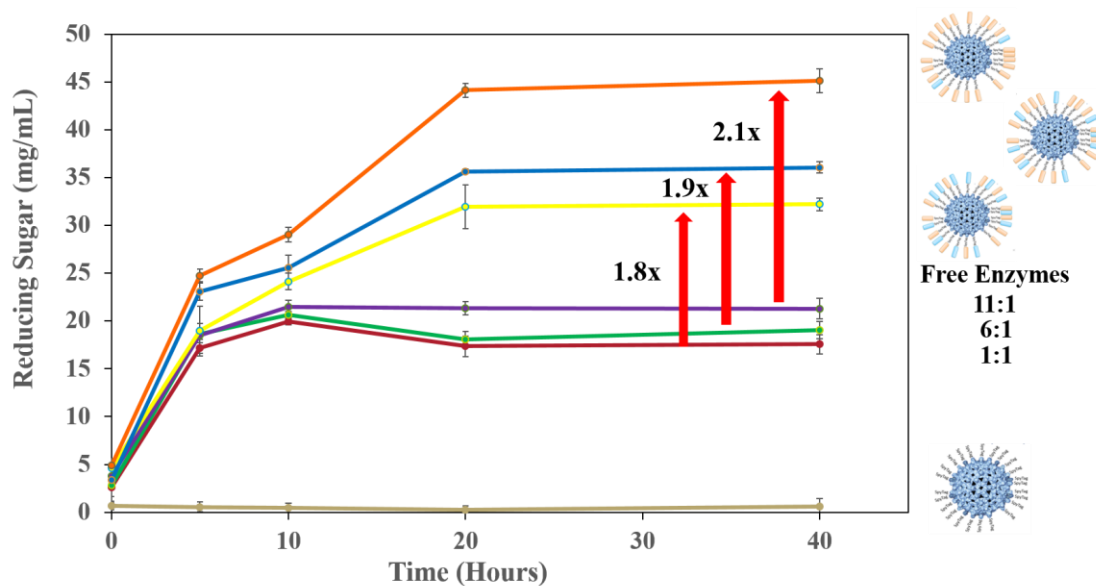


Figure 4.6 **Decorated HBV nanocage increases cellulose hydrolysis.** HBV nanocage was used as a scaffolding platform for the synthesis of mini cellulosomes. The HBV-SpyTag nanocage alone showed background levels of reducing sugar production (brown). While the free floating enzymes in either the 1:1 (red), 6:1 (green) and 11:1 (purple) ratio showed moderate levels of reducing sugar production. Scaffolding the CelA and CBD enzymes onto the HBV-SpyTag nanocage resulted in a 1.8-fold increase for the 1:1 ratio (yellow), 1.9 fold increase in the 6:1 ratio (blue) while the 11:1 ratio showed the most impressive fold enhancement at 2.1 (orange).

It is interesting to note that the level of fold enhancement is lower than using the E2 nanocage platform. This may be the result of the increased spacing between the ligated CelA due to the larger HBV particle size. This result is consistent with our previous observations using different quantum dots for CelA conjugation.⁹³ While the level of synergy may be smaller for the HBV conjugates, the overall rate of hydrolysis is faster due to the higher number of CelA per HBV nanocage.

4.5 Conclusions

In summary, we presented here a new platform to co-localize enzymes using the self-assembling E2 and HBV protein nanocages. By controlling enzyme concentrations, we are able to achieve a range of ligation ratios onto the nanocages capable of enhancing cellulose hydrolysis. These results confirm the ability to co-localize multiple proteins onto the same capsid using the self-assembling SpyCatcher bioconjugation. While there may be a need for optimization of the platforms, the overall concept shows promise, especially for metabolic pathways that require multiple pathway enzymes. For *in vivo* applications, the choice of nanocage used may depend on the overall expression level and the number of enzymes required for the cascade reaction.

Chapter 5

CONCLUSIONS AND FUTURE WORK

5.1 Moving Towards a Fully Dynamic System

Ideally, a dynamic scaffolding system would possess the ability to toggle between an on-state (scaffolded) and an off-state (not scaffolded). While static scaffolding systems improve product titers, they often have negative effects on host cells such as growth retardation, suboptimal productivity and a starvation or cell suicide response. Expanding our modular dCas9 platforms (Chapters 2 and 3) in such a way that a host cell can sense the accumulation of a critical intermediate inside and elicit the co-localization of downstream pathways that convert that intermediate to the final product. In fact, a wide range of genetic circuits have been designed in recent years to provide dynamic control of gene expression and pathway fluxes, resulting in an impressive 50 to 100-fold improvement in product titers.¹²⁵ The majority of these dynamic designs focus on traditional methods of metabolic engineering and were engineered as the foundation to provide the required input-output relationships at both the transcriptional and translational levels.^{126,127} While these dynamic strategies exhibit rapid upregulation, the down regulation is much slower as it relies on the natural degradation of biological components. Expanding the benefits of our modular dCas9 platform while expanding to a dynamic design that focuses both fast turn-on and turn-off rates could be the best of both worlds.

One reliable way to obtain this rapid conditional control would be to employ *ssrA* degradation tags onto the dCas9 proteins. The *ssrA* degradation tags are 11 amino acids peptides recognized by a native protease complex in *E. coli* (Figure 5.1A). The ClpX protein recognizes the *ssrA*-tagged proteins and begins to pull the protein through its pore unfolding it, and the ClpP domain where the unfolded protein is fragmented and released. The endogenous protease complex has an adaptor molecule, SspB which directly attaches the *ssrA*-tagged protein to the ClpXP complex for increased degradation.¹²⁸ The *ssrA* degradation system is of particular interest because the last three amino acids in the tag have been engineered to confer a different rate of degradation, from minutes to hours.¹²⁹ When combined with our dCas9-mediated scaffolding system, we have the ability to tune the rate of synthesis with the rate of degradation, such that when a promoter is induced (via endogenous signaling or added components) the overall expression and subsequent scaffolding is greater than the rate of degradation Figure 5.1B. However, when the promoter is not induced the rate of degradation, controlled by the *ssrA* tags is greater than the rate of synthesis and such the protein would no longer be scaffolded Figure 5.1C.

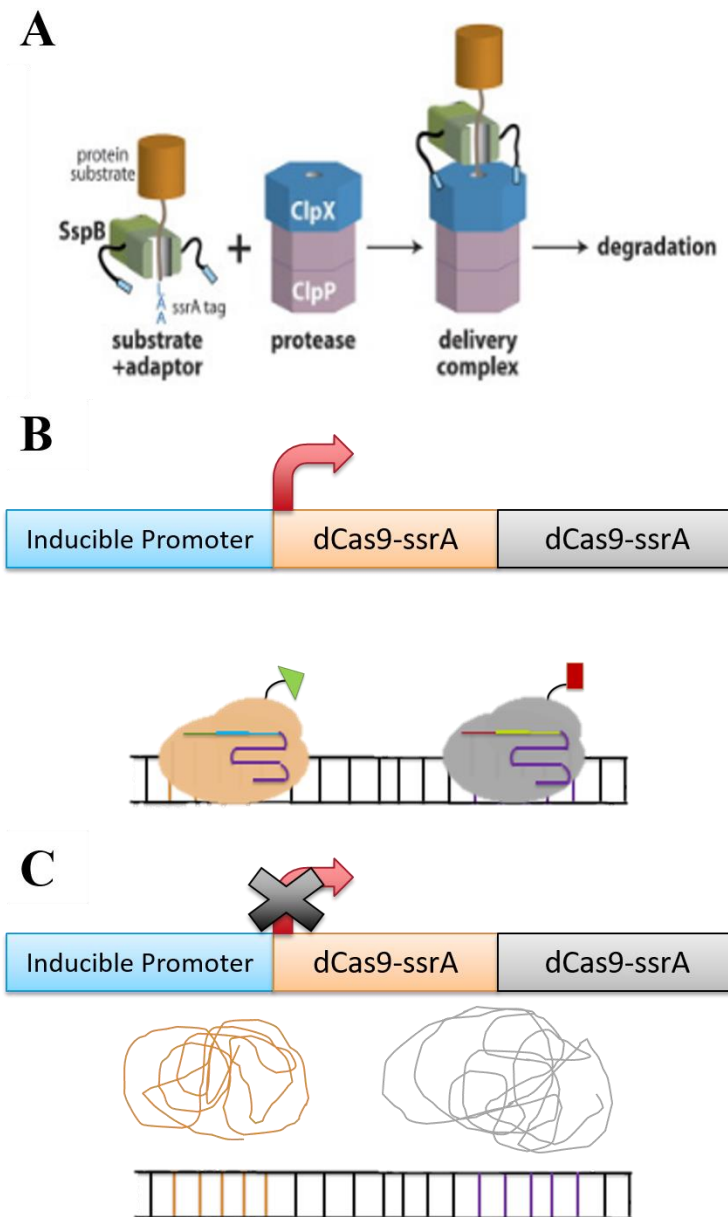


Figure 5.1 **Fully dynamic dCas9 scaffolding system.** (A) Endogenous E.coli degradation complex, ClpXP. (B) Scaffold-on system, translation of dCas9 proteins is greater than the rate of degradation thus synthetic metabolon formation occurs. (C) Scaffold-off system, with no induction to the system, proteins are degraded via the ClpXP system and directed their by the ssrA tags.

5.2 Expanding to Alternative Metabolic Pathways

In this thesis, we have employed both the dCas9 mediated and the self-assembling nanocage platforms for *in vitro* applications. However one of the key features of both platforms is their versatility for *in vivo* applications as well. We envision expanding beyond cellulose hydrolysis and studying other metabolic pathways that are more complex such as the previously studied mevalonate pathway that has shown benefits from scaffolding, as well as the IAA biosynthesis pathway and violacein pathway. We hope to explore the differences in production of the various pathways as they respond to the alternative platforms we have engineered, the dCas-mediated, the E2 nanocage, and the HBV nanocage platforms.

REFERENCES

1. Paddon, C.J. et al. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**, 528-532 (2013).
2. Atsumi, S., Hanai, T. & Liao, J. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86–89 (2008).
3. Yim, H., Haselbeck, R., Niu, W. et al. Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat Chem Biol* **7**, 445–452 (2011).
4. Tai, M. & Stephanopoulos, G. Engineering the push and pull of lipid biosynthesis in oleaginous yeast *Yarrowia lipolytica* for biofuel production. *Metabolic Engineering* **15**, 1-9 (2013).
5. Lan, E.I. & Liao, J.C. ATP drives direct photosynthetic production of 1-butanol in cyanobacteria. *Proceedings of the National Academy of Sciences* **109**, 6018-6023 (2012).
6. Agapakis, C., Boyle, P. & Silver, P. Natural strategies for the spatial optimization of metabolism in synthetic biology. *Nat Chem Biol* **8**, 527–535 (2012).
7. Velot, C., Mixon, M.B., Teige, M. & Srere, P.A. Model of a quinary structure between Krebs TCA cycle enzymes: A model for the metabolon. *Biochemistry* **36**, 14271-14276 (1997).
8. Good, M.C., Zalatan, J.G. & Lim, W.A. Scaffold Proteins: Hubs for Controlling the Flow of Cellular Information. *Science* **332**, 680-686 (2011).
9. Conrado, R.J. et al. DNA-guided assembly of biosynthetic pathways promotes improved catalytic efficiency. *Nucleic Acids Research* **40**, 1879-1889 (2012).
10. Lee, H., DeLoache, W.C., Dueber, J.E. Spatial organization of enzymes for metabolic engineering. *Metabolic Engineering* **14**, 242-251 (2012)

11. Dueber J., Wu G., Malmirchegini G. *et al.* Synthetic protein scaffolds provide modular control over metabolic flux. *Nature Biotechnology* **27**, 753–759 (2009).
12. Albertsen L, Chen Y, Bach LS, Rattleff S, Maury J, Brix S *et al.* Diversion of flux toward sesquiterpene production in *Saccharomyces cerevisiae* by fusion of host and heterologous enzymes. *Applied and Environmental Microbiology* **77**, 1033-1040 (2011).
13. Whitaker WR, Davis SA, Arkin AP, Deuber JE. Engineering Robust control of two-component system phosphotransfer using modular scaffolds. *PNAS* **109**, 18090-18095 (2012).
14. Y. Mori, S. Ozasa, M. Kitaoka, S. Noda, T. Tanaka, H. Ichinose, *et al.* Aligning an endoglucanase Cel5A from *Thermobifida fusca* on a DNA scaffold: potent design of an artificial cellulosome. *Chemical Communications* **49**, 6971-6973 (2013).
15. Condro RJ, Wu GC, Boock JT, Xu H, Chen SY, Lebar T *et al.* DNA-guided assembly of biosynthetic pathways promotes improved catalytic efficiency. *Nucleic Acid Research* **40**, 1879–1889 (2011).
16. Sun Q and Chen W. HaloTag mediated artificial cellulosome assembly on rolling circle amplification DNA template for efficient cellulose hydrolysis. *Chemical Communications* **52**, 6701-6704 (2016).
17. Sun Q., Madan B., Tsai S.-L., DeLisa M.P. and Chen W., Creation of artificial cellulosomes on DNA scaffolds by zinc finger protein-guided assembly for efficient cellulose hydrolysis, *Chemical Communications* **50**, 1423-1425 (2014).
18. Miles EW. Tryptophan synthasae: a multienzyme complex with an intramolecular tunnel. *The Chemical Record* **1**, 140-151 (2001).
19. An S., Kumar S., Sheets E.D. & Benkovic S.J. Reversible compartmentalization of de novo purine biosynthetic complexes in living cells. *Science* **320**, 103-106 (2008).
20. Agapakis, C., Boyle, P. & Silver, P. Natural strategies for the spatial optimization of metabolism in synthetic biology. *Nature Chemical Biology* **8**, 527–535 (2012).

21. Chen R., Chen Q., Kim H., Siu, K., Sun Q., Tsai, S. & Chen W. Biomolecular scaffolds for enhanced signaling and catalytic efficiency. *Current Opinion in Biotechnology* **28**, 59-68 (2014).
22. Bulow L., Ljungcrantz P. & Mosbach K. Preparation of a Soluble Bifunctional Enzyme by Gene Fusion. *Nature Biotechnology* **3**, 821-823 (1985).
23. Salles I.M., Forchhammer N., Croux C., Girbal L. & Soucaille P. Evolution of a *Saccharomyces cerevisiae* metabolic pathways in *Escherichia coli*. *Metabolic Engineering* **9**, 152-159 (2007).
24. Wang J.H., Tsai M.Y., Lee G.C. & Shaw J.F. Construction of a recombinant thermostable beta-amylase-trehalose synthase bifunctional enzyme for facilitating the conversion of starch to trehalose. *Journal of Agricultural and Food Chemistry* **55**, 1256-1263 (2007).
25. Seo H.S., Koo Y.J., Lim J.Y., Song J.T., Kim C.H., Kim J.K., Lee J.S. & Choi Y.D. Characterization of a bifunctional enzyme fusion of trehalose-6-phosphate synthetase and trehalose-6-phosphate phosphatase of *Escherichia coli*. *Applied and Environmental Microbiology* **66**, 2484-2490 (2000).
26. Kim Y.H., Kwon T.K., Park S., Seo H.S., Cheong J.J., Kim C.H., Kim J.K., Lee J.S. & Choi Y.D. Trehalose synthesis by sequential reactions of recombinant maltooligosyltrehalose synthase and maltooligosyltrehalose trehalohydrolase from *Brevibacterium helvolum*. *Applied and Environmental Microbiology* **66**, 4620-4624 (2000).
27. de Pascale, .D., Di Lernia, .I., Sasso, .M. *et al.* A novel thermophilic fusion enzyme for trehalose production. *Extremophiles* **6**, 463–468 (2002).
28. Prachayasittikul, V., Ljung, S., Isarankura-Na-Ayudhya, C., & Bülow, L. NAD(H) recycling activity of an engineered bifunctional enzyme galactose dehydrogenase/lactate dehydrogenase. *International journal of biological sciences* **2**, 10–16 (2006).
29. Riedel K. & Bronnermeier K. Intramolecular synergism in an engineered exo-endo-1,4- β -glucanase fusion protein. *Molecular Microbiology* **28**, 767-775 (2002).
30. Conrado R.J., Varner J.D. & DeLisa M.P. Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. *Current Opinion in Biotechnology* **19**, 492-299 (2008).

31. Kourtz. L., Dillon K., Daughtry S., Madison L.L., Peoples O. & Snell K.D. A novel thiolase-reductase gene fusion promotes the production of polyhydroxybutyrate in Arabidopsis. *Plant Biotechnology Journal* **3**, 435-447 (2005).
32. Fierobe H., Mechaly A., Tardif C., Belaich A., Lamed R., Shoham Y., Belaich J. & Bayer E.A. Design and Production of Active Cellulosome Chimeras. *Journal of Biological Chemistry* **276**, 21257-21261 (2001).
33. Moon T.S, Dueber J.E., Shiue E. & Prather K.L.J. Use of modular, synthetic scaffolds for the production of glucarid acid in engineered E.coli. *Metabolic Engineering* **12**, 298-305 (2010).
34. Moon T.S., Yoon S.-H., Lanza A.M. Roy-Mayhew J.D. & Prather K.L.J. Production of glucaric acid from a synthetic pathways in recombinant Escherichia coli. *Applied and Environmental Microbiology* **75**, 589-595 (2009).
35. Kim, A.S., Kakalis, L.T., Abdul-Manan, N., Liu, G.A. & Rosen, M.K. Autoinhibition and activation mechanisms of the Wiskott-Aldrich syndrome protein. *Nature* **404**, 151-158 (2000).
36. Harris, B.Z., Hillier, B.J. & Lim, W.A. Energetic determinants of internal motif recognition by PDZ domains. *Biochemistry* **40**, 5921-5930 (2001).
37. Nguyen, J.T., Turck, C.W., Cohen, F.E., Zuckermann, R.N. & Lim, W.A. Exploiting the basis of proline recognition by SH3 and WW domains: design of N-substituted inhibitors. *Science* **282**, 2088-2092 (1998).
38. Baneyx, F., Mujacic, M. Recombinant protein folding and misfolding in *Escherichia coli*. *Nature Biotechnology* **22**, 1399–1408 (2004)
39. Wilner O., Shimron S. & Weizmann Y. Self-assembly of enzymes on DNA scaffolds: en route to biocatalytic cascades and the synthesis of metallic nanowires. *Nano Letters* **9**, 2040-2043 (2009).
40. Fu J., Liu M., Liu Y., Woodbury N.W. & Yan H. Interenzyme Substrate Diffusion for an Enzyme Cascade Organized on Spatially Addressable DNA Nanostructures. *Journal of the American Chemical Society* **134**, 5516-5519 (2012).

41. V Los G., Encell L.P., Mcdougall M.G., Hartzell D.D., Karassina N., Zimprich C., Wood M.G. Learish R., Ohana R.F. Urh M., Simpson D., Mendez J., Zimmerman K., Otto P., Vidugiris G., Zhu J., Darzins A., Klaubert D.H., Bulleit R.F. & Wood K.V. HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chemical Biology* **3**, 373-382 (2008).
42. Sun Q. and Chen W. HaloTag mediated artificial cellulosome assembly on rolling circle amplification DNA template for efficient cellulose hydrolysis. *Chemical Communications* **52**, 6701-6704 (2016).
43. Maeder M.L., Thibodeau-Beganny S., Osiak A., Wright D.A., Anthony R.M., Eichtinger M., Jiang T., Foley J.E., Winfrey R.J., Townsend J.A., Unger-Wallace E., Sander J.D., Muller-Lerch F., Fu F., Pearlberg J., Gobel C., Dassie J.P., Pruett-Miller S.M., Porteus M.H., Sgroi D.C., Iafrate A.J., Dobbs D., McCray P.B., Cathomen T., Voytas D.F. & Joung J.K. Rapid “open-source” engineering of customized zinc-finger nucleases for highly efficient gene modification. *Molecular Cell* **31**, 294-301 (2008).
44. Greisman H.A. & Pabo C.O. A General Strategy for Selecting High-Affinity Zinc Finger Proteins for Diverse DNA Target Sites. *Science* **275**, 657-661 (1997).
45. Sun Q., Madan B., Tsai S., DeLisa M.P. & Chen W. Creation of artificial cellulosomes on DNA scaffolds by zinc finger protein-guided assembly for efficient cellulose hydrolysis. *Chemical Communications* **50**, 1423-1425 (2014).
46. Jansen R., van Embden J.D.A., Gaastra W. & Schouls L.M. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology* **43**, 1565-1575 (2002).
47. Jansen R., van Embden J.D.A., Gaastra W. & Schouls L.M. Identification of Novel Family Sequence Repeats among Prokaryotes. *OMICS: A Journal of Integrative Biology* **6**, 23-22 (2002).
48. Bolotin A., Quinquis B., Sorokin A. & Ehrlich S.D. Clustered regularly interspaced short palindrome repeats(CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**, 2551–2561 (2005).
49. Mojica F.J., Diez-Villasenor C., Garcia-Martinez J. & Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution* **60**, 174–182 (2005).

50. Pourcel C., Salvignol G. & Vergnaud G. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151**, 653–663 (2005).
51. Barrangou R., Fremaux C., Deveau H., Richards M., Boyaval P, et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).
52. Chylinski K., Makarova K.S., Charpentier E. & Koonin E.V. Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Research* **42**, 6091-6105 (2014).
53. Barrangou R. and Horvath P. CRISPR: new horizons in phage resistance and strain identification. *Annual Review of Food Science and Technology* **3**, 143-162 (2012)
54. Makarova K.S., Wolf Y.I., Alkhnbashi O.S., Costa F., Shah S.A., Saunders S.J., Barrangou R., Brouns S.J.J., Charpentier E., Haft D.H., Horvath P., Moineau S., Mojica F.J.M., Terns R.M., Terns M.P., White M.F., Yakunin A.F., Garrett R.A., van der Oost J., Backofen R. & Koonin E.V. An updated evolutionary classification of CRISPR-Cas systems. *Nature* **13**, 722-736 (2015).
55. Jiang F., Zhou K., Ma L., Gressel S. & Doudna J.A. A Cas9-guide RNA complex preorganized for target DNA recognition. *Science* **348**, 1477-1481 (2015).
56. Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A. & Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821 (2012).
57. Gasiunas G., Barrangou R., Horvath P. & Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *PNAS. USA.* **109** (2012)
58. Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A. & Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821 (2012)
59. Mojica F.J., Díez-Villaseñor C., García-Martínez J. & Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155**, 733-740 (2009).

60. Esvelt K.M., Mali P., Braff J.L., Moosburner M., Yaung S.J. & Church G.M. Orthogonal Cas9 Proteins for RNA-Guided Gene Regulation and Editing. *Nature* **10**, 1116-1121 (2013).
61. Gao Y., Xiong X., Wong S., Charles E.J., Lim W.A. & Qi L.S. Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. *Nature Methods* **13**, 1043-1049 (2016).
62. Ma H., Naseri A., Reyes-Gutierrez P., Wolfe S.A., Zhang S. & Pederson T. Multicolor CRISPR labeling of chromosomal loci in human cells. *PNAS* **112**, 3002-3007 (2015).
63. Qi L.S., Larson M.H., Gilbert L.A., Doudna J.A., Weissman J.S., Arkin A. & Lim W.A. Repurposing CRISPR as an RNA-guided platform for sequence specific control of gene expression. *Cell Press* **152**, 1173-1183 (2013).
64. Chen, B. et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479–1491 (2013).
65. Chavez, A., Scheiman, J., Vora, S. et al. Highly efficient Cas9-mediated transcriptional programming. *Nature Methods* **12**, 326–328 (2015).
66. Gilbert, L. A. et al. Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* **159**, 647–661 (2014).
67. Bikard, D. et al. Programmable repression and activation of bacterial gene expression using an engineered CRISPR–Cas system. *Nucleic Acids Res.* **41**, 7429–7437 (2013). Gilbert, L. A. et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013). Repressor
68. Lawhorn, I. E., Ferreira, J. P. & Wang, C. L. Evaluation of sgRNA target sites for CRISPR-mediated repression of *TP53*. *PLoS ONE* **9**, e113232 (2014).
69. Fujita, T. & Fujii, H. Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. *Biochem. Biophys. Res. Commun.* **439**, 132–136 (2013).
70. Fujita, T. et al. Identification of telomere-associated molecules by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP). *Sci. Rep.* **3**, 3171 (2013).

71. Polstein, L. R. & Gersbach, C. A. A light-inducible CRISPR–Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol.* **11**, 198–200 (2015).
72. Nihongaki, Y., Yamamoto, S., Kawano, F., Suzuki, H. & Sato, M. CRISPR–Cas9-based photoactivatable transcription system. *Chem. Biol.* **22**, 169–174 (2015).
73. Siu K. & Chen W. Riboregulated toehold-gated gRNA for programmable CRISPR-Cas9 function. *Nature Chemical Biology* **15**, 217-220 (2019).
74. Fontes C.M.G.A. & Gilbert H.J. Cellulosomes: Highly Efficient Nanomachines Designed to Deconstruct Plant Cell Wall Complex Carbohydrates. *Annual Review of Biochemistry* **79**, 655-681 (2010).
75. Rajagopal, D.; Zilberman, D. Review of Environmental , Economic and Policy Aspects of Biofuels. *Foundations and Trends in Microeconomics* **4**, 353-468 (2007).
76. Lynd, L. R.; Weimer, P. J.; Zyl, W. H. Van; Pretorius, I. S. Microbial Cellulose Utilization : Fundamentals and Biotechnology. *Microbiology and Molecular Biology Reviews* **66**, 506–577 (2002).
77. Bayer EA, Kenig R, Lamed R. Adherence of *Clostridium thermocellum* to cellulose. *Journal Bacteriology* **156**, 818–827 (1983).
78. Lamed R, Setter E, Bayer EA. Characterization of a cellulose-binding, cellulase-containing complex in *Clostridium thermocellum*. *Journal Bacteriology* **156**, 828–836 (1983)
79. Lamed R, Setter E, Kenig R, Bayer EA. The cellulosome: a discrete cell surface organelle of *Clostridium thermocellum* which exhibits separate antigenic, cellulose-binding and various cellulolytic activities. *Biotechnology and Bioengineering Symposium* **13**, 163–181 (1983).
80. Schwarz, W. H. The Cellulosome and Cellulose Degradation by Anaerobic Bacteria. *Applied Microbiology and Biotechnology* **56**, 634–649 (2001).
81. Tsai, S.; Dasilva, N. A.; Chen, W. Functional Display of Complex Cellulosomes on the Yeast Surface via Adaptive Assembly. *ACS Synthetic Biology* **2**, 14–21 (2013).

82. Fierobe H., Bayer E.A., Tardif C., Czjzek M., Mechaly A., Belaich A., Lamed R., Shoham. & Belaich J. Degradation of Cellulose Substrates by Cellulosome Chimeras: Substrate Targeting Versus Proximity of Enzyme Components. *Journal of Biological Chemistry* **277**, 49621-48630 (2002).
83. Cha J., Matsuoka S., Chan H., Yukawa H., Inui M * Doi R.H. Effect of Multiple Copies of Cehesins on Cellulase and Hemicellulase Activities of *Clostridium cellulovorans* Mini-cellulosomes. *Journal of Microbiology and Biotechnology* **17**, 1782-1788 (2007).
84. Mingardon F., Chanal A., Lopez-Contreras A.M., Dray C., Bayer E.A. & Fierobe H. Incorporation of Fungal Cellulases in Bacterial Minicellulosomes Yields Viable, Synergistically Acting Cellulolytic Complexes. *Applied and Environmental Microbiology* **73**, 3822-3832 (2007).
85. Fierobe H., Mingardon F., Mechaly A., Belaich A., Rincon M.T., Pages S., Lamed R., Tardif C., Belaich J. & Bayer E.A. Action of Designer Cellulosome on Homogenous Versus Complex Substrates controlled incorporation of three distinct enzymes into a defined trifunctional scaffoldin. *Journal of Biological Chemistry* **280**, 16325-16334 (2005).
86. Mingardon F., Chanal A., Tardif C., Bayer E.A. & Fierobe H. Exploration of New Geometries in Cellulosome-Like Chimeras. *Applied and Environmental Microbiology* **73**, 7138-7149 (2007).
87. Zakeria B., Fierer J.O., Celik E., Chittock E.C., Schwarz-Linek U., Moy V.T. & Howarth M. Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *PNAS* **109**, 690-697 (2012).
88. Veggiani G., Nakamura T., Brenner M.D., Gayet R.V., Yan J., Robinson C.V. & Howarth M. Programmable polyproteins built using twim peptide superglues. *PNAS* **113**, 1202-1207 (2016).
89. Kim J.S., Valencia C.A., Liu R. & Lin W. Highly-efficient purification of native polyhistidine-tagged proteins by multivalent NTA-modified magnetic nanoparticles. *Bioconjug. Chem.* **18**, 333-341 (2007).
90. Lauer S.A. & Nolan J.P. Development and characterization of Ni-NTA-bearing microspheres. *Cytometry* **48**, 136-145 (2002).

91. Kim D., Umetsu M., Takai k., Matsuyama T., Ishida N., Takahashi H., Asano R.* Kumagai I., Enhancement of Cellulolytic Enzyme Activity by Clustering Cellulose Binding Domains on Nanoscaffolds. *Small* **7**, 656-664 (2011).
92. Blanchette C., Lacayo C.I., Fischer N.O., Hwang M & Thelen M.P. Enhanced Cellulose Degradation Using Cellulase-Nanosphere Complexes. *PLoS ONE* **7**, 1-7 (2012).
93. Tsai S.L., Park M. & Chen W. Size-modulated synergy of cellulose clustering for enhanced cellulose hydrolysis. *Biotechnology Journal* **8**, 257-261 (2013).
94. Mitsuzawa S., Kagawa H., Li Y., Chan S.L., Paavola C.D. & Trent J.D. The rosettazyme: a synthetic cellulosome. *Journal of Biotechnology* **143**, 139-144 (2009).
95. Chen Q., Sun Q., Molino N.M., Wang S-W., Boder E.T. & Chen W. Sortase A-mediated multi-functionalization of protein nanoparticles. *Chemical Communications* **51**, 12107-12110 (2015).
96. IZARD T., Aevansson A., Allen M.D., Westphal A.H., Perham R.N., deKok A. & Hol W.G. Principles of quasi-equivalence and Euclidean geometry govern the assembly of cubic and dodecahedral cores of pyruvate dehydrogenase complexes. *PNAS* **96**, 1240-1245 (1999).
97. Milne J.L.D., Shi D., Rosenthal P.B., Sunshine J.S., Domingo G.J., Wu X., Brooks B.R., Perham R.N., Henderson R. & Subramaniam S. Molecular Structure of a 9-MDa Icosahedral Pyruvate Dehydrogenase Subcomplex Containing the E2 and E3 Enzymes Using Cryoelectron Microscopy. *Journal of Biological Chemistry* **21**, 5587-5598 (2002).
98. Ren D., Dalmau M., Randall A., Shindel M.M., Baldi P. & Wang S-W. Biomimetic Design of Protein Nanomaterials for Hydrophobic Molecular Transport. *Advanced Functional Matter* **22**, 3170-3180 (2012).
99. Molino N.M., Anderson A.K.L., Nelson E.L. & Wang S.W. Biomimetic Protein Nanoparticles Facilitate Enhanced Dendritic Cell Activation and Cross-Presentation. *ACS Nanotechnology* **7**, 9743-9752 (2013).
100. Peng T., Paramelle D., Sana B., Lee C.F. & Lim S. Designing Non-Native Iron-Binding Site on a Protein Cage for Biological Synthesis of Nanoparticles. *Small* **10**, 3131-3138 (2014).

101. Ren D., Kratz F. & Wang S.-W. Engineered drug- protein nanoparticle complexes for folate receptor targeting. *Biochemical Engineering Journal* **89**, 33-41 (2014).
102. Murata M., Narahara S., Kawano T., Hamano N., Piao J, et al. Design and function of engineered protein nanocages as a drug delivery system for targeting pancreatic cancer cells via neuropilin-1. *Molecular Pharmacology* **12**, 1422-1430 (2015).
103. Gerin J.L., Holland P.V. & Purcell R.H. Australian antigen: larger-scale purification from human serum and biochemical studies of its proteins. *Journal of Virology* **7**, 569-576 (1971).
104. Glasgow J. & Tullman-Ercek D. Production and applications of engineered viral capsids. *Applied Microbiology and Biotechnology* **98**, 5847-5858 (2014).
105. Guss B., Eliasson M., Olsson A., Uhlen M., Frej A.K., et al. Structure of the IGG-binding regions of streptococcal protein-G. *EMBO J.* **5**, 1567–1575 (1986).
106. Choi K.M., Choi S.H., Jeon H., Kim I.S. & Ahn H.J. Chimeric capsid protein as a nanocarrier for siRNA delivery: stability and cellular uptake of encapsulated siRNA. *ACS Nanotechnology* **5**, 8690–8699 (2011).
107. Pontrelli S., Chiu T.-Y., Lan E.I., Chen F. Y.-H., Chang P. and Liao J.C. Escherichia coli as a host for metabolic engineering. *Metabolic Engineering* **50**, 16–46 (2018).
108. Peralta-Yahya P., Zhang F., del Cardayre S.B & J.D. Keasling. Microbial engineering for the production of advanced biofuels. *Nature* **488**, 320-328 (2012).
109. Jorgensen K., Rasmussen A.V., Morant M., Nielsen A.H., Bjarnholt N., Zagrobelny M., Bak S. & Moller B.L. Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Current Opinion in Plant Biology* **8**, 280-291 (2005).
110. Garbett D. & Bretscher A. The suprising dynamics of scaffolding proteins. *Molecular Biology of the Cell* **25**, 2315-2319 (2014).
111. Smirnoff N. Engineering of Metabolic Pathways Using Synthetic Enzymes Complexes. *Plant Physiology* **179**, 918-928 (2019).

112. Wheeldon I., Minter S.D., Banta S., Barton S.C., Atanassov P. & Sigman M. Substrate channeling as an approach to cascade reactions. *Nature Chemistry* **8**, 299-309 (2016).
113. Sweetlove L.J. & Fernie A.R. The role of dynamic enzyme assemblies and substrate channeling in metabolic regulation. *Nature Communications* **9**, 2136 (2018).
114. Pinheiro A.V., Han D., Shih W.M. & Yan H. Challenges and opportunities for structural DNA nanotechnology. *Nature Nanotechnology* **6**, 736-722 (2011).
115. Adli M. The CRISPR tool kit for genome editing and beyond. *Nature Communications* **9**, 1911 (2018).
116. Mali P., Esvelt K.M. & Church G.M. Cas9 as a versatile tool for engineering biology. *Nature Methods* **10**, 957-963 (2013).
117. Liu F., Tsai S.-L., Madan B. & Chen W. Engineering a high-affinity scaffold for non-chromatographic protein purification via intein-mediated cleavage. *Biotechnology and Bioengineering* **109**, 2829-2835 (2012).
118. Semenova E., Kuznedelov K., Minakhin L, Mekler V. & Severinov K. Kinetics of the CRISPR-Cas9 effector complex assembly and role of the 3'-terminal segment of guide RNA. *Nucleic Acids Research* **44**, 2837-2845 (2016).
119. Hall M.P., Unch J., Binkowski B.F., Valley M.P., Butler B.L., Wood M.G., Otto P., Zimmerman K., Vidugiris G., Machleidt T., Roberts M.B Benink H.A., Eggers C.T., Slater M.R., Meisenheimer P.L. Klauert D.H., Fan F. Encell L.P. & Wood K.V. Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chemical Biology* **7**, 1848-1857 (2012).
120. Dixon A.S., Schwinn M.K., Hall M.P., Zimmerman K., Otto P., Lubben T.H., Butler B.L., Binkowski B.F., Machleidt T., Kirland T.A., Wood M.G., Eggers C.T., Encell L.P. & Wood K.V. NanoLuc complementation reporter optimized for accurate measurement of protein interaction in cells. *ACS Chemical Biology* **11**, 400-408 (2016).
121. Moller B.L. Plant Science. Dynamic metabolons. *Science* **330**, 1328-2329 (2010).

122. Chen R.P., Blackstock D., Sun Q. & Chen W. Dynamic protein assembly by programmable DNA strand displacement. *Nature Chemistry* **10**, 474-481 (2018).
123. Berckman E. & Chen W. Exploiting dCas9 fusion proteins for dynamic assembly of synthetic metabolons. *Chemical Communications* **55**, 8219-8222 (2019).
124. Peter H., Sorensen K. & Mortensen K. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *Journal of Biotechnology* **115**, 112-128 (2005).
125. Xu, P., Li, L., Zhang, F., Stephanopoulos, G. & Koffas, M. Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. *Proceedings of the National Academy of Sciences* **111**, 11299-11304 (2014).
126. Zhang, F., Carothers, J.M. & Keasling, J.D. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nature Biotechnology* **30**, 354-359 (2012).
127. Win, M.N. & Smole, C.D. A modular and extensible RNA-based gene regulatory platform for engineering cellular function. *Proceedings of the National Academy of Sciences* **104**, 14283-14288 (2007).
128. Grottesman S., Roche E., Zhou Y.N. & Sauer R.T. The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes and Development* **12**, 1338-1347 (1998).
129. McGinness K.E., Baker, T.A. & Sauer R.T. Engineering Controllable Protein Degradation. *Cell Press* **22**, 701-707 (2006).
130. Tsai. S-L., Oh J., Singh S., Chen R. and Chen W. Functional Assembly of Minicellulosomes on the *Saccharomyces cerevisiae* Cell Surface for Cellulose Hydrolysis and Ethanol Production. *Applied and Environmental Microbiology* **75**, 6087-6093 (2009).

Appendix A

LIST OF ABBREVIATIONS

dCas9	nuclease- null Cas9
DNA	deoxyribonucleic acid
RNA	ribonucleic acid
SgRNA	single-guide RNA
thgRNA	toe-hold guide RNA
Tag	small peptide tag
Catcher	protein binding partner
LacZ	β -galactosidase
GalK	galactokinase
NADH	Nicotinamide adenine dinucleotide
SH3	Src homology 3
PDZ	post-synaptic density, disc large, zo-1 protein
GDB	GTPase binding domain
K _d	dissociation constant
ZFP	zinc finger protein
CRISPR	clustered regularly interspaced short palindromic repeats
Cas	CRISPR-associate genes
crRNA	CRISRP RNA
tracrRNA	trans-activating crRNA
dsDNA	double stranded DNA
PAM	protospacer adjacent motif
CBM	cellulose binding module
CelA	endoglucanase
CH	chlorohexane
SpyTag	13 amino acid peptide from <i>Streptococcus pyogenes</i>
SpyCatcher	15 kDa protein binding partner from <i>Streptococcus pyogenes</i>

SnoopTag	13 amino acid peptide from <i>Streptococcus pneumoniae</i>
SnoopCatcher	15 kDa protein binding partner from <i>Streptococcus pneumoniae</i>
AtoB	acetoacetyl-CoA thiolase
HMGS	hydroxy-methylglutaryl-CoA synthase
HMGR	hydroxymethylglutaryl-CoA reductase
E2	pyruvate dehydrogenase complex of the thermophilic bacterium <i>Bacillus stearothermophilus</i>
HBV	Hepatitis B virus
TEM	transmission electron microscopy
SpdCas9	nuclease-null Cas9 proteins <i>S.pyogenes</i>
ST1dCas9	nuclease-null Cas9 proteins <i>S.thermophilus</i>
SadCas9	nuclease-null Cas9 proteins <i>S.aureus</i>
PCR	polymerase chain reaction
LB	Luria-Bertani
IPTG	isopropyl-thiogalactopyranoside
SDS	sodium dodecyl sulfate
PAGE	polyacrylamide gel electrophoresis
TBS	Tris Buffered Saline
EMSA	Electrophoretic mobility shift assay
PASC	Phosphoric acid-swollen cellulose
bp	basepair
Nluc	Nano luciferase
LgBit	159 amino acid fragment of Nluc
SmBit	11 amino acid fragment of Nluc
ssrA	11 amino acids peptide degradation tags
SspB	adaptor molecule
ClpXP	endogenous protein comple in <i>E. coli</i>
IAA	Indole-3-acetic acid

Appendix B

BEACON SEQUENCES

<u>Beacon</u>	<u>Name</u>	<u>Sequence (5' →3')</u>
SpT2 Beacon	oCR_t_30 FQ SpT2 FP	ataGGTATCACATGACTAAACGA
	oCR_t_31 Sp PAM	AGGtctagaGCTAGC
	oCR_t_32 SpT2 FAM RP	GCTAGCtctagaCCTTCGTTTAGTCAT GTGATACCtat
ST1 Beacon	oCR_t_33 FQ St1T1 FP	tatATGGAATGGAATGGAATGGA
	oCR_t_34 ST1 PAM	TTGGAAaga
	oCR_t_35 ST1T1 FAM RP	tctTTCCAATCCATTCCATTCCATTC CATata
Sa Beacon	oCR_t_36 FQ Sa FP	ataGAAGTGTCGGTGACAGGAAG
	oCR_t_37 Sa PAM	AGGGGAtct
	oCR_t_38 Sa FAM RP	agaTCCCCTCTTCCTGTCACCGACA CTTCtct
SpT2 15 bp St1 Dual Beacon	oCR_t_044 Sp PAM 15 bp St1 protospacer Cy5	AGGtctagagactacgacTTCCAAtccattccatt ccattccatTAT

	oCR_t_043 SpT2 protospacer 15 bp St1 PAM FAM	TTGGAAgtcgtagtctctagaCCTTCGTTT AGTCATGTGATACctat
	oCR_t_30 FQ SpT2 FP	ataGGTATCACATGACTAAACGA
	oCR_b_26 St1 Iowa Black RQ	aggtaaggtaaggtaaggtaata
SpT2 25 bp St1 Dual Beacon	oCR_t_046 25 bp spacer St1 protospacerCy5	TTGGAAgtaatgatcggtcgtagtctctagaCCT TCGTTTAgTCATGTGATACctat
	oCR_t_045 SpT2 protospacer 25 bp St1PAM FAM	AGGtctagagactacgaccgatcattacTTCCAA tccattccattccattccatTAT
	oCR_t_30 FQ SpT2 FP	ataGGTATCACATGACTAAACGA
	oCR_c_13 Iowa Black RQ St1T1 bottom	ataATGGAATGGAATGGAATGGA
SpT2 25 bp SaT1 Dual Beacon	oCR_b_20 Sa PAM 25 bp Sp protospacer FAM	AGGGGTgtaatgatcggtcgtagtctctagaCCT TCGTTTAgTCATGTGATACctat
	oCR_b_15 Sp APM 25 bp Sa protospacer Cy5	GACCGATCATTACACCCCTCTTCC TGTCACCGACACTTCTAT

	oCR_b_17 Iowa Black RQ Sa	ataGAAGTGTCGGTGACAGGAAG
	oCR_b_24 Sp PAM 25 bp	AGGTCTAGAGACTAC
	oCR_t_30 FQ SpT2 FP	ataGGTATCACATGACTAAACGA
SaT1 25 bp ST1T1 Dual Beacon	oCR_b_10 5' Iowa Black Sa Beacon FP	ataGAAGTGTCGGTGACAGGAAG
	oCR_b_12 St1 PAM 25 bp Sa protospacer FAM	TTGGAAgtaatgatcggctgtagtctctagaACC CCTCTTCCTGTCACCGACACTTctat
	oCR_b_23 Sa PAM 25 bp	AGGGGTtctagag
	oCR_b_11 Sa PAM 25 bp St1 protospacer Cy5	actacgaccgatcattacTTCCAATCCATTC CATTCCATTCCATtat
	oCR_c_13 Iowa Balck RQ St1T1 bottom	ataATGGAATGGAATGGAATGGA
Triple Beacon	oCR_b_25 Sp PAM 25 bp Sa Protospacer 25 bp St1 PAM	AGGtctagagactacgaccgatcattacaccctcttc ctgtcaccgacacttcttagctaactagtagtaccatcgattcg aaggtt
	oCR_b_26 St1 Iowa Black RQ	aggttaaggttaaggttaaggttaata

	oCR_b_29 25 bp SpT2 Protospacer FAM	tcggtcGTAGTCTCTAGACCTTCGTTT AGTCATGTGATACCTAT
	oCR_b_28 25 bp Sa Sa PAM 25 bp	ttagctaagaagtgtcggtgacaggaagaggggtgtaa tgatcggtc
	oCR_b_27 25 bp St1 protospacer Cy5	tattaccttaccttaccttacctaaccttcgaatcgaatgta ctag
	oCR_t_30 FQ SpT2 FP	ataGGTATCACATGACTAAACGA