Protein-Mediated Colloidal Assembly

Thesis by

Bradley Ross Silverman

In Partial Fulfillment of the Requirements for
the degree of
Doctor of Philosophy in Chemical
Engineering

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2020
Defended May 29, 2020

ACKNOWLEDGEMENTS

It is a cliché, but no thesis is solely the product of the author. I have often described a PhD as being a marathon, not a sprint, but perhaps a better analogy is a (long) relay race. Thus, I owe a debt of gratitude to many for making my stay at Caltech fruitful, productive, and enjoyable.

First and foremost, I would like to thank my advisor, Prof. David Tirrell. Dave is the paragon of what I feel like an advisor should be. He gave me an incredible amount of intellectual freedom to explore topics that I found interesting, while also giving me the guidance to not go too far down any particular rabbit hole (or occasionally, to be supportive of that somewhat deep dive, for example, the work on pH responsive coils in Chapter 4). Dave also has the unique ability to see problems (if not their solutions) incredibly cleanly, and has helped me greatly to see problems in new and fruitful ways. I hope that my abilities develop in the same way.

I would also like to thank the other members of my committee: Prof. Rustem Ismagilov, Prof. John Brady, and Prof. Michael Elowitz. All of them brought incredibly unique perspectives to my committee meetings, and I always enjoyed our discussions. I would like to particularly acknowledge Rustem’s lively contributions at our biweekly BRICS supergroup meetings, where his feedback related to this work substantially contributed to my ideas during a formative period. Along with Rustem, John and Michael have among the keenest insights of anyone I have encountered and were very helpful in focusing my (sometimes scattered) thinking about how to put a bow on ideas. Beyond my committee members, I’d like to thank my other professors at Caltech, but especially Prof. Zhen-Gang Wang and Dr. Justin Bois who taught me how to approach hard problems in a characteristically rigorous way.

Dave’s other major contribution, of course, is that over the years he has built a lab culture second to none. I’d like to thank all the Tirrell lab members who I have interacted with, past and present, who have all provided valuable insights, perspectives, and friendships over the years. In particular though, I would first like to thank Dr. Peter Rapp, who took me under his wing when I first joined the lab and taught me the joys of microscopy, hydrogels, and polymer physics. I would also like to thank my co-authors on some of the work described in this thesis, Maiko Obana (Chapter 2), and Dr. Mark Kozlowski (Chapter 3), both of whom are amazing scientists and friends (and also good softball players). My office mates over the years, Dr. Katharine Fang, Dr. (soon to be Prof.) Joshua Baccile, and Dr. Mark Kozlowski, were always around to lend a kind ear and their thoughts when my bacteria (again) did not do what I wanted them to.

There are many others at Caltech who helped the work in this thesis become a reality. I would like to acknowledge Dr. Andres Collazo, Steven Wilbert, Dr. Blaise Ndjamen, and Dr. Giada Spigolon for their work with the confocal microscopes in the BIF, which were essential for almost every part of this work. Thanks also to Dr. Mona Shahgoli at the CCE mass spectrometry facility for assistance with the LC-MS done in Chapter 5 of this thesis. I would also like to thank Dr. Kimberly Mayer and Julie Kelly along with Prof. Frances Arnold for their stewardship over the Rosen Bioengineering Center and the Biotechnology Leadership Program, which I was fortunate to take part in during my stay here. Finally, thank you to the people behind the scenes at Caltech who make everything happen smoothly, Kathy Bubash, Allison Kinard, Anne Hormann, Irina Meininger, Elizabeth Garcia, Joe Drew, Armando Villasenor, Greg Rolette, and many others that I have forgotten.

I would like to thank my teachers and professors from before I came to Caltech. Thanks are due most particularly to Prof. Julie Champion at Georgia Tech. Julie began mentoring me even before I started at Georgia Tech, and has been an exemplary scientific mentor ever since then; whether that was teaching me to pipette, or by encouraging me to go to graduate school (and join Dave’s lab), she was instrumental in helping me reach this point. I would also like to thank my colleagues (and friends) in the Champion Lab while I was there, but especially Drs. Lina Herrera-Estrada and Trudy Padmore, who taught me a lot about science and life, and made our work incredibly fun. Finally I would like to gratefully acknowledge here the late Allen Szigety, my middle-school algebra and geometry teacher, who I know would have loved to see me here. He always believed that I was capable of more than I thought, and I miss him.

Finally, thanks to all my friends and family for their support over the years. Whether that was many lunches at the turtle pond (thanks Rachel!), board games (Heidi and Nick), or trivia back in Atlanta (go Team Bullmoose!), my friends have helped keep me sane through the tough times in graduate school. My family have provided so much support (emotional and otherwise) over the years, and I never could have made it here without them. Thanks to my mother for her "substantive" comments on my manuscripts (they always needed to be more rigorous!), my father for his constant refrain of "When are you graduating again?" and my brother for his commiseration about graduate school and complaints about Final Jeopardy wagering (looks like I am graduating before you after all.). Thanks to the chemical engineering cadre (Grandpa Bernie, Uncle Charles, Cousin Jeff, and Brother Adam); I do not know how many other Thanksgiving celebrations include a description of the importance of high Biot numbers when frying a turkey (as well as the proper PPE), but ours do…

Thank you to everyone.

ABSTRACT

The assembly of colloidal-sized particles into larger structures by the manipulation of inter-particle forces has been a subject of significant research towards applications in materials science, soft matter physics, and synthetic biology. To date, much of this work has utilized manipulation of electrostatic or depletion interactions to drive the aggregation of the particles. More recently, specific (bio)-chemical interactions have been harnessed, particularly the use of deoxyribonucleic acid (DNA) linkers to program particle interactions by Watson-Crick base-pairing. In this thesis, we will demonstrate the use of an alternative set of biochemical interactions, protein-protein interactions, which have useful properties (in particular, their ability to be completely genetically-programmable).

In Chapter 2, we discuss the development of a model system for the protein-mediated assembly of colloidal micro-particles. Associative proteins are grafted onto the surface of polystyrene micro-particles, enabling their assembly into aggregates either through reversible coiled-coil interactions or by irreversible isopeptide linkages. The sizes of the resulting aggregates are tunable and can be controlled by the concentration of the immobilized associative proteins on their surface. Further, we show that particles grafted with different protein pairs show excellent self-sorting into separate aggregates. Finally, we demonstrate that these protein-protein interactions can be used to assemble complex core-shell aggregates. The principles of protein-mediated colloidal assembly learned in this chapter will be instructive as we attempt the more complex assembly of living microbial cells.

In Chapter 3, we discuss the implementation of a protein-driven aggregation system in living bacterial cells. Similarly to Chapter 2, we demonstrate that we can drive the aggregation of bacteria by the surface display of proteins enabling reversible coiled-coil interactions or irreversible isopeptide bonds. The sizes of these aggregates are tunable by titration of surface expression levels by standard synthetic biology techniques. Finally, we show that this programmable aggregation of bacteria may have physiological consequences

for the cells, in particular, the activation of a quorum sensing circuit due to a higher local concentration of bacteria.

In Chapter 4, we further investigate how the properties of the aggregates described in Chapter 3 can be controlled and how these relate to the underlying properties of the associative proteins and shear field. we demonstrate control of the assembly kinetics and equilibrium sizes of the resulting flocs over several orders of magnitude using different associating proteins and expression levels. Finally, we show that a single point mutation in the associative protein leads to an unexpected ultra-sensitive pH-responsive coil, demonstrating the importance of molecular-scale interactions on the macro-scale properties of the aggregates.

In Chapter 5, we discuss the ability of the bacterial aggregates described in Chapters 3 and 4 to enable substrate channeling between bacterial strains, leading to enhancement of titers in multi-step biosynthetic pathways. When biosynthetic pathways are split into separate bacterial strains, dilution of the intermediate compound into the bulk media may decrease reaction flux. By aggregating the bacteria, the intermediate compound is able to rapidly diffuse into the downstream cell without being diluted, enabling higher reaction fluxes. we demonstrate through the model flavonoid synthesis pathway that aggregation can lead to substantially higher titers of the desired compound without pathway re-engineering, and develop a mathematical model by which this result can be understood.

PUBLISHED CONTENT AND CONTRIBUTIONS

M. Obana, B. R. Silverman and D. A. Tirrell, Protein-mediated colloidal assembly, *J. Am. Chem. Soc.*, 2017, **139**, 14251–14256 doi: 10.1021/jacs.7b07798

Reused with permission from the American Chemical Society. B.R.S contributed new protein reagents, performed experiments, performed quantitative image analysis, and participated in the writing of the manuscript.

* Co-first authors

M. T. Kozlowski*, B. R. Silverman*, C. P. Johnstone, and D. A. Tirrell. Genetically programmable control of microbial assembly. In preparation.

B.R.S. contributed new experimental bacterial strains, performed experiments, modeled quorum sensing reaction diffusion system, performed quantitative image analysis, and participated in the writing of the manuscript.

* Co-first authors

B. R. Silverman and D. A. Tirrell, Control of engineered microbial flocs. In preparation.

B.R.S. contributed new experimental bacterial strains and protein reagents, performed experiments, performed quantitative image analysis, and participated in the writing of the manuscript.

B. R. Silverman and D. A. Tirrell, Genetically programmable bacterial assembly enhances microbial production of specialty chemicals. In preparation.

B.R.S. contributed new experimental bacterial strains, performed experiments, analyzed data, and participated in the writing of the manuscript.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS AND/OR TABLES

Chapter 1

INTRODUCTION: PROGRAMMABLE COLLOIDAL ASSEMBLY AND SPATIAL STRUCTURE IN MICROBIAL COMMUNITIES

The overall purpose of the work described in this thesis is to develop a new set of tools to program and drive assembly of colloidal particles using protein-protein interactions, with the major application of enabling the genetically programmable assembly of living bacterial cells in order to enable emergent function of multi-cell bacterial aggregates. In this chapter, we discuss previous work in programmable colloidal assembly, and the importance of spatial structure in microbial consortia.

Colloidal Forces and Flocculation

A colloid is a substance consisting of a dispersed phase of particles, generally between 1 nanometer and several microns in size, suspended in a continuous solvent phase.[1] In this dispersed state, the repulsive interactions between colloidal particles must be greater in magnitude than any attractive forces between particles. In many cases, these repulsive forces are dominated by relatively long-range electrostatic interactions, which are generally characterized by the ζ-potential (zeta-potential), the potential in the electrical double-layer at the interface at the slip plane of the particle with respect to the bulk continuous phase. At high absolute values of the zeta-potential, particle suspensions are stabilized by electrostatic repulsion between particles of the same charge.[2] Flocculation, the assembly of colloidal particles into aggregates, takes place when attractive interactions dominate over the repulsive interactions. Flocculation may occur from the dispersed phase when the repulsive interactions between particles are decreased, such as by the addition of salt, which will screen the electrostatic repulsion between particles,[2,3] or by increasing the attractive forces between particles, for example by introducing particles with the opposite charge into the solution,[4] adding a multi-functional chemical crosslinker that bridges particles,[5] or by the depletion force which is caused by macromolecules creating an entropically-derived osmotic pressure between particles.[6,7]

Colloidal Crystals

When the attractive forces between particles overcome the repulsive interactions, flocculation occurs, but does not in general lead to the thermodynamic product, a colloidal crystal.[8] A colloidal crystal is an ordered structure (analogous to a molecular crystal), where colloidal particles are arranged in such a way as to lead to long-range order of the assembly. For monodisperse colloidal spheres, the close-packed crystalline structure is the face-centered cubic (fcc) crystal lattice, but more complex crystalline structures may also be obtained by utilizing electrostatic attractive forces between two different types of particles,[9] by using bidisperse-sized particles,[10] or by using non-spherical particles.[11] However, in all cases when attractive interactions are much stronger than the thermal energy scale k_BT, formation of particle aggregates are essentially irreversible, and are not able to anneal to the lowest energy, crystalline structure, and thus a metastable disordered colloidal aggregate is obtained.[12] In this case, particle aggregates will generally exhibit a fractal structure characteristic of diffusion-limited cluster aggregation (DLCA), so called because associative forces are strong enough that only the rate of diffusion limits the size of aggregates.[13] Thus, in order to form colloidal crystals, which are useful in many applications,[8,14] the attractive forces must be carefully tuned. In many cases, in order to form structures with a high degree of long-range order, complex heating and cooling regimens must be performed to push the assemblies out of metastable energetic minima.[15] It is particularly difficult to obtain colloidal crystals for larger, micron-scale particles because association energies are often of larger magnitude, so annealing is much more difficult.[16,17]

Programmable Colloidal Assembly

Because there is a great deal of interest in creating complex colloidal structures with long-range order, much effort has gone into programming additional interactions between particles that enable the precise control of inter-particle forces. Approximately two decades ago, Mirkin and co-workers developed methods for DNA-programmable assembly of colloidal particles into colloidal crystals.[18] In this set of techniques, DNA is grafted onto colloidal particles, which can then program inter-particle interactions by Watson-Crick base-pairing between complementary DNA strands on different particles (**Figure 1.1**). Because DNA hybridization is highly predictable, large numbers of interactions are able to be programmed into the colloids. Additionally, because the association energy of Watson-Crick base pairing is well-understood, and is essentially a linear function of the length and composition, it is relatively simple to tune the association energies to achieve a desired result.[19] A large number of studies have resulted, enabling the production of many different crystalline architectures,[18,20] dynamic colloidal phase transitions based on DNA strand displacement,[21] and assembly of particles consisting of many different polymeric[22,23] and metallic materials.[24,25]

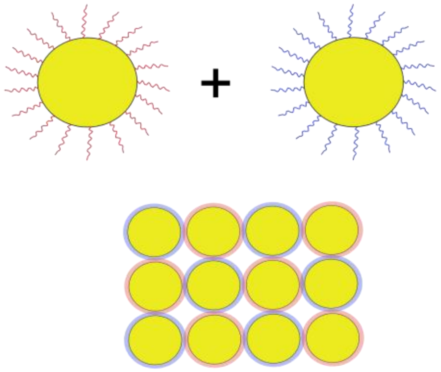


Figure 1.1: DNA Mediated Assembly: by grafting DNA molecules to colloidal particles (often gold nanoparticles), specific interactions can be programmed using Watson-Crick base-pairing. Because many strands are grafted onto each particle, each particle may bind to several other particles. If the

energy is O(kT), then a colloidal crystal with long-range order may form. For this, the pairing region of each DNA oligonucleotide is often only 4-6 base-pairs.

The great deal of work in DNA-coated colloids has led to a large amount of conceptual understanding of the mechanisms of crystallization of colloidal particles, their thermodynamics, and colloidal physics in general. However, all DNA-programmable colloidal assemblies share certain properties, including high sensitivity to temperature and nucleases and a narrow range of association free energies.[26] Thus, there is a need to develop alternative programmable elements for colloidal assembly. One method is the use of "patchy" Janus colloids where small parts of colloidal particles are functionalized with complementary chemical groups enabling the specific association of particles.[1] Another method is to use anisotropic particles whereby the shapes of the particles themselves direct the assembly of the desired super-structure through steric interactions.[27,28] However, both of these methods suffer from the inability to direct many orthogonal interactions in a way similar to the promise of DNA-programmable assembly.

Instead, by analogy with DNA-programmable assembly, some groups have begun to experiment with protein-programmable assembly.[26] Protein-programmable assembly enables the use of a large variety of protein-protein interactions that have evolved over millions of years to enable assembly of proteins into complexes. In theory, because many of these proteins exist in a complex cellular milieu, they are highly evolved to be specific, and they span a large set of interaction energies. Much of the work with protein-programmable colloidal assembly has utilized the coiled-coil protein motif.[29,30] Coiled coils (or helical bundles) are a common protein interaction domain, in which alpha helices are held together by internal hydrophobic interactions, usually leucine residues at their internal positions.[31] Due to their simple structures and design rules, coiled coils are highly amenable to synthetic design, and many examples of coiled coils have been computationally designed and used for various applications.[32,33] By grafting coiled-coils onto the surface of nanoparticles,[34] microparticles,[30] or by adding soluble coiled-coil linker proteins,[35] researchers have been able to program the specific assembly of colloidal particles. Other protein-protein or protein-ligand interactions have also been used to drive assembly of nanoparticles including the biotin-streptavidin interaction,[36] and the barnase-barnstar interaction.[37] Multi-step assembly of hierarchically structured inorganic materials (nanoflowers) via coiled-coil and electrostatic interactions has also been demonstrated.[38] Because the structures of proteins vary so widely, interaction affinities may vary over many orders of magnitude. In addition, in an analogous way for DNA-mediated interactions, dis-assembly can often be programmed by addition of soluble competitor proteins (similar to the use of strand displacement reactions for DNA).[30]

However, to date, protein-mediated colloidal assemblages have not yet achieved the same degree of long-range order obtained from DNA-mediated colloidal assemblies. This is because even though protein interaction energies span such a large range, they are still uniformly stronger than the DNA linkers utilized in DNA-driven assembly. In addition, the techniques developed for annealing of DNA-mediated colloids, especially temperature-driven melting and annealing, generally do not work with protein-mediated interactions, as these interactions are much less strongly influenced by temperature.

In Chapter 2 of this thesis, we describe our results for the protein-mediated control of colloidal assemblies. We demonstrate the use of a heterodimeric designed coiled coil[32] and the SpyTag-SpyCatcher[39] interaction to drive well-controlled colloidal assembly (and dis-assembly) of colloidal micro-particles.

Microbial Consortia and Biofilms

Though the idealized picture of bacteria is of free-living (also called planktonic) organisms swimming through liquid media (and much microbiological research has centered on this setting), in their natural environment, bacteria much more often live in multi-cellular communities, often in biofilms surrounded by a tough polymeric extracellular matrix.[40,41] Biofilms are assemblages of (often multi-species) microbes that live together symbiotically, often with a high degree of spatial organization. Biofilms represent a significant medical challenge, particularly in the hospital setting, where biofilms may form on many surfaces in the hospital, and assist in spreading pathogens to patients.[42] Because biofilms consist of bacteria surrounded by a mechanically tough extracellular matrix, they are often difficult to remove from surfaces, and may require both chemical and mechanical treatment to remove effectively.[43] Bacterial biofilms are also present in animals (for example, the oral microbiome contains biofilms on the enamel of teeth), in soil, in hot springs, and in many other natural (and artificial) environments.[40]

Emergent Properties of Bacterial Communities

The ubiquity of the sorts of multi-organism communities described above implies that the organisms must obtain substantial fitness benefits from living in these communities. The simplest advantages result simply from multiple species living in the same environment, but not necessarily in a biofilm or aggregate (i.e. where diffusion of metabolites is not limited by the environment). In this circumstance, the different species of bacteria may benefit from the ability to share metabolites that they produce/consume in a division of labor-type approach. For example, one bacterium may be capable of fixing carbon dioxide into organic compounds, while another may be capable of nitrogen fixation. If both bacterial species are grown together, the resulting consortium will be able to fix both carbon and nitrogen, enabling it to be more metabolically flexible.[44] Different species may also compete in such an environment, whether that is passively by competing for limited resources, or actively by secreting toxins that actively disrupt the competitor cells.[45]

However, many of the emergent properties of microbial consortia result from the bacteria living in closely packed, highly organized biofilms or aggregates. For example, relevant to the hospital context, biofilms are highly resistant to antibiotics and antiseptics.[46] There are several reasons for this lack of effectiveness which are highly instructive for the work described in this thesis. First, due to the tightly packed nature of biofilms, the diffusion coefficients of small molecule substrates are somewhat decreased in biofilms, which will decrease the diffusive flux of antibiotics into the interior of biofilms.[47] However, this decrease of diffusion coefficient is not sufficient to account for the significantly lower concentrations of antibiotics found in biofilms. Instead, researchers have found that antibiotics may bind to biofilm extracellular matrix components, which has the effect of decreasing the diffusion rates into the biofilm even further.[48] In another case, researchers have found that chlorine actually appears to be consumed by cells in the biofilm.[49] Finally, some researchers have suggested that the resistance of biofilms towards antibiotic agents may be related to physiological changes of the cells in the biofilms, for example the induction of a stress response or slow growth state.[46] The restricted diffusion of substrates inside aggregates is utilized in Chapter 3 to enhance quorum sensing in engineered bacterial aggregates, while the consumption of molecules inside aggregates is used to enhance titers of biosynthetic pathways in Chapter 5.

Another property of spatially organized biofilms and aggregates is their ability to sense and signal their surroundings, enabling them to act in concert. The most common way for bacteria to do this is using the “quorum sensing” system (**Figure 1.2**).[50] In the canonical quorum sensing pathway, bacteria are able to both secrete and sense small molecules called autoinducers. When bacteria sense autoinducer in the environment, they will both produce more autoinducer (such that the production of autoinducer is under positive feedback), and, when the concentration is high enough, engage in high cell-density specific activities (often production of virulence factors, motility genes, biofilm synthesis genes, etc.).[50–52] When bacteria are found in a biofilm, their ability to quorum sense is enhanced, as the autoinducer is more easily retained in the biofilm (compared to a planktonic culture), and thus is more able to accumulate. In addition to within-species signaling, many bacteria can also recognize a cross-species autoinducer (AI-2), potentially enabling cross-species communication, though this is controversial.[53,54] In the past several years, bacteria in biofilms have also been found to utilize electrical signaling.[55] In these studies, researchers have found that expression of bacterial ion channels and subsequent membrane depolarization enables long-range signaling in an analogous manner to eukaryotic neurons.

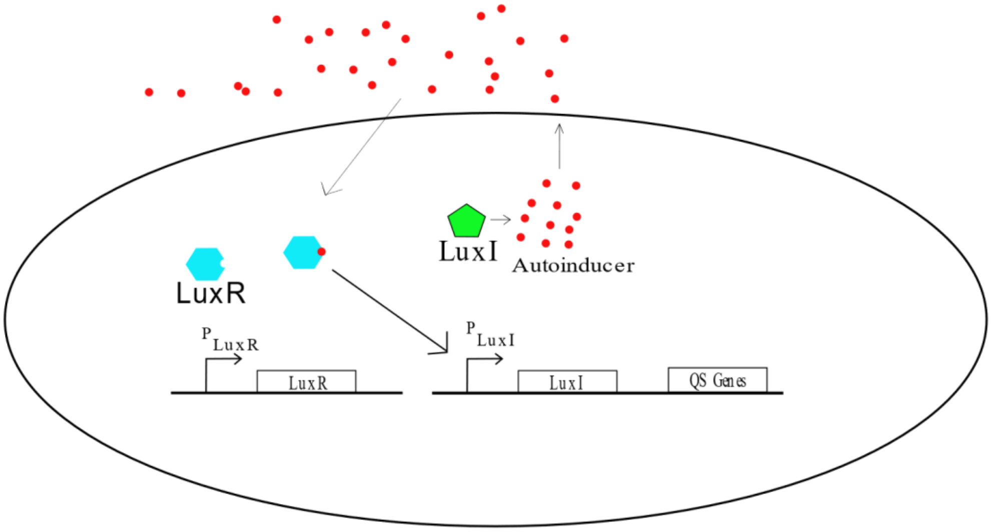


Figure 1.2 Quorum Sensing: in a canonical quorum sensing system (here the LuxI/LuxR system from *Vibrio fischeri*), the bacteria constitutively make a response regulator protein LuxR and an autoinducer synthesis protein LuxI. When the autoinducer is at high concentration in the cell, it binds to LuxR, activating it, leading to activation of the P_{LuxI} promoter and higher expression of LuxI, and subsequently activation of other quorum sensing genes. The autoinducer is cell-permeable, so it will diffuse to other cells, leading to concerted behavior of a cellular community.

Finally, an additional property of biofilms (which is mechanistically related to the resistance to antimicrobials), is the ability of biofilms to create gradients of chemicals due to consumption or synthesis of substrates.[56] This sort of gradient formation is not possible in a well-mixed system. A descriptive example of this property (**Figure 1.3**) is the creation of a hypoxic (or anaerobic) zone in the interior of a biofilm or aggregate. If aerobic (oxygen-consuming) bacteria are present on the outside of a biofilm, they will consume much or all of the oxygen present in the environment, creating an anerobic micro-environment in an otherwise normoxic environment. This may enable obligate anaerobic microbes to live in the film. Similarly, if bacteria in the biofilm are making a metabolite, this metabolite will accumulate in the interior of the biofilm such that the inside of the biofilm will have much higher accumulation than the surface (from which the metabolite will diffuse into the bulk solution). These gradients will accumulate regardless of any gradient in the actual production of the metabolite.

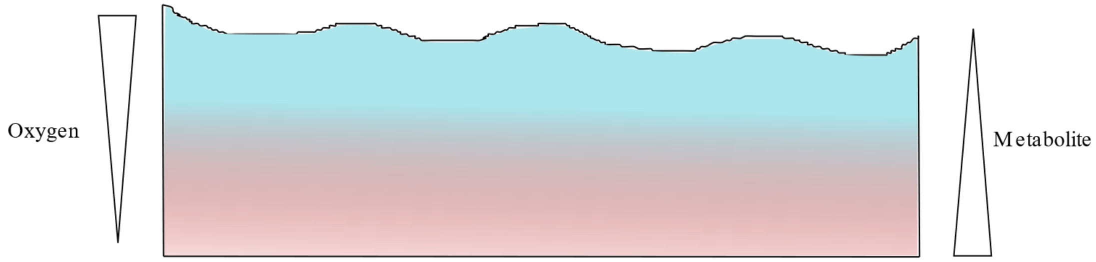


Figure 1.3. Biofilm Gradients: in biofilms, gradients of both oxygen and metabolites can form due to consumption and/or production of the species by cells in the biofilm. In some cases, this could lead to essentially anaerobic conditions in the interior of biofilms.

Imposing Spatial Structure on Cells

Because the spatial structure of groups of cells is so important for both natural and engineered systems, many groups have attempted to impose spatial structure on both prokaryotic and eukaryotic cells. Broadly, these techniques can be separated into two approaches, top-down and bottom-up assembly.[57] In top-down assembly, smaller structures are assembled by starting with a bulk material and then decreasing the size of the structure by various physical techniques. Exemplars of these processing technologies are extrusion, milling, and emulsification. By contrast, in bottom-up assembly, smaller structures are brought together to form larger structures. This is the same colloidal assembly process discussed above.

Both bottom-up and top-down methods have been used for the synthesis of bacterial communities. An exemplar of the latter set of methods is one by Ismagilov and co-workers where an extrusion process was used to form a core-shell type community, in which an interior species was protected from environmental insult by the consumption of a toxin by the exterior species.[58] In other work, the Ismagilov group used microfluidic technology to impose spatial structure on a group of three microorganisms to enable them to survive nutrient starvation and antibiotic stress.[59] In both cases, the use of defined spatial structure enabled emergent function that would not have been possible in simple mixed culture.

Bottom-up assembly has also been used to assemble multi-celled aggregates. In one exemplary study, Bertozzi and co-workers grafted DNA strands onto azide-modified sugars on the surface of mammalian cells using click chemistry.[60,61] By grafting strands with complementary sequences, the researchers could program cell-cell interactions and generate tissues with well-defined cellular connectivity. Although the sizes of the resulting clusters were small (generally less than 10 cells) and the technique requires in-situ chemical modification of the cells, this still represents an important advancement, and one that is highly relevant to our attempts at generating multicellular aggregates in Chapters III, IV, and V of this thesis.

In another more relevant example, Riedel-Kruse and coworkers genetically engineered microbes to surface-display associative nanobodies on their surfaces, driving specific aggregation of small clusters.[62,63] This is similar in approach to our efforts in later chapters of this thesis. This approach has the advantage of being purely genetically programmable; the programmable interactions are stored entirely in the organisms’ genetic information and can be manipulated using standard molecular biology techniques. The techniques we describe in this thesis represent an alternative set of protein-protein interactions, including the previously unused SpyTag-SpyCatcher interaction.[39] We will further demonstrate how we can understand the properties of the resulting aggregates from the molecular scale interactions of the underlying associative proteins, and then how we can use these aggregates to enhance the ability of co-cultured strains to produce desired substances in metabolic engineering through intermediate channeling between bacteria in aggregates.

References

1 F. Li, D. P. Josephson and A. Stein, Colloidal assembly: the road from particles to colloidal molecules and crystals., *Angew. Chemie*, 2011, **50**, 360–88.

2 D. Hanaor, M. Michelazzi, C. Leonelli and C. C. Sorrell, The effects of carboxylic acids on the aqueous dispersion and electrophoretic deposition of ZrO2, *J. Eur. Ceram. Soc.*, 2012, **32**, 235–244.

3 Y. Adachi, Dynamic aspects of coagulation and flocculation, *Adv. Colloid Interface Sci.*, 1995, **56**, 1–31.

4 M. E. Leunissen, C. G. Christova, A.-P. Hynninen, C. P. Royall, A. I. Campbell, A. Imhof, M. Dijkstra, R. van Roij and A. van Blaaderen, Ionic colloidal crystals of oppositely charged particles., *Nature*, 2005, **437**, 235–40.

5 A. K. Boal, F. Ilhan, J. E. Derouchey, T. Thurn-albrecht, T. P. Russell and V. M. Rotello, Self-assembly of nanoparticles into structured spherical and network aggregates, 2000, **404**, 746–748.

6 G. Dorken, G. P. Ferguson, C. E. French and W. C. K. Poon, Aggregation by depletion attraction in cultures of bacteria producing exopolysaccharide, *J. R. Soc. Interface*, 2012, **9**, 3490–3502.

7 S. Asakura and F. Oosawa, On interaction between two bodies immersed in a solution of macromolecules, *J. Chem. Phys.*, 1954, **22**, 1255.

8 O. D. Velev and S. Gupta, Materials fabricated by micro- and nanoparticle assembly - the challenging path from science to engineering, *Adv. Mater.*, 2009, **21**, 1897–1905.

9 A. Rugge and S. H. Tolbert, Effect of electrostatic interactions on

crystallization in binary colloidal films, *Langmuir*, 2002, **18**, 7057–7065.

10 Z. Zhou, Q. Yan, Q. Li and X. S. Zhao, Fabrication of binary colloidal crystals and non-close-packed structures by a sequential self-assembly method, *Langmuir*, 2007, **23**, 1473–1477.

11 M. Rycenga, J. M. McLellan and Y. Xia, Controlling the assembly of silver nanocubes through selective functionalization of their faces, *Adv. Mater.*, 2008, **20**, 2416–2420.

12 J. Dhont, C. Smits and H. Lekkerkerker, A time resolved static light scattering study on nucleation and crystallization in a colloidal system, *J. Colloid Interface*, 1992, **152**, 386-401

13 T. A. Witten, Diffusion-limited aggregation, a kinetic critical phenomenon, *Phys. Rev. Lett.*, 1981, **47**, 1400–1402.

14 N. Rosi and C. Mirkin, Nanostructures in biodiagnostics, *Chem. Rev.*, 2005, **105**, 1547–1562.

15 L. Di Michele, F. Varrato, J. Kotar, S. H. Nathan, G. Foffi and E. Eiser, Multistep kinetic self-assembly of DNA-coated colloids., *Nat. Commun.*, 2013, **4**, 2007.

16 V. Kodali, W. Roos, J. Spatz and J. Curtis, Cell-assisted assembly of colloidal crystallites, *Soft Matter*, 2007, 337–348.

17 Y. Wang, Y. Wang, X. Zheng, É. Ducrot, J. S. Yodh, M. Weck and D. J. Pine, Crystallization of DNA-coated colloids., *Nat. Commun.*, 2015, **6**, 7253.

18 C. Mirkin, R. Letsinger, R. Mucic and J. Storhoff, A DNA-based method for rationally assembling nanoparticles into macroscopic materials, *Nature*, 1996,

382, 607-609

19 J. N. Zadeh, C. D. Steenberg, J. S. Bois, B. R. Wolfe, M. B. Pierce, A. R. Khan, R. M. Dirks and N. A. Pierce, Software news and updates NUPACK : analysis and design of nucleic acid systems, *J. Comput. Chem.*, 2010.

20 R. J. Macfarlane, B. Lee, M. R. Jones, N. Harris, G. C. Schatz and C. A. Mirkin, Nanoparticle superlattice engineering with DNA, *Science*, 2011, **334**, 204–209.

21 W. B. Rogers and V. N. Manoharan, Programming colloidal phase transitions with DNA strand displacement, *Science*, 2015, **347**, 639–642.

22 Y. Wang, Y. Wang, X. Zheng, É. Ducrot, M. G. Lee, G. R. Yi, M. Weck and D. J. Pine, Synthetic strategies toward DNA-coated colloids that crystallize, *J. Am. Chem. Soc.*, 2015, **137**, 10760–10766.

23 A. J. Kim, P. L. Biancaniello and J. C. Crocker, Engineering DNA-mediated colloidal crystallization, *Langmuir,* 2006, **22**, 1991–2001.

24 Y. Zhang, F. Lu, K. G. Yager, D. Van Der Lelie and O. Gang, A general strategy for the DNA-mediated self-assembly of functional nanoparticles into heterogeneous systems, *Nat. Nanotechnol.*, 2013, **8**, 865–872.

25 J. S. Lee, A. K. R. Lytton-Jean, S. J. Hurst and C. A. Mirkin, Silver nanoparticle - oligonucleotide conjugates based on DNA with triple cyclic disulfide moieties, *Nano Lett.*, 2007, **7**, 2112–2115.

26 S. Mann, W. Shenton and M. Li, Biologically programmed nanoparticle assembly, *Adv. Mater.*, 2000, **4**, 147–150.

27 P. F. Damasceno, M. Engel and S. C. Glotzer, Crystalline assemblies and

densest packings of a family of truncated tetrahedra and the role of directional entropic forces, *ACS Nano*, 2012, **6**, 609–614.

28 S. Sacanna, M. Korpics, K. Rodriguez, L. Colón-Meléndez, S. H. Kim, D. J. Pine and G. R. Yi, Shaping colloids for self-assembly, *Nat. Commun.*,2013 ,**4**

29 M. M. Stevens, N. T. Flynn, C. Wang, D. A. Tirrell and R. Langer, Coiled-coil peptide-based assembly of gold nanoparticles, *Adv. Mater.*, 2004, **16**, 915–918.

30 A. P. Schoen, B. Hommersom, S. C. Heilshorn and M. E. Leunissen, Tuning colloidal association with specific peptide interactions, *Soft Matter*, 2013, **9**, 6781-6785.

31 J. M. Mason and K. M. Arndt, Coiled coil domains: Stability, specificity, and biological implications., *Chembiochem*, 2004, **5**, 170–176.

32 A. W. Reinke, R. A. Grant and A. E. Keating, A synthetic coiled-coil interactome provides heterospecific modules for molecular engineering, *J. Am. Chem. Soc.*, 2010, **132**, 6025–6031.

33 D. Woolfson, The design of coiled-coil structures and assemblies, *Adv. Protein Chem.*, 2005, **70**, 79–112.

34 M. M. Stevens, N. T. Flynn, C. Wang, D. A. Tirrell and R. Langer, Coiled-coil peptide-based assembly of gold nanoparticles, *Adv. Mater.*, 2004, **16**, 915–918.

35 M. G. Ryadnov, B. Ceyhan, C. M. Niemeyer and D. N. Woolfson, “ Belt and Braces ”: a peptide-based linker system of de novo design, *J. Am. Chem. Soc.*, 2003, **125**, 9388–9394.

36 S. A. Connolly and D. J. Fitzmaurice, Programmed assembly of gold nanocrystals in aqueous solution, *Adv. Mater.*1999, **11**, 1202-1215

37 M. P. Nikitin, T. A. Zdobnova, S. V. Lukash, O. A. Stremovskiy and S. M. Deyev, Protein-assisted self-assembly of multifunctional nanoparticles, *Proc. Natl. Acad. Sci*, 2010, **107**, 5827–32.

38 W. M. Park and J. A. Champion, Colloidal assembly of hierarchically structured porous supraparticles from flower-shaped protein-inorganic hybrid nanoparticles, *ACS Nano*, 2016, **10**, 8271–80.

39 B. Zakeri, J. Fierer, E. Celik, E. Chittock, U. Schwarz-Linek, V. Moy and M. Howarth, Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin, *Proc. Natl. Acad. Sci.*,2012 , **109,** 12

40 L. Hall-Stoodley, J. W. Costerton and P. Stoodley, Bacterial biofilms: From the natural environment to infectious diseases., *Nat. Rev. Microbiol.*, 2004, **2**, 95–108.

41 S. Elias and E. Banin, Multi-species biofilms: Living with friendly neighbors., *FEMS Microbiol. Rev.*, 2012, **36**, 990–1004.

42 J. D. Bryers, Medical biofilms *Biotechnol. Bioeng.*, 2008, **100**, 1–18.

43 S. L. Percival, L. Suleman, C. Vuotto and G. Donelli, Healthcare-associated infections, medical devices and biofilms: Risk, tolerance and control, *J. Med. Micro. Biol.*, 2015, **64,** 323-334

44 G. D’Souza, S. Shitut, D. Preussger, G. Yousif, S. Waschina and C. Kost, Ecology and evolution of metabolic cross-feeding interactions in bacteria, *Nat. Prod. Rep.*, 2018, **35**, 455–488.

45 C. D. Nadell, K. Drescher and K. R. Foster, Spatial structure, cooperation, and competition in biofilms, *Nat. Rev. Microbiol.*, 2016, **14**, 589–600.

46 T. Mah and G. O'Toole, Mechanisms of biofilm resistance to antimicrobial agents, *Trends Microbiol.*, 2001, **9**, 34–39.

47 P. S. Stewart, Diffusion in biofilms, *J. Bacteriol.*, 2003, **185**, 1485–1491.

48 P. A. Suci, M. W. Mittelman, F. P. Yu and G. G. Geesey, Investigation of ciprofloxacin penetration into Pseudomonas aeruginosa biofilms, *Antimicrob. Agents Chemother.*, 1994, **38**, 2125–33.

49 D. De Beer, R. Srinivasan and P. S. Stewart, Direct measurement of chlorine penetration into biofilms during disinfection, *Appl. Environ. Microbiol.*, 1994, **60**, 4339–4344.

50 M. Miller and B. Bassler, Quorum sensing in bacteria, *Annu. Rev. Microbiol*, 2001, **55**, 165-199

51 C. A. Swofford, N. Van Dessel and N. S. Forbes, Quorum-sensing Salmonella selectively trigger protein expression within tumors, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 3457–62.

52 B. K. Hammer and B. L. Bassler, Quorum sensing controls biofilm formation in Vibrio cholerae, *Mol. Microbiol.*, 2003, **50**, 101–104.

53 F. Rezzonico and B. Duffy, Lack of genomic evidence of AI-2 receptors suggests a non-quorum sensing role for luxS in most bacteria, *BMC Microbiol.*, 2008, **8,** 154

54 S. P. Diggle, A. Gardner, S. A. West and A. S. Griffin, Evolutionary theory of bacterial quorum sensing: When is a signal not a signal, *Philos. Trans. R. Soc.*

B Biol. Sci., 2007, **362**, 1241–1249.

55 A. Prindle, J. Liu, M. Asally, S. Ly, J. Garcia-Ojalvo and G. M. Süel, Ion channels enable electrical communication in bacterial communities, *Nature*, 2015, **527**, 59–63.

56 H. C. Flemming, J. Wingender, U. Szewzyk, P. Steinberg, S. A. Rice and S. Kjelleberg, Biofilms: An emergent form of bacterial life, *Nat. Rev. Microbiol.*, 2016, **14**, 563–575.

57 G. W. Crabtree and J. L. Sarrao, Opportunities for mesoscale science, *MRS Bull.*, 2012, **37**

58 H. J. Kim, W. Du and R. F. Ismagilov, Complex function by design using spatially pre-structured synthetic microbial communities: Degradation of pentachlorophenol in the presence of Hg(II), *Integr. Biol.*, 2011, **3**, 126–33.

59 H. J. Kim, J. Q. Boedicker, J. W. Choi and R. F. Ismagilov, Defined spatial structure stabilizes a synthetic multispecies bacterial community, *Proc Natl. Acad. Soc* 2008, **105**, 18188–18193

60 R. A. Chandra, E. S. Douglas, R. A. Mathies, C. R. Bertozzi and M. B. Francis, Programmable cell adhesion encoded by DNA hybridization, *Angew. Chem. Int. Ed. Engl.*, 2006, **45**, 896–901.

61 Z. J. Gartner and C. R. Bertozzi, Programmed assembly of 3-dimensional microtissues, *Proc. Natl. Acad. Sci.*, 2009,

62 D. S. Glass and I. H. Riedel-Kruse, A synthetic bacterial cell-cell adhesion toolbox for programming multicellular morphologies and patterns, *Cell*, 2018, **174**, 649-658

63 X. Jin and I. H. Riedel-Kruse, Biofilm lithography enables high-resolution cell patterning via optogenetic adhesin expression, *Proc. Natl. Acad. Sci.*, 2018, 201720676.

Chapter 2

PROTEIN-PROGRAMMABLE ASSEMBLY OF PASSIVE COLLOIDAL PARTICLES

Abstract

Programmable colloidal assembly enables the creation of mesoscale materials in a bottom-up manner. Although DNA oligonucleotides have been used extensively as the programmable units in this paradigm, proteins, which exhibit more diverse modes of association and function, have not been widely used to direct colloidal assembly. Here we use protein-protein interactions to drive controlled aggregation of polystyrene microparticles, either through reversible coiled-coil interactions or through intermolecular isopeptide linkages. The sizes of the resulting aggregates are tunable and can be controlled by the concentration of immobilized surface proteins. Moreover, particles coated with different protein pairs undergo orthogonal assembly. We demonstrate that aggregates formed by association of coiled-coil proteins, in contrast to those linked by isopeptide bonds, are dispersed by treatment with chemical denaturants or soluble competing proteins. Finally, we show that protein-protein interactions can be used to assemble complex core–shell aggregates. This work illustrates a versatile strategy for engineering colloidal systems for use in materials science and biotechnology.

Introduction

Assembly of colloidal particles into mesoscale aggregates has been a topic of continuing interest because of its relevance to applications in photonics, drug delivery, and synthetic biology.[1-4] Many strategies have been used to drive assembly of colloidal particles, including evaporation,[5,6] depletion,[7] and electrostatic interactions.[8,9] Recently, research on colloidal assembly has focused on the use of specific interactions between grafted biomolecules to direct the assembly of particles into more complex architectures.[10] DNA oligonucleotides have been used extensively for this purpose because inter-particle interactions can be programmed simply and directly through Watson-Crick base-pairing.[11-14] Many researchers have used this approach to form colloidal crystals and aggregates.[11-14] By tuning the relative sizes of the underlying colloidal particles and complementarity of the oligonucleotides, researchers have been able to create a striking variety of ordered super-lattices.[12]

An alternative approach is to use associative proteins or peptides to program colloidal assembly.[15-17] Stevens and coworkers demonstrated the assembly of gold nanoparticles using coiled-coil peptide domains as associative units.[17] Coiled-coil domains are especially well suited to the task of directing colloidal assembly; they form helical bundles of small and predictable aggregation number, they are relatively easily engineered through variation in amino acid sequence,[18,19] and many examples in the natural world can serve as starting points for new designs.[20] Self-assembly of nanoparticles by homo-oligomeric[16] and hetero-dimeric[17] coiled-coil peptides grafted to particle surfaces, as well as by peptide linkers,[21] has been reported. Multistep assembly of hierarchically structured inorganic nanoparticles has also been shown.[22]

Work on protein-mediated assembly of micron-sized colloidal particles has been more limited. Schoen and coworkers used self-associating coiled-coil peptides to drive formation of small (~20 particle) clusters.[15] They were able to reverse cluster formation by addition of excess soluble peptide. Deyev and coworkers have also used the barnase-barstar interaction to form complex structures that span multiple length scales.[23,24]

Here we examine two systems for protein-mediated assembly of colloidal microparticles – one based on a pair of high-affinity coiled-coil proteins,[19] the other on the SpyTag/SpyCatcher system, which forms covalent isopeptide bonds between associative protein domains.[25] We show that these interactions can be used to drive large-scale aggregation of particles and to determine the conditions that allow dispersal of aggregates. Aggregate size can be controlled by stoichiometry or by competition with soluble peptide. Finally, we show that these interactions can be used to create complex architectures such as core-shell aggregates. The strategies described here should enable the assembly of particulate and cellular systems for applications in catalysis, drug delivery and tissue engineering.[26]

Results and Discussion

Design of Associative Proteins

We employed two pairs of associative protein domains, designated SYNZIP17/SYNZIP18 (Z17/Z18) and SpyTag/SpyCatcher, in this study (Figure 2.1b). Z17 and Z18 are coiled-coil peptides derived from the SYNZIP library introduced by Keating and coworkers,[19] and are reported to form anti-parallel coiled-coil dimers with high (<10 nM) affinity. SpyTag and SpyCatcher were derived from the Streptococcus pyogenes fibronectin-binding protein FbaB by Howarth and coworkers,[25] who showed that association of the two domains leads to formation of an isopeptide bond between a lysine residue in SpyCatcher and an aspartic acid residue in SpyTag. The SpyTag/SpyCatcher interaction has been used to prepare cross-linked hydrogels,[27] to control protein topology,[28] to analyze expression and localization of proteins,[29,30] and to create long, extended proteins by linking together multiple polypeptides.[31] Each of the associative protein domains was genetically fused at its C-terminus to an elastin-like polypeptide bearing a C-terminal cysteine residue for site-specific conjugation to particle surfaces. Hexahistidine tags were added to N- and C-termini of each protein to facilitate purification. These proteins, along with a control protein containing no associative domain (denoted E), were expressed in Escherichia coli and purified by affinity chromatography. Yields of purified protein were in all cases at least 50 mg/L.

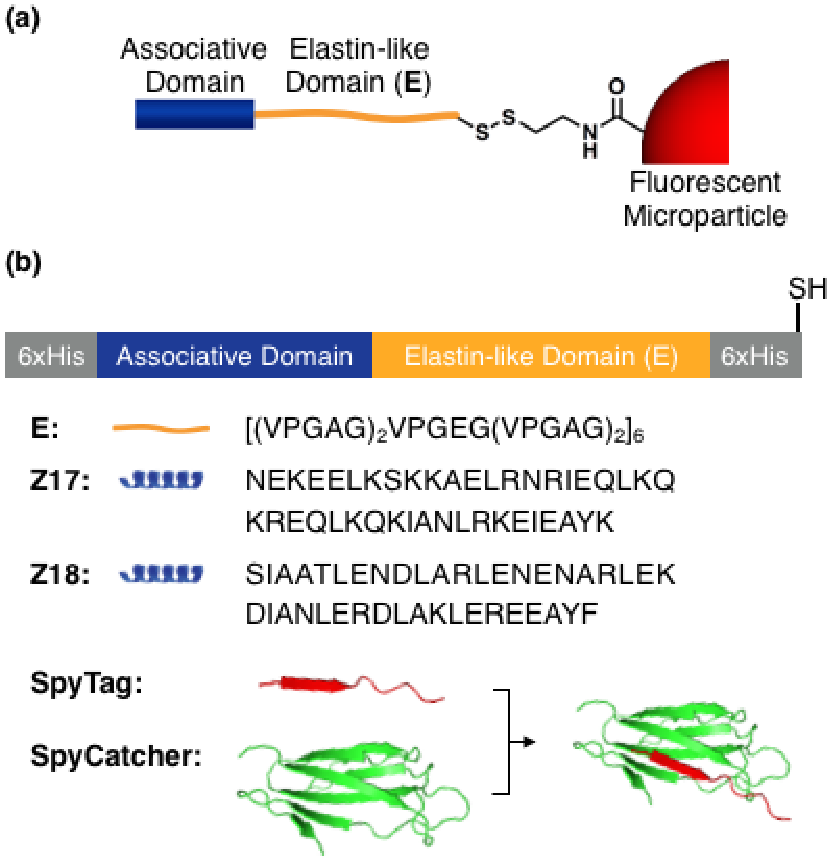


Figure 2.1. Schematic of protein-mediated assembly (a) Schematic illustration of polystyrene particles functionalized with associative proteins. (b) Designs of artificial proteins used in this study. Complete amino acid sequences are given in Table S2.1. Crystal structures of SpyTag and SpyCatcher are adapted from PDB (ID: 4MLI).

Assembly of Protein-Functionalized Particles

Carboxylated fluorescent polystyrene particles (d = 2.0 µm) were activated by treatment with 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) and N-hydroxysuccinimide (NHS). 2-(2-Pyridinyldithio)ethaneamine hydrochloride (PDEA) was added to introduce thiol-reactive pyridyl disulfide functionality to the particle surface.[32] After removal of excess reagents, proteins bearing C-terminal cysteine residues were grafted to the particle surface via thiol exchange.

To initiate colloidal assembly, protein-functionalized particles were mixed in phosphate-buffered saline (PBS, pH 7.5) with 0.005% tween 20 (v/v) and continuously mixed at 25 ºC. After 30 min, particle suspensions bearing either the Z17/Z18 pair or the SpyTag/SpyCatcher pair contained visible aggregates. Suspensions were cast between glass cover slips separated by a 120-µm spacer and imaged by fluorescence confocal microscopy. Mixtures of Z17- and Z18-functionalized particles formed aggregates, broadly distributed in size with an average projected area of 1300 μm^2 (**Figure 2.2a-b**). Aggregates formed from mixtures of SpyTag- and SpyCatcher-functionalized particles were larger (average projected area 3100 μm^2; **Figure 2.2d-e**). Cross-association of Z17 and Z18 particles, and of SpyTag and SpyCatcher particles, was apparent in fluorescence images (**Figure 2.2j-k**) and in the results of colocalization analysis (**Figure 2.2l**). In contrast, no clustering was observed when particles functionalized with Z17 were mixed with those bearing SpyTag (**Figure 2.2g-h**). Colocalization analysis of Z17/SpyTag suspensions revealed negative correlation of green and red fluorescence signals (ρ = -0.51) indicating no substantial cross-association (**Figure 2.2i**). Aggregates formed using either protein pair do not grow without bound, probably because larger aggregates are fragmented by the constant mixing during the aggregation process.

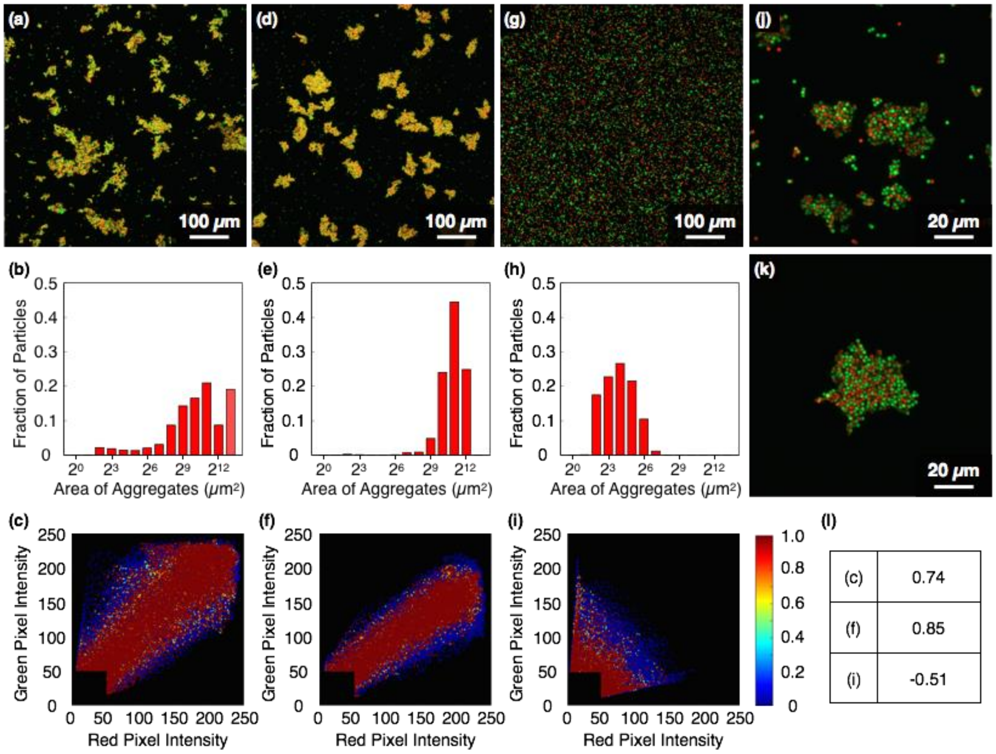


Figure 2.2. Assembly of microparticles functionalized with associative proteins: Assembly of microparticles functionalized with (a) Z17 (green) and Z18 (red), (d) SpyCatcher (green) and SpyTag (red), and (g) SpyTag (green) and Z17 (red). (b, e, h) Size distributions of aggregates shown in (a, d, g). (c, f, i) Colocalization plots of aggregates shown in (a, d, g). (j, k) Magnified images of aggregates of (j) Z17/Z18 and (k) SpyTag/SpyCatcher. (l) Pearson correlation coefficients of colocalization plots.

To investigate the dependence of cluster size on the density of grafting of associative proteins, we prepared sets of fluorescent particles functionalized with different ratios of Z17 and SpyCatcher; each set was then mixed with particles functionalized with SpyTag **(Figure S2.1)**. In this way, the effective number of protein interactions between particles could be varied, although we were unable to determine the absolute surface densities of grafted proteins by flow cytometry, bicinchoninic acid (BCA) assay, or other means. When particles were functionalized by treatment with Z17 and SpyCatcher at a 1:3 ratio (v/v) and mixed with SpyTag-functionalized particles, we found aggregates of average projected area 2900 μm^2. When the concentration of SpyCatcher was reduced to half that of Z17, the aggregate size decreased substantially (average projected area 310 μm^2).

Dissociation of Particle Aggregates

The Z17/Z18 pair drives particle association through physical protein-protein interactions whereas the SpyTag/SpyCatcher pair is expected to form covalent interparticle bonds. We anticipated that the former pair would be dissociated by chemical denaturants and by excess soluble protein competitors, and that the latter would be resistant to such treatments. To test these expectations, we added guanidine hydrochloride (GuHCl) or soluble Z17 protein (identical to the protein that was conjugated to particles) to suspensions of assembled particles, mixed at 25 ºC, and analyzed the resultant aggregates by fluorescence confocal microscopy. Aggregates of Z17- and Z18-coated particles were effectively dispersed both by 5 M GuHCl (3 h) and by 1.0 mg/mL soluble Z17 (24 h) (**Figure 2.3a-d**). Notably, the sizes of the aggregates were controlled by the amounts of denaturant and soluble protein (**Figures S2.2, S2.3**). In contrast, aggregates assembled through interaction of grafted SpyTag and SpyCatcher domains remained intact upon addition of GuHCl or free SpyTag, indicative of stable covalent bond formation between surface-bound proteins (**Figure S2.4**).

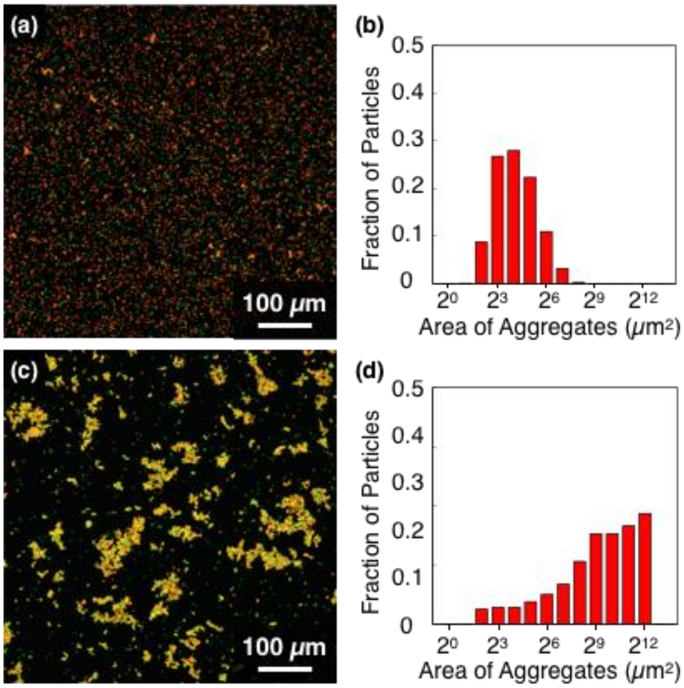


Figure 2.3. Dissociation of particle assemblies (a) Dissociation of Z17–Z18 particle aggregates by (a) 1 mg/mL and (c) 0 mg/mL (control) soluble Z17 mixed at 25 °C for 24h. (b, d) Size distributions of Z17–Z18 particle aggregates shown in (a, c).

Orthogonal Assembly and Selective Dissociation

In light of the specificity of the SYNZIP proteins and the SpyTag/SpyCatcher pair, we expected mixtures of particles coated with Z17, Z18, SpyTag, and SpyCatcher to undergo orthogonal assembly (**Figure 2.4a**). To test this hypothesis, red fluorescent particles functionalized with Z17 or Z18 and green particles functionalized with SpyTag or Spy-Catcher were mixed in PBS with 0.005% tween 20 for 40 min until visible particle aggregates were formed. The resulting particle suspensions were imaged by confocal microscopy. As shown in **Figure 2.4b**, aggregates of red particles and green particles formed separately, and the colocalization plot revealed a strong negative correlation (ρ = -0.6062) between red and green fluorescence channels. Notably, aggregates formed by association of Z17 and Z18 were smaller than those formed by SpyTag and SpyCatcher, consistent with the results of aggregation experiments with separate particle mixtures (**Figure 2.2a-b**).

Because the SpyTag/SpyCatcher pair forms aggregates that are stable with respect to denaturants and excess competitive protein (**Figure 2.3, S2.4**), we imagined that Z17/Z18 aggregates would be selectively dissociated in mixtures of all four particles. We prepared such mixtures, and then added 5 M GuHCl or 1 mg/mL free Z17 to investigate their dissociation behavior (**Figure 2.4c, S2.5**). After mixing for 24 h in 1 mg/mL free Z17, significant dissociation of aggregates of Z17- and Z18-functionalized particles was observed, whereas no dissociation of SpyTag- and SpyCatcher-functionalized particles was noted (**Figure 2.4c**). Similarly, selective dissociation of aggregates of Z17- and Z18-functionalized particles was observed upon treatment with 5 M GuHCl **(Figure S2.5**).

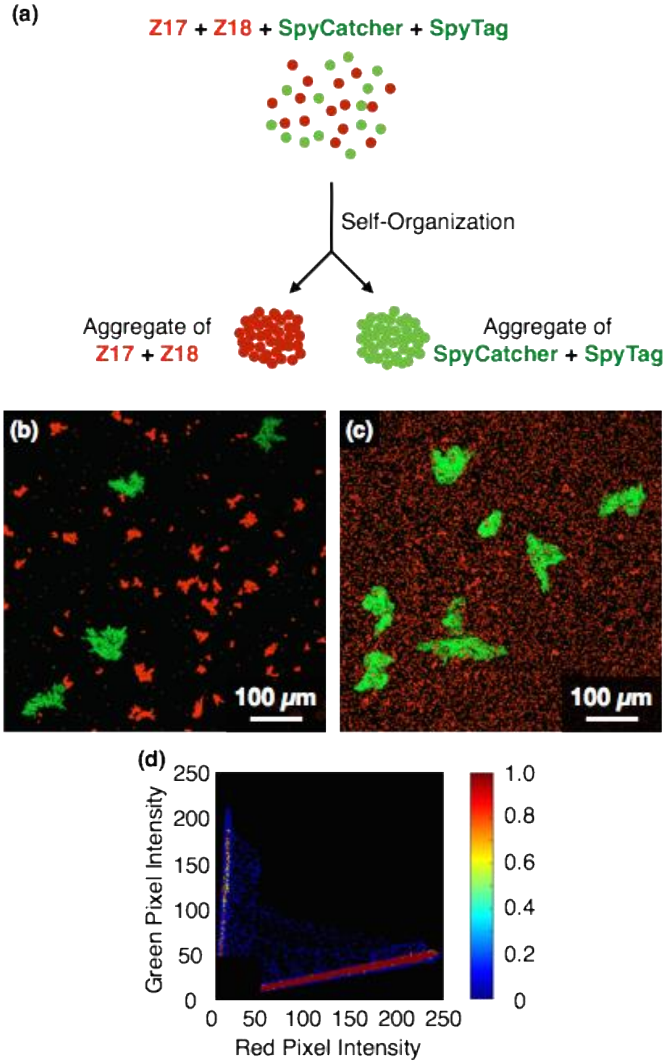


Figure 2.4. Orthogonal assembly of aggregates: (a) Schematic illustration of orthogonal assembly of protein-functionalized particles. Red fluorescent particles are coated with Z17 or Z18, and green fluorescent particles are coated with SpyTag or SpyCatcher. (b) Orthogonal assembly in a 1:1:1:1 particle mixture of Z17 (red), Z18 (red), SpyCatcher (green), and SpyTag(green) in PBS with 0.005% tween 20 mixed at 25 °C for 40 min. (c) Selective dissociation of Z17–Z18 aggregates by 1 mg/mL soluble Z17 for 24 h. (d) Colocalization plot of red and green particles shown in (b).

Formation of Core-Shell Architecture

In drug delivery and tissue engineering applications, it may be useful to form core-shell aggregates to control diffusion of materials into or out of particle clusters. For example, core-shell structures enable the controlled sequential delivery of multiple drugs[33] or delivery of hydrophilic drugs.[34] Toward this end, we constructed core-shell aggregates by exploiting the strong, selective interaction between SpyTag and SpyCatcher (**Figure 2.5a**). Red fluorescent particles were functionalized with SpyTag or SpyCatcher and mixed in PBS with 0.005% tween 20 to form covalent core structures. After 30 min, green fluorescent particles coated with SpyCatcher were added, and the suspension was mixed for 1 h to form the shell. Confocal fluorescence microscopy confirmed the formation of core–shell aggregates with surface-confined green fluorescent particles surrounding the red core structure, although the surface coverage is incomplete (**Figure 2.5b**). Z-stacked images (**Figure 2.5d**) show that the cores are formed exclusively by red particles, which exclude the green particles added subsequently. Moreover, radial fluorescence intensity profiles reveal decreasing red fluorescence near the aggregate surface, where green fluorescence increases (**Figure 2.5c**). In contrast, addition of green E-functionalized particles to red SpyTag/SpyCatcher cores did not yield shell layers (**Figure 2.5f**); assembly of the shell appears to require specific interaction between SpyTag and SpyCatcher. Quantitative analysis of core-shell aggregates showed that 60 ± 8% (n = 9) of the surface was occupied by green particles (**Figure S2.7**), while in the control images, only 7 ± 4% (n = 9) of the surface was green (**Figure S2.8**).

We were unable to construct core–shell structures by SYNZIP-driven assembly. In the SYNZIP system, the core and shell layers were poorly defined; the reversibility of the interaction between Z17 and Z18 appears to enable inter-mixing of the core and shell.

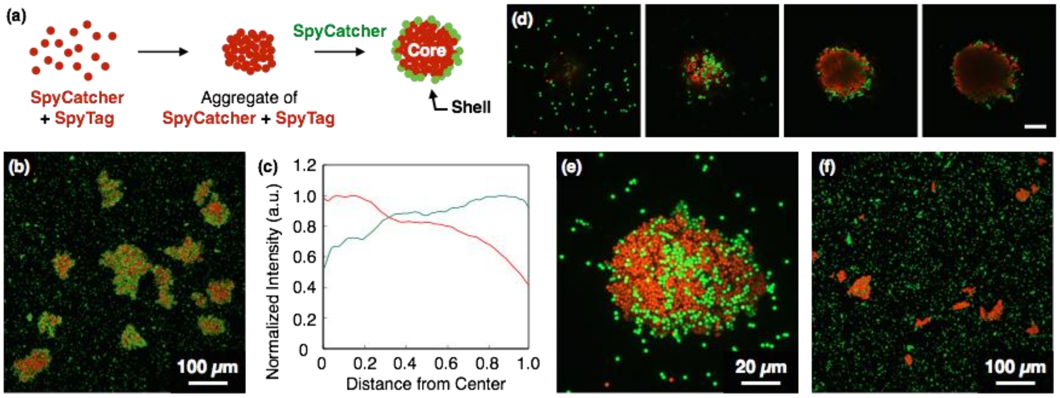


Figure 2.5. Core-shell assembly of particles (a) Schematic illustration of formation of core–shell architecture. (b) Core–shell structure formed by SpyTag- and SpyCatcher-functionalized particles. SpyCatcher-coated particles (green) were added to the aggregates of SpyTag- and SpyCatcher-functionalized particles (red) in PBS with 0.005% tween 20 at 25 °C. (c) Fluorescence intensity of core–shell structure shown in (b), plotted against the distance from center of the aggregates. (d) Z-stack of magnified image of core–shell structure formed by SpyTag and SpyCatcher. Images are shown with 3.87 μm slice spacing. Total thickness: 11.6 μm. Scale bar: 20 μm. (e) Orthogonal projection image of (d). (f) Control experiment for core–shell formation. E-functionalized particles (green) were added to the aggregates of SpyTag- and SpyCatcher-functionalized particles (red).

Conclusions

In this study, we demonstrated programmed assembly of microparticles using two associative protein pairs (Z17/Z18 and SpyTag/SpyCatcher), as well as selective dissociation of mixed aggregates and the formation of core–shell architectures. The methods developed in this report represent a new strategy for the synthesis of mesoscale materials using programmable protein-protein interactions. The strategy is general and easily expanded, owing to the diversity of associative protein domains.[31,35,36] The preparation of complex colloidal aggregates in a scalable, programmable manner should find application in catalysis, health technologies and environmental remediation.[37-39] In addition, the lessons learned in this study are important principles in the development of protein-programmable assembly of living cells, as will be discussed in the rest of this thesis.

Acknowledgements

We thank Dr. Trudy Padmore for technical assistance with particle conjugation and Dr. Andres Collazo for assistance with confocal microscopy. This work was supported by Defense Advanced Research Projects Agency Biological Robustness in Complex Settings Contract HR001-15-C-0093. Imaging was performed in the Biological Imaging Facility, with the support of the Caltech Beckman Institute and the Arnold and Mabel Beckman Foundation.

Materials and Methods

General

Restriction enzymes and ligase were purchased from New England Biolabs (Beverly, NJ). PfuUltraII polymerase was purchased from Agilent Technologies (Santa Clara, CA). Nickel NTA was purchased from Qiagen (Hilden, Germany). N-Hydroxysuccinimide (NHS) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) were purchased from Life Technologies (Carlsbad, CA). 2-(2-Pyridinyldithio)ethaneamine hydrochloride (PDEA) was purchased from GE Healthcare (Piscataway, NJ). Fluorescent, carboxylated polystyrene particles were purchased from Sigma Aldrich (St. Louis, MO). Fluorescence images were obtained with a Zeiss LSM 800 laser scanning confocal microscope, using 20x/0.8 and 100x/1.46 Plan-Apochromat objectives and acquired with Zeiss's Zen software. Images were taken as 30-80 z-stacks with a spacing of 1.1 µm per slice for 20x images and 0.4 µm per slice for 100x images and shown as maximum intensity projections unless otherwise stated.

Image Analysis

All image analysis code was written in Matlab 2015a. Images were generally saved as 16 bit .czi files. Czi files were opened using the Bioformats toolbox and custom-written code.[40]

Cluster size analysis was performed as follows: for simplicity, confocal z-stack images were collapsed into maximum intensity projections. These projections were manually thresholded based on the intensity in each fluorescence channel. Pixels above the threshold in either channel were defined to be "bright." The projected areas of aggregates containing contiguous "bright" pixels were extracted. The projected areas were then converted into area-weighted distributions and area-weighted averages according to the equations:

$$P_{w,i} = \frac{iN_i}{\sum_{i=1}^{N} iN_i}$$

$$\overline{A_W} = \frac{\sum_{i=1}^{N} i^2 N_i}{\sum_{i=1}^{N} iN_i}$$

where $P_{w,i}$ is the area-weighted probability of a cluster of projected area i, and N_i is the number of clusters of this size. These probabilities were binned logarithmically (base 2), and plotted as histograms. The height of a bar represents the probability that a particle chosen at random is found in an aggregate of projected area between the number for that bar and the number for the next bar (e.g., the bar for 128 μm^2 contains aggregates between 128 and 256 μm^2). Aggregate "volumes" may be extracted in similar fashion, but in our experience, projected areas can be determined more accurately because the laser intensity is attenuated in the cores of larger aggregates.

Colocalization analysis was performed as follows: confocal z-stack images were collapsed into maximum intensity projections. Because individual particles extend beyond a single pixel, images were blurred using a mean filter acting on a disc of radius five pixels (representing approximately two particle diameters). In this way, adjacent pixels were blurred into each other, while leaving the larger-scale structure of the aggregate intact. Dark pixels (those below threshold) were then excluded from the analysis, and the colocalization between fluorescence channels was plotted using scatplot.[41] Pearson correlation coefficients were calculated to provide a measure of colocalization of green and red fluorescence signals.

Core-shell fluorescence intensity profiles were created as follows: In maximum intensity projections, large aggregates were identified by thresholding in a manner similar to that used for cluster-size analysis. For each large aggregate, z-stacks with high levels of fluorescence were combined using a mean-intensity projection. Then, starting at the centroid of each aggregate, 100 radii representing equally spaced directions were drawn to the edge of the aggregate (determined by thresholding), extracting the fluorescence intensities from each channel. The fluorescence intensities were then scaled (with a value of 1 representing the maximum fluorescence intensity in each aggregate) and plotted along a “location” axis from 0 to 1 (with 0 representing the centroid and 1 representing the edge of the aggregate for each radius).

Surface coverage of core-shell aggregates was analyzed as follows. Three-dimensional z-stacks of core-shell aggregates were obtained at 100x magnification and with 0.39 μm z-spacing. Then, a maximum intensity projection was thresholded, and the projected area of the large cluster was identified. Starting from the top of the image, in each z-slice bright pixels in the projected area of the cluster were identified and classified as red or green. Locations of these bright pixels are then preserved for successive z-slices of the aggregate such that only pixels that are on the outside of the aggregate are counted. The ratio of “shell” pixels to the total number of bright pixels represents the surface coverage. This process is illustrated in **Figures S2.7** and **S2.8.**

All image analysis code can be obtained from http://tirrell-lab.caltech.edu/Code.

Cloning of Recombinant Proteins

Recombinant fusion proteins were produced by standard recombinant DNA technology. DH10b or Mach1 strains of *Escherichia coli* were used for all cloning steps. A modified pQE-80L expression vector was used for expression. The *Xho*I site upstream of the T5 promoter was deleted and a *C*-terminal in-frame cysteine residue was added down-stream of the multiple cloning site by two rounds of site-directed mutagenesis. The modified pQE-80L expression vector was denoted pQE-80X-Cys.

The DNA sequence encoding the elastin-like polypeptide segment of each fusion protein was synthesized in two 250 bp segments denoted E3 and E3' (Genscript) with 5' *Bam*HI and *Sal*I sites and 3' *Xho*I and *Hind*III sites. These segments have identical amino acid sequences but different degenerate DNA sequences to decrease sequence similarity at the DNA level. A polyhistidine tag was added between the *Xho*I and *Hind*III sites. These DNA fragments were inserted into pQE-80X-Cys between the *Bam*HI and *Hind*III sites to make plasmids pQE-80X-E_3-Cys and pQE-80X-E_3'-Cys and transformed into chemically competent *E. coli.* The full-length elastin-like polypeptide was synthesized via recursive directional ligation.[42] Briefly, pQE-80X-E_3-Cys was digested with *Xho*I and *Hind*III to linearize the plasmid, while pQE-80X-E_3'-Cys was digested with *Sal*I and *Hind*III. The resulting DNA fragments were ligated together to fuse the two elastin-like sequences. Because *Sal*I and *Xho*I have complementary overhangs, the restriction sites are ablated, leaving a two-amino acid scar, but retaining the 5' *Bam*HI and *Sal*I sites and 3' *Xho*I and *Hind*III sites in the plasmid. This construct is denoted pQE-80X-E-Cys.

The SYNZIP17 and SYNZIP18 domains were ordered from IDT (Coralville, IA) as G-blocks with 5' *Bam*HI and *Sal*I sites, 3' *Xho*I and *Hind*III sites, and a polyhistidine tag between the 3' restriction sites. These constructs were inserted into pQE-80X-Cys by digestion with *Bam*HI and *Hind*III and ligation as above. The elastin-like polypeptide was then added to the *C*-termini of these coding sequences by recursive directional ligation by the procedure described above. The resulting constructs are denoted pQE-80X-Z17-E-Cys and pQE-80X-Z18-E-Cys.

The coding sequence for SpyCatcher was amplified from pQE-EB,[28] with the addition of 5' *Bam*HI and *Sal*I sites, 3' *Xho*I and *Hind*III sites, and a polyhistidine tag between the 3' restriction sites. This construct was inserted into pQE-80X-Cys by digestion with *Bam*HI and *Hind*III and ligation; recursive directional ligation was used to add the elastin-like polypeptide to the *C*-terminus to yield pQE-80X-SpyCatcher-E-Cys.

Single-stranded oligonucleotides encoding SpyTag, 5' *Bam*HI and *Sal*I sites, 3' *Xho*I and *Hind*III sites, and a polyhistidine tag between the 3' restriction sites were ordered from IDT, and annealed by cooling from 95 °C to room temperature. The annealed oligonucleotides were inserted into a digested pQE-80X-Cys vector, and recursive directional ligation was used to add the elastin-like polypeptide to the *C*-terminus to yield pQE-80X-SpyTag-E-Cys.

All constructs were confirmed by sequencing; full protein sequences are given in Table S2.1.

Protein Expression and Purification

Constructs were transformed into *E. coli* strain BL21 for expression. Expression was performed in Terrific Broth (12 g/L casein, 24 g/L yeast extract, 0.4% w/v glycerol, 0.017 M monobasic potassium phosphate, 0.072 M dibasic potassium phosphate). Cultures were induced at an optical density of 0.6-0.9 to a final concentration of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Expression was allowed to proceed for 5 h, after which cells were harvested by centrifugation. For all proteins other than SpyCatcher-E-Cys, cultures were resuspended in denaturing lysis buffer (8 M urea, 0.1 M Na_2HPO_4, 10 mM imidazole; pH 8.0), and lysed by sonication. Lysates were cleared by centrifugation and incubated with NiNTA resin. The resin was washed with lysis buffer and wash buffer (8 M urea, 0.1 M Na_2HPO_4, 25 mM imidazole; pH 6.3). Protein was eluted with elution buffer (8 M urea, 0.1 M Na_2HPO_4, 250 mM imidazole; pH 3.5). Purity was confirmed with SDS-PAGE. Proteins were then extensively dialyzed against water and lyophilized for storage.

SpyCatcher-E-Cys was purified under native conditions. Cultures were resuspended in native lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, 1 mg/mL lysozyme; pH 8.0). Cells were lysed by sonication, and cleared lysates were incubated with NiNTA resin. The resin was washed with native wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 25 mM imidazole; pH 8.0), and protein was eluted with native elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole; pH 8.0). Purity was confirmed with SDS-PAGE, and purified SpyCatcher-E-Cys was dialyzed against water and lyophilized.

Immobilization of Associative Proteins

Fluorescent carboxylated polystyrene microparticles (d = 2.0 µm; Sigma-Aldrich) were dispersed in 50 mM MES buffer (pH 6.8). A solution of 30 mM NHS and 20 mM EDC in MES buffer was added and mixed at 25 ºC. After 30 min, particles were collected by centrifugation and washed with PBS (pH 7.5). PDEA was dissolved in sodium acetate buffer (100 mM, pH 4.2) and added to particles to a final concentration of 25 mM. After mixing at 25 ºC for 30 min, particles were collected and washed with PBS. Protein solution (1 mg/mL) in PBS with 10 mM sodium azide was added to particles and samples were mixed at 25 ºC for 22 h. Particles were washed with PBS to remove unreacted proteins, and dispersed in PBS with 10 mM sodium azide for use in assembly experiments.

Supplemental Tables and Figures

Table S2.1: Protein Sequences

Protein:	Sequence
E-Cys	MRGSHHHHHHGSVDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLEHHHHHHKLC
Z17-E-Cys	MRGSHHHHHHGSVDGSGSGSGSGSGANEKEELKSKKAELRNRIEQLKQKREQLKQKIANLRKEIEAYKGSGSGSGSGSGALDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLEHHHHHHKLC
Z18-E-Cys	MRGSHHHHHHGSVDGSGSGSGSGSGASIAATLENDLARLENENARLEKDIANLERDLAKLEREEAYFGSGSGSGSGSGALDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLEHHHHHHKLC
SpyTag-E-Cys	RGSHHHHHHGSVDGSGSGSGSGSGAAHIVMVDAYKPTKGSGSGSGSGSGALDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLEHHHHHHKLC
SpyCatcher-E-Cys	MRGSHHHHHHGSVDGSGSGSGSGSGAAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIDGSGSGSGSGSGALDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLEHHHHHHKLC

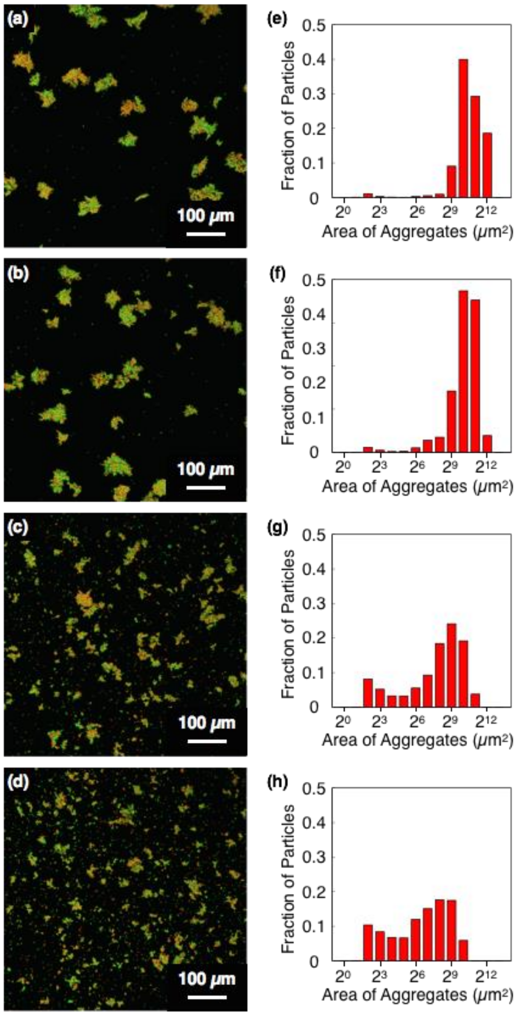


Figure S2.1. Control of aggregate size with surface content: assembly of particles functionalized with SpyCatcher and Z17 (green) and SpyTag-functionalized particles (red). For the preparation of green particles, 1 mg/mL solutions of SpyCatcher and Z17 were mixed at ratios of (a) 3:1, (b) 2:1, (c) 1:1 and (d) 0.5:1 (v/v), and proteins were immobilized via thiol exchange. Each population of green particles was mixed with red particles treated with SpyTag and suspensions were continuously mixed for 30 min at 25 ºC in PBS (pH 7.5) with 0.005% tween 20 ([green particle] = [red particle] = 1 × 10^5 particle/μL). (e-h) Size distribution histograms of particle aggregates shown in (a-d).

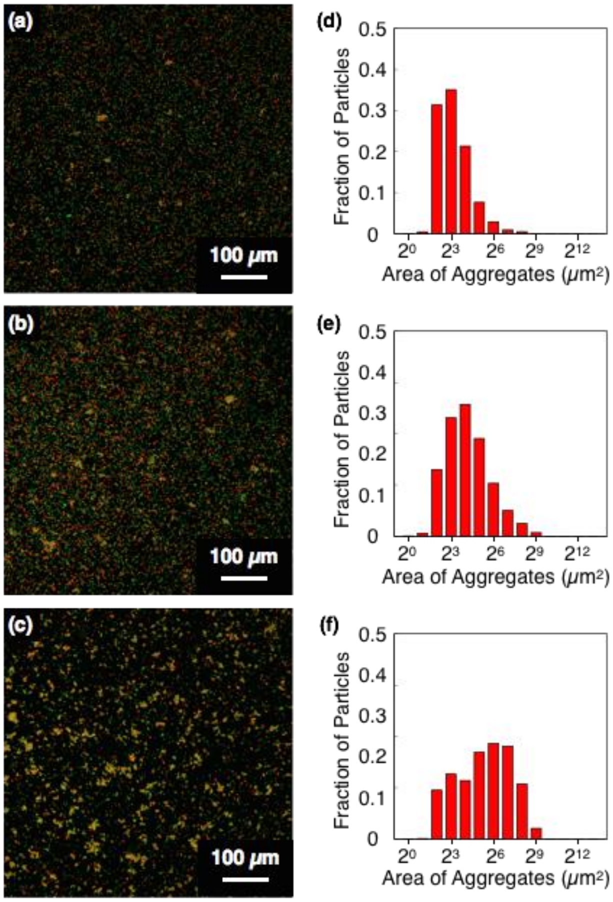


Figure S2.2. Aggregate dissociation with soluble protein: dissociation of the aggregates of Z17- (green) and Z18- (red) functionalized particles by addition of soluble Z17. Particle aggregates were prepared by continuously mixing Z17- (green) and Z18- (red) functionalized particles for 40 min at 25 °C in PBS (pH 7.5) with 0.005% tween 20 ([Z17 (green) particle] = [Z18 (red) particle] = 1×10^5 particle/μL). Then, (a) 0.2 mg/mL (b) 0.04 mg/mL and (c) 0.008 mg/mL of free Z17 solution was added to the particle aggregates and mixed at 25 °C for 24 h. (d-f) Size distribution histograms of particle aggregates shown in (a-c).

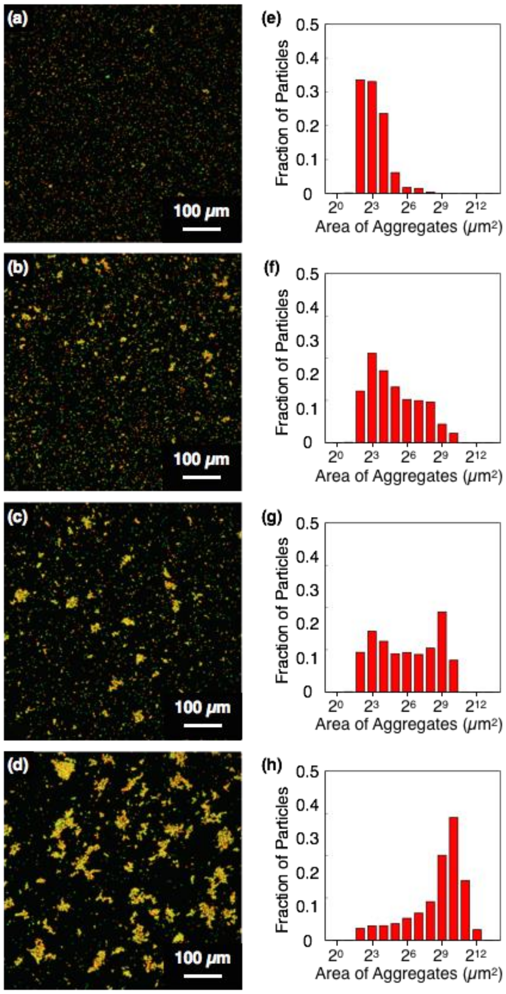


Figure S2.3. Aggregate dissociation with GuHCl: dissociation of the aggregates of Z17- (green) and Z18- (red) functionalized particles by addition of GuHCl. Particle aggregates were prepared by mixing Z17- (green) and Z18- (red) functionalized particles for 40 min at 25 ºC in PBS (pH 7.5) with 0.005% tween 20 ([Z17 (green) particle] = [Z18 (red) particle] = 1 × 10^5 particle/μL). Then, (a) 5 M (b) 1 M, (c) 0.2 M and (d) 0.04 M of GuHCl was added to the particle aggregates and mixed at 25 ºC for 3 h. (e-h) Size distribution histograms of particle aggregates shown in (a-d).

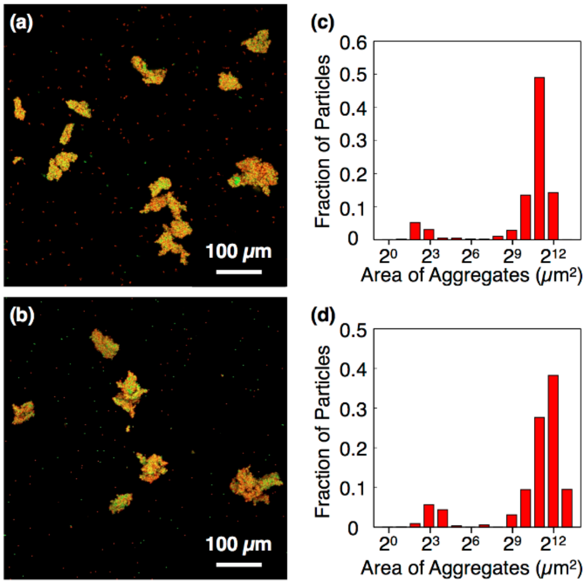
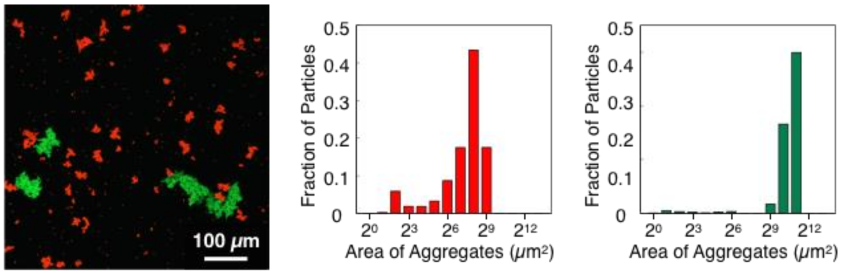
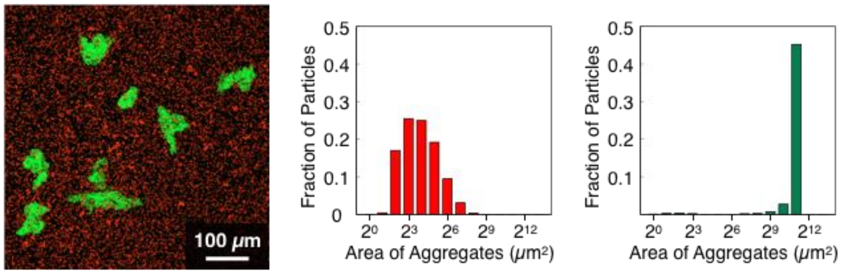


Figure S2.4. SpyTag/SpyCatcher dissociation stability: aggregates of SpyCatcher- (green) and SpyTag- (red) functionalized particles are stable in the presence of GuHCl and soluble SpyTag. Particle aggregates were prepared by mixing SpyCatcher- (green) and SpyTag- (red) functionalized particles for 40 min at 25 °C in PBS (pH 7.5) with 0.005% tween 20 ([SpyCatcher (green) particle] = [SpyTag (red) particle] = 1×10^5 particle/μL). Then, (a) 5 M GuHCl and (b) 1 mg/mL soluble SpyTag were added to the aggregate suspension and mixed at 25 °C for 3 h and 24 h, respectively. (c, d) Size distribution histograms of (a, b).

(a) 0.005% Tween 20 (control)



(b) 1 mg/mL Soluble Z17



(c) 5M GuHCl

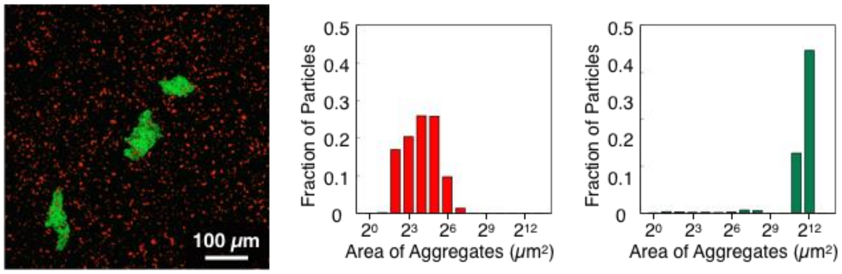


Figure S2.5. Selective aggregate dissociation (a) A 1:1:1:1 particle mixture of Z17 (red), Z18 (red), SpyCatcher (green) and SpyTag(green) in PBS with 0.005% tween 20 at 25 ºC ([Z17 (red) particle] = [Z18 (red) particle] = [SpyCatcher (green) particle] = [SpyTag (green) particle] = 0.5 × 10^5 particle/μL). (b) 0.1% Soluble Z17 or (c) 5 M GuHCl was added to the assembled particle mixture. Aggregates of Z17 and Z18 (red) were selectively dissociated while aggregates of SpyCatcher and SpyTag (green) were stable.

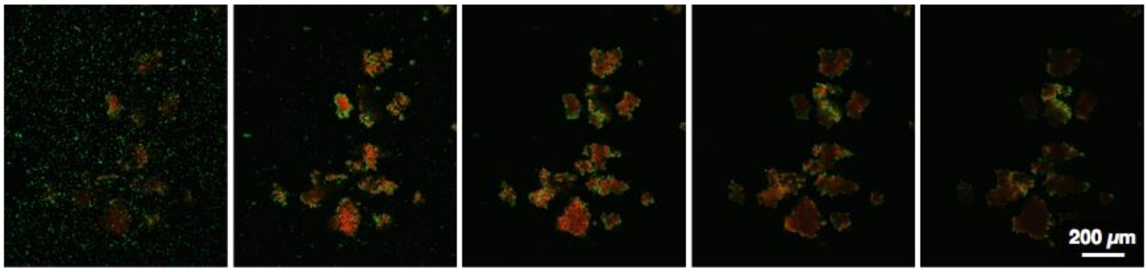


Figure S2.6 Core-shell Z-stack images: z-stack images of core–shell structure formed by SpyTag- and SpyCatcher-functionalized particles related to Figure 5b. SpyTag- (red) and SpyCatcher- (red) functionalized particles were mixed at 25 ºC for 30 min in PBS (pH 7.5) with 0.005% tween 20 to form core structures. Then, SpyCatcher-coated particles (green) were added to the aggregates and mixed at 25 ºC for 1 h to form the shell. ([SpyCatcher (red) particle] = [SpyTag (red) particle] = [SpyCatcher (green) particle] = 0.7 × 105 particle/μL). Images were shown with 3.42 μm slice spacing. Total thickness is 13.7 μm.

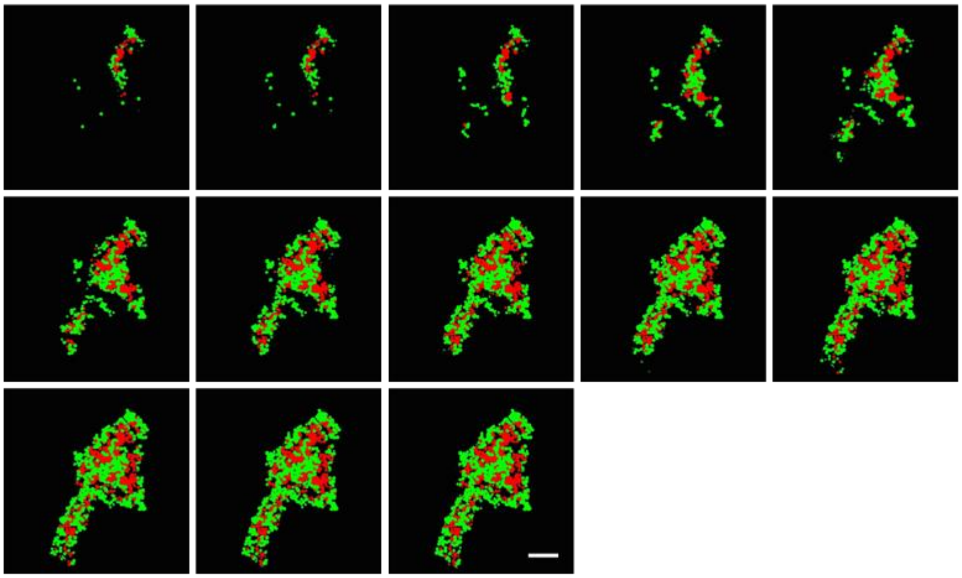


Figure S2.7. Illustration of Surface Projection Algorithm Core-Shell: successive steps of the surface projection algorithm of a core–shell structure formed by SpyTag- and SpyCatcher-functionalized particles. SpyTag- (red) and SpyCatcher- (red) functionalized particles were mixed at 25 ºC for 30 min in PBS (pH 7.5) with 0.005% tween 20 to form core structures. Then, SpyCatcher-coated particles (green) were added to the aggregates and mixed at 25 ºC for 1 h to form the shell. ([SpyCatcher (red) particle] = [SpyTag (red) particle] = [SpyCatcher (green) particle] = 0.7×10^5 particle/μL). Successive z-slices from each step of the algorithm are shown with 2.34 μm of slice spacing. Total thickness: 28.1 μm. Scale bar: 20 μm.

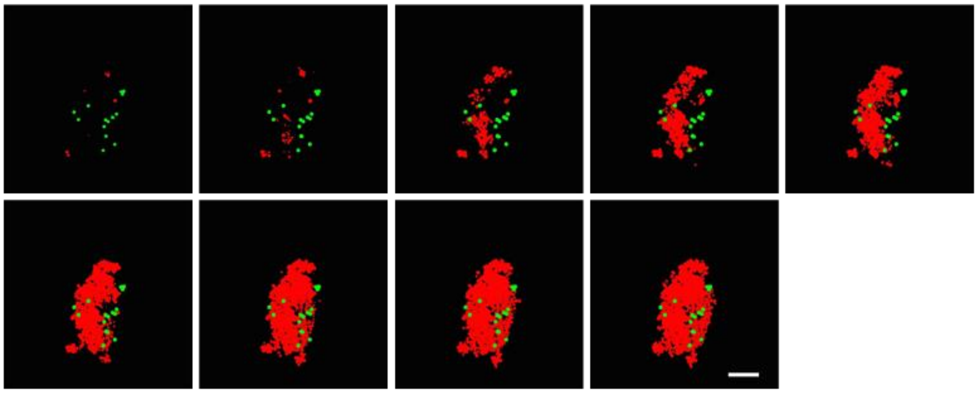


Figure S2.8. Illustration of surface projection algorithm control: successive steps of surface projection algorithm for control experiment for core–shell formation. E-functionalized particles (green) were added to the aggregates of SpyTag- (red) and SpyCatcher- (red) functionalized particles and mixed at 25 ºC for 1 h in PBS (pH 7.5) with 0.005% tween 20 ([SpyCatcher (red) particle] = [SpyTag (red) particle] = [E (green) particle] = 0.7×10^5 particle/μL). Successive z-slice images from each step of the algorithm are shown with 2.34 μm of slice spacing. Total thickness: 18.7 μm.

References

1 F. Li, D. P. Josephson and A. Stein, Colloidal assembly: The road from particles to colloidal molecules and crystals., *Angew. Chemie*, 2011, **50**, 360–88.

2 O. D. Velev and S. Gupta, Materials fabricated by micro- and nanoparticle assembly - the challenging path from science to engineering, *Adv. Mater.*, 2009, **21**, 1897–1905.

3 M. Y. Lin, H. M. Lindsay, D. A. Weitz, R. C. Ball, R. Klein and P. Meakin, Universality in colloid aggregation, *Nature*, 1989, **339**, 360–362.

4 N. Rosi and C. Mirkin, Nanostructures in biodiagnostics, *Chem. Rev.*, 2005, **105**, 1547–1562.

5 F. Fan and K. Stebe, Assembly of colloidal particles by evaporation on surfaces with patterned hydrophobicity, *Langmuir*, 2004, **20**, 3062–3067.

6 Z. Zhou, Q. Yan, Q. Li and X. S. Zhao, Fabrication of binary colloidal crystals and non-close-packed structures by a sequential self-assembly method, *Langmuir*, 2007, **23**, 1473–1477.

7 K. Lin, J. Crocker, V. Prasad and A. Schofield, Entropically driven colloidal crystallization on patterned surfaces, *Phys. Rev. Lett.*, 2000, **85**, 1–4.

8 A. Rugge and S. H. Tolbert, Effect of electrostatic interactions on crystallization in binary colloidal films, *Langmuir*, 2002, **18**, 7057–7065.

9 M. E. Leunissen, C. G. Christova, A.-P. Hynninen, C. P. Royall, A. I. Campbell, A. Imhof, M. Dijkstra, R. van Roij and A. van Blaaderen, Ionic colloidal crystals of oppositely charged particles, *Nature*, 2005, **437**, 235–40.

10 C. Snyder, M. Ong and D. Velegol, In-solution assembly of colloidal water, *Soft Matter*, 2009, **5**, 1263–1268.

11 C. Mirkin, R. Letsinger, R. Mucic and J. Storhoff, A DNA-based method for rationally assembling nanoparticles into macroscopic materials, *Nature*, 1996, **382,** 607-609

12 R. J. Macfarlane, B. Lee, M. R. Jones, N. Harris, G. C. Schatz and C. A. Mirkin, Nanoparticle superlattice engineering with DNA, *Science* 2011, **334**, 204–209.

13 Y. Wang, Y. Wang, X. Zheng, É. Ducrot, J. S. Yodh, M. Weck and D. J. Pine, Crystallization of DNA-coated colloids, *Nat. Commun.*, 2015, **6**, 7253.

14 L. Di Michele, F. Varrato, J. Kotar, S. H. Nathan, G. Foffi and E. Eiser, Multistep kinetic self-assembly of DNA-coated colloids, *Nat. Commun.*, 2013, **4**, 2007.

15 A. P. Schoen, B. Hommersom, S. C. Heilshorn and M. E. Leunissen, Tuning colloidal association with specific peptide interactions, *Soft Matter*, 2013, **9**, 6781.

16 D. Ernenwein, P. Ghosh, V. Rotello and J. Chmielewski, Gold nanoparticle self-assembly promoted by a non-covalent, charge-complemented coiled-coil peptide, *J. Mater. Chem.*, 2010, **20**, 5608.

17 M. M. Stevens, N. T. Flynn, C. Wang, D. A. Tirrell and R. Langer, Coiled-coil peptide-based assembly of gold nanoparticles, *Adv. Mater.*, 2004, **16**, 915–918.

18 D. Woolfson, The design of coiled-coil structures and assemblies, *Adv. Protein Chem.*, 2005, **70**, 79–112.

19 A. W. Reinke, R. A. Grant and A. E. Keating, A synthetic coiled-coil interactome provides heterospecific modules for molecular engineering, *J. Am. Chem. Soc.*, 2010, **132**, 6025–6031.

20 J. M. Mason and K. M. Arndt, Coiled coil domains: Stability, specificity, and

biological implications, *Chembiochem*, 2004, **5**, 170–6.

21 M. G. Ryadnov, B. Ceyhan, C. M. Niemeyer and D. N. Woolfson, “Belt and Braces ”: A peptide-based linker system of de novo design, *J. Am. Chem. Soc.*, 2003, **125**, 9388–9394.

22 W. M. Park and J. A. Champion, Colloidal assembly of hierarchically structured porous supraparticles from flower-shaped protein-inorganic hybrid nanoparticles, *ACS Nano*, 2016, **10**, 8271–80.

23 M. P. Nikitin, T. A. Zdobnova, S. V Lukash, O. A. Stremovskiy and S. M. Deyev, Protein-assisted self-assembly of multifunctional nanoparticles, *Proc. Natl. Acad. Sci.*, 2010, **107**, 5827–32.

24 U. F. Aghayeva, M. P. Nikitin, E. V. Korostylev, S. V. Lukash, S. M. Deyev and R. V. Petrov, Self-assembly of magnetic and fluorescent colloidal constructs based on protein-protein interactions, *Dokl. Biochem. Biophys.*, 2012, **445**, 210–212.

25 B. Zakeri, J. Fierer, E. Celik, E. Chittock, U. Schwarz-Linek, V. Moy and M. Howarth, Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin, *Proc. Natl. Acad. Sci.*, 2012, **109**, 690-697

26 R. Perez and H.W. Kim, Core-shell designed scaffolds for drug delivery and tissue engineering, *Acta Biomater.*, 2015, **21**, 2–19.

27 F. Sun, W.-B. Zhang, A. Mahdavi, F. H. Arnold and D. A. Tirrell, Synthesis of bioactive protein hydrogels by genetically encoded SpyTag-SpyCatcher chemistry, *Proc. Natl. Acad. Sci.*, 2014, **111**, 11269–74.

28 W. Zhang, F. Sun, D. A. Tirrell and F. H. Arnold, Controlling macromolecular topology with genetically encoded SpyTag–SpyCatcher chemistry, *J. Am. Chem Soc.*, 2013, **135**, 13988–13997.

29 C. N. Bedbrook, M. Kato, S. Ravindra Kumar, A. Lakshmanan, R. D. Nath,

F. Sun, P. W. Sternberg, F. H. Arnold and V. Gradinaru, Genetically encoded spy peptide fusion system to detect plasma membrane-localized proteins in vivo, *Chem. Biol.*, 2015, **22**, 1108–21.

30 D. Dovala, W. S. Sawyer, C. M. Rath and L. E. Metzger, Rapid analysis of protein expression and solubility with the SpyTag-SpyCatcher system., *Protein Expr. Purif.*, 2016, **117**, 44–51.

31 G. Veggiani, T. Nakamura, M. D. Brenner, R. V Gayet, J. Yan, C. V Robinson and M. Howarth, Programmable polyproteams built using twin peptide superglues, *Proc. Natl. Acad. Sci.*, 2016, **113**, 1202–1207.

32 S. Bubenikova, I. C. Stancu, L. Kalinovska, E. Schacht, E. Lippens, H. Declercq, M. Cornelissen, M. Santin, M. Amblard and J. Martinez, Chemoselective cross-linking of alginate with thiol-terminated peptides for tissue engineering applications, *Carbohydr. Polym.*, 2012, **88**, 1239–1250.

33 S. Narayanan, M. Pavithran, A. Viswanath, D. Narayanan, C. C. Mohan, K. Manzoor and D. Menon, Sequentially releasing dual-drug-loaded PLGA-casein core/shell nanomedicine: Design, synthesis, biocompatibility and pharmacokinetics, *Acta Biomater.*, 2014, **10**, 2112–2124.

34 M. P. A. Lim, W. L. Lee, E. Widjaja and S. C. J. Loo, One-step fabrication of core-shell structured alginate-PLGA/PLLA microparticles as a novel drug delivery system for water soluble drugs, *Biomater. Sci.*, 2013, **1**, 486–493.

35 K. E. Thompson, C. J. Bashor, W. A. Lim and A. E. Keating, SYNZIP protein interaction toolbox: In vitro and in vivo specifications of heterospecific coiled-coil interaction domains, *ACS Synth. Biol.*, 2012, **1**, 118–129.

36 S. Kim, J. H. Kim, J. S. Lee and C. B. Park, Beta-sheet-forming, self-assembled peptide nanomaterials towards optical, energy, and healthcare applications., *Small*, 2015, **11**, 3623–40.

37 Z. J. Gartner and C. R. Bertozzi, Programmed assembly of 3-dimensional microtissues, *Proc. Natl. Acad. Sci.*, 2009, 1–5.

38 S. Elias and E. Banin, Multi-species biofilms: Living with friendly neighbors., *FEMS Microbiol. Rev.*, 2012, **36**, 990–1004.

39 H. J. Kim, W. Du and R. F. Ismagilov, Complex function by design using spatially pre-structured synthetic microbial communities: Degradation of pentachlorophenol in the presence of Hg(II)., *Integr. Biol.*, 2011, **3**, 126–133.

40 M. Linkert, C. T. Rueden, C. Allan, J.-M. Burel, W. Moore, A. Patterson, B. Loranger, J. Moore, C. Neves, D. Macdonald, A. Tarkowska, C. Sticco, E. Hill, M. Rossner, K. W. Eliceiri and J. R. Swedlow, Metadata matters: Access to image data in the real world, *J. Cell Biol.*, 2010, **189**, 777–82.

41 A. Sanchez-Barba, Scatplot. 2005.

42 J. R. Mcdaniel, J. A. Mackay and F. Garcı, Recursive directional ligation by plasmid reconstruction allows rapid and seamless cloning of oligomeric genes, *Macromolecules*, 2010, **11**, 944–952.

Chapter 3

GENETICALLY PROGRAMMABLE MICROBIAL ASSEMBLY

Abstract

Engineered microbial communities show promise in a wide range of applications, including environmental remediation, microbiome engineering, and synthesis of fine chemicals. Here we present methods by which bacterial aggregates can be directed into several distinct architectures by inducible surface expression of hetero-associative protein domains (SpyTag/SpyCatcher and SynZip17/18). Programmed aggregation can be used to activate a quorum-sensing circuit, and aggregate size can be tuned via control of the amount of the associative protein displayed on the cell surface. We further demonstrate reversibility of SynZip-mediated assembly by addition of soluble competitor peptide. Genetically programmable bacterial assembly provides a starting point for the development of new applications of engineered microbial communities in environmental technology, agriculture, human health, and bioreactor design.

Introduction

Bacteria and other microorganisms form complex, multispecies consortia in a wide variety of environments such as marine sediments,[1] soils,[2] biofilms,[3] and the human gut.[4] Living in a consortium affords important advantages for the member species, such as protection from toxins and antibiotics,[5] cross-feeding relationships that allow more flexible utilization of nutrients,[6,7] and efficient division of labor.[8] Recently, there has been substantial interest in the development of artificial consortia for use in environmental remediation,[9,10] biofuel production,[11] and construction of microbial fuel cells.[12] By dividing metabolic tasks across multiple organisms, the genetic and metabolic stresses placed on individual organisms can be minimized, leading to improved yields.[13]

Imposing spatial organization on microbial consortia has the potential to provide further advantages. For example, pentachlorophenol (PCP) is commonly found in sites that contain significant amounts of mercury,[14,15] and the concentration of mercury is often high enough to kill microorganisms tasked with remediation of PCP. To remedy this problem, Ismagilov and co-workers used extrusion to construct a coaxial consortium in which a central cylinder of *Sphingobium chlorophenolicum* oxidizes PCP and a shell of *Ralstonia metallidurans* provides protection from toxic mercury ions.[16] Other techniques for organizing bacterial consortia include inkjet[17] and 3D printing.[18] Each of these methods requires “top-down” processing to impose structure. An alternative approach would encode the capacity for controlled assembly into the genetic material of the consortium, such that aggregation could be triggered in response to biochemical or optogenetic stimuli.[19-20]

Fernandez and co-workers reported that surface display of the complementary Junβ and Fosβ leucine zipper pair via fusion to the C-terminal region of the adhesin protein EhaA of *Escherichia coli* could be used to drive bacterial aggregation.[21] More recently, we demonstrated selective, orthogonal assembly of micro-particles functionalized by covalently-attached associating proteins.[22] Here we build on these results to direct assembly of *Escherichia coli* into aggregates of controlled size and structure. Furthermore, we demonstrate that genetically-programmed assembly of bacterial cells can lead to triggering of a quorum-sensing circuit at a cell density that does not ordinarily support quorum sensing, and we present a reaction-diffusion model by which this result can be understood.

Results and Discussion

Design of the bacterial aggregation system

To prepare bacterial aggregates, we expressed two sets of associating proteins on the *E. coli* cell surface, building on the autodisplay system first reported by Maurer, Jose, and Mayer.[23] This system has been used to display a wide variety of proteins, including hydrolases, esterases, enzyme inhibitors, and epitopes for vaccine development on bacterial cell surfaces.[24,25] Here we used the autotransporter system to display two pairs of cross-associating proteins, SynZip17/18 and SpyTag/SpyCatcher (sequences shown in Table S3.2). The SynZip proteins were adapted from a library of leucine-zipper peptides reported by the Keating laboratory.[26] SynZip17 and 18 are reported to form anti-parallel coiled-coil dimers with high (< 10 nM) affinity and cross-association specificity. SpyTag and SpyCatcher were derived from the fibronectin-binding protein FbaB of *Staphylococcus pyogenes*, as first reported by Howarth and coworkers.[27,28] After splitting the full length protein into two polypeptide chains, Howarth and coworkers showed that the resulting SpyTag and SpyCatcher fragments undergo spontaneous coupling via formation of an isopeptide bond between lysine residue K31 in SpyCatcher and aspartic acid residue D117 in SpyTag. SpyTag/SpyCatcher chemistry has been used to control protein topology,[29] crosslink protein hydrogels,[30] engineer novel protein vaccines,[31] and cyclize enzymes for enhanced thermal stability.[32]

The expression constructs are shown in **Figure S3.1**. In each construct, the target associative domain is fused to a 6xHis tag (for immunostaining) and inserted between a PelB secretion sequence and the autotransporter. Expression was controlled either by a T5-Lac or by an araBAD promoter, to enable induction by isopropyl-β-D-1-thiogalactopyranoside (IPTG) or L-arabinose, respectively. Plasmids bearing SynZip17, SynZip18, SpyTag, or SpyCatcher, under control of the T5-Lac promoter on a pQE80 backbone, are referred to as pAT-17, pAT-18, pAT-ST, and pAT-SC, respectively. The same protein constructs under control of an arabinose promoter on a pBAD33 backbone are referred to as pBAD-17, pBAD-18, pBAD-ST, and pBAD-SC, respectively. Expression plasmids were introduced into *E. coli* strain DH10B for aggregation experiments. Cells were co-transformed with plasmids encoding mWasabi or mCherry to allow the aggregation process to be monitored by fluorescence confocal microscopy.

Procedures for forming bacterial aggregates

Individual colonies chosen from Luria-Bertani (LB) plates were grown overnight to stationary phase in LB medium supplemented with 100 mg/L ampicillin or 35 mg/L chloramphenicol, then used to inoculate fresh cultures at 100:1 dilution. When the optical density (OD_{600}) reached 0.6-0.8, cultures were induced with 0.1 mM IPTG (for pQE-80-based plasmids) or 0.1% L-arabinose (for pBAD-33-based plasmids). and allowed to express for 90 min at 37°C and 300 RPM agitation speed (slower speeds would often cause settling of the aggregates). The induced cells were then mixed and placed in the shaking incubator at 37^oC and 300 RPM for an additional 90 min. Aliquots were spotted on glass cover-slips for confocal imaging. Depending on the level of surface protein expression and the nature of the associative protein, we observed aggregates ranging from 10^3 to 10^5 μm^3 in volume (see below). In some cases, aggregates were visible to the naked eye.

Control of aggregate size by control of expression levels

For many applications of bacterial clusters, the average size of the clusters is an important design parameter. We expected that cluster size would be sensitive to the amount of associative protein displayed on the cell surface. To test this idea, we reduced the levels of surface display of the SpyTag and SpyCatcher proteins by modifying their ribosome-binding sites (RBS). This approach allowed us to maintain inducible control of bacterial assembly while enabling separate control of the expression levels of each associative protein.

Starting from the arabinose-inducible constructs pBAD-ST and pBAD-SC, we engineered an RBS predicted to be significantly weaker than the wild-type sequence.[33] Expression levels were quantified by immunostaining and subsequent flow cytometry (**Figure 3.1A**). We found that use of the weaker RBS led to an approximately four-fold decrease in surface expression for both SpyTag and SpyCatcher. All combinations of the wild-type and attenuated RBS constructs were then subjected to aggregation conditions and imaged (**Figure 3.1 B-F**). Automated image analysis showed that aggregate size could be varied by more than an order of magnitude in volume by control of expression level (**Figure 3.1B**).

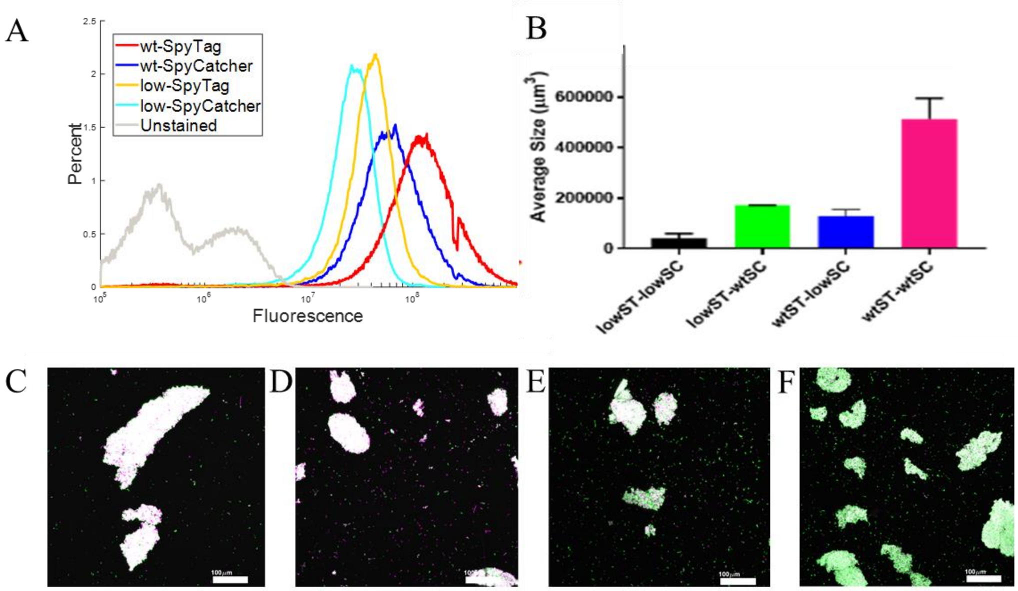


Figure 3.1: Expression levels control size of bacterial aggregates. **(A)** Flow cytometry analysis enables quantification of protein expression levels; mutant RBS's lead to an approximate 4-fold decrease in expression levels. **(B)** Volume-weighted averages of aggregate sizes. **(C)** Aggregation of wt-SC (magenta) and wt-ST (green). **(D)** Aggregation of wt-SC (magenta) and low-ST (green). **(E)** Aggregation of low-SC (magenta) and wt-ST (green). **(F)** Aggregation of low-SC (magenta) and low-ST (green)

Unless otherwise mentioned, the pQE-80-based plasmids with the wild-type RBS are used in the remainder of this work in order to demonstrate principles of programmable assembly with the highest expression levels available.

Dissociation of bacterial aggregates

Two different mechanisms – the physical association of SynZip leucine zippers and the formation of covalent isopeptide bonds between SpyTag and SpyCatcher – drive cellular aggregation in the systems introduced here. We expected SynZip-mediated aggregation to be reversible in the presence of excess (soluble) competing protein, and the SpyTag/SpyCatcher aggregation to be irreversible due to the permanence of the isopeptide covalent bonds over the experimental time-scale. To test these expectations, soluble SynZip 17 and SpyCatcher proteins containing just the associative domain and an elastin linker as a solubility tag were expressed in *E. coli* and purified via Ni-NTA affinity chromatography using methods described in previous work.[22] Aggregates mediated by the SynZip system were formed from DH10b *E. coli* containing pAT-17 and pAT-18 plasmids, and then the soluble SynZip 17 protein was added to a final concentration of 0, 0.001, or 0.1 mg/mL in LB medium. Three biological replicates were examined for each disaggregation condition. Representative micrographs and aggregate sizes (reported as volume-weighted averages, see Methods) are shown in **Figure 3.2A-D**. Titration of soluble SynZip17 into SynZip17/18 cultures decreased the size of the aggregates in a dose-dependent manner, consistent with the hypothesis that aggregated cells are bridged by specific biomolecular interactions.

In contrast, when soluble SpyCatcher was added to aggregates mediated by SpyTag/SpyCatcher interactions, no significant changes in aggregate size were observed, even when 1 mg/mL of protein was added (**Figure 3.2E-H**). This result suggests that the clusters mediated by SpyTag and SpyCatcher are held together by covalent bonds that cannot be disrupted by introduction of a competing protein. Clusters mediated by SpyTag and SpyCatcher are also substantially larger than those mediated by reversible SynZip interactions, as these clusters are more stable to shear-induced disruption.[34]

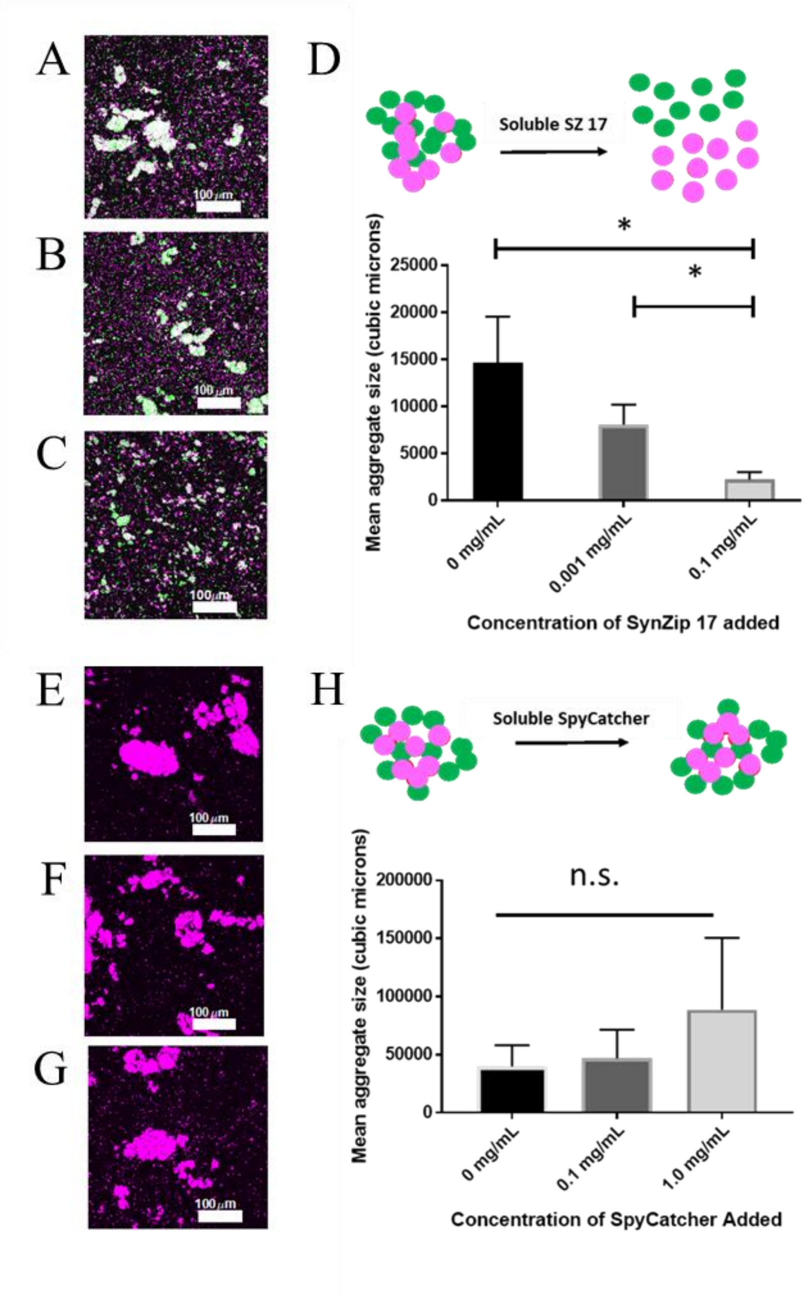


Figure 3.2: Dissociation and non-dissociation of bacterial aggregates by the addition of competing recombinant protein **(A)** pAT-Z17 (green) and pAT-Z18 (magenta) cells aggregate when IPTG is added. These cells can be disassociated by the addition of **(B)** 0.01 mg/mL soluble Z17 and **(C)** 0.01 mg/mL soluble Z17. **(D)** Volume-weighted average aggregate sizes of the resulting suspension (* $p<0.05$ by one-sided Student's T-test). By comparison, SpyTag/SpyCatcher aggregates do not dissociate when competitor protein is added. **(E)** pAT-ST and pAT-SC cells (magenta) aggregate when IPTG is added. These cells do not dissociate when **(F)** 0.1 mg/mL soluble SpyCatcher protein or **(G)** 1 mg/mL

soluble SpyCatcher protein is added. **(H)** Quantification of average aggregate sizes. Differences are not significantly different.

Formation of core-shell architectures

Many potential applications of bacterial aggregates require protection of a member of the consortium from environmental insult. We assembled mCherry-labeled cellular "cores" by the SpyTag/SpyCatcher interaction. We then added mWasabi-labeled cells that either expressed SpyCatcher or contained an empty pQE80 plasmid. After 30 min incubation, we found that cells carrying the SpyCatcher plasmid formed distinct green shells around magenta cores (**Figure 3.3A-B**). No core-shell structures were observed for control cultures (**Figure 3.3C-D**). These structures were characterized by line profiles drawn through the centroid of each aggregate. Core-shell aggregates showed increases in mWasabi fluorescence from the centroid to the surface of the aggregate, while control samples exhibited no correlation between radial location and fluorescence. Line profiles for individual core-shell clusters and a 63x image of a core-shell structure are shown in the supporting information (**Figure S3.3**).

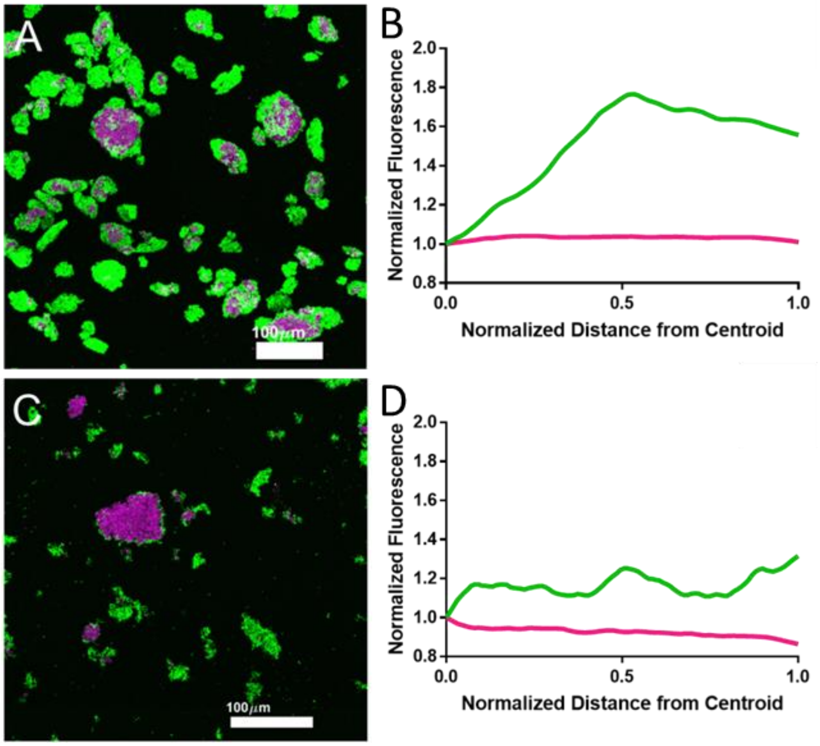


Figure 3.3: Formation of aggregates with a core-shell architecture **(A)** Cores of SpyTag and SpyCatcher surface-expressing bacteria, along with an mCherry fluorescent marker, were aggregated. Next, cells expressing SpyCatcher and mWasabi were added. A green shell can be observed around the magenta core **(B)** When mWasabi expressing cells that did not express surface protein were added to mCherry expressing cores, no green shell was observed. **(C)** Construction of line profiles from the center of the cores outward, and averaging over all aggregates in the image demonstrates that core-shell structures are formed, where magenta is observed in the middle of aggregates, with the green content increasing as the radial coordinate increases **(D)** If no core-shell structure is formed, there is no correlation between radial coordinate and fluorescence values. Scale bars correspond to 100 microns.

Triggering quorum sensing in clusters

To demonstrate the functional consequences of programmable microbial assembly, we investigated whether aggregation could be used to activate the LuxI-LuxR quorum-sensing circuit derived from *Vibrio harveyi*.[35-37] The LuxR-LuxI system is activated when LuxR binds the autoinducer N-(3-oxohexanoyl)-L-homoserine lactone (HSL), the concentration of which correlates with the density of bacterial cells (**Figure 3.4A**). We expected that aggregation could be used to increase the local concentration of HSL and trigger a quorum-sensing response under conditions where the average cell density in the system was below the threshold for quorum sensing.

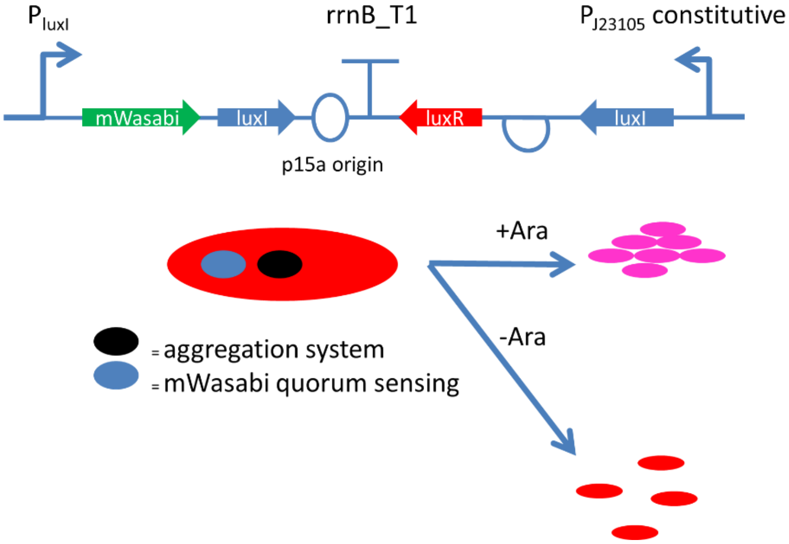


Figure 3.4: Illustration of quorum-sensing circuit **(A)** In the quorum-sensing circuit, the activator protein LuxR and acyl homoserine lactone (AHL) synthetase LuxI are driven by a constitutive promoter. When a sufficient concentration of AHL is present, LuxR is activated, binds to the pLuxI promoter, and recruits RNA polymerase, leading to expression of mWasabi as well as additional copies of LuxI, thus generating a positive-feedback loop. The E. coli strain has an integrated, constitutively-expressing mCherry cassette for confocal imaging. **(B)** In our scheme, upon aggregation, quorum sensing is activated, leading to the joint expression of mWasabi and mCherry (represented by pink cells). However, without induction of aggregation, the cells do not express mWasabi and express only mCherry.

We first tested this expectation by developing a reaction-diffusion model of quorum sensing. In the model, we consider an isolated aggregate of bacteria that can produce the autoinducer with positive feedback. The autoinducer diffuses within and out of the aggregate. We found that accumulation of the autoinducer in the aggregate is predicted to be much faster than its accumulation in the planktonic case (the unstructured curve) over a large range of parameter values (**Figure 3.5A, S3.6-7**), and that it was most strongly controlled by the magnitude of the Thiele modulus, which characterizes the length scale of the aggregate with respect to the diffusion and production rates (**Figure 3.5A**).[38] A large value of the Thiele modulus implies that the aggregate is large compared to the rate of diffusion and reaction, and as such, the autoinducer is able to accumulate strongly in the aggregate before being diluted by the bulk. By contrast, for small values of the Thiele modulus, the aggregate is small compared to the diffusion rate, and as such there is little accumulation in the aggregate. In the limit of small Thiele modulus, the model reduces to planktonic bacteria, where there is no benefit of the spatial structure. It is likely in this system that we are in the $\varphi>1$ regime, because the aggregates are closely packed leading to significantly restricted diffusion of substrates within the aggregates, and a correspondingly high φ.

To prepare the experimental strain, we first integrated a gene encoding mCherry into the *E. coli* DH10B chromosome under control of a T5 promoter to serve as a cell-marker.[39] Our quorum-sensing plasmid was a modification of pLuxRI2, which was a generous gift of the laboratory of Frances Arnold.[40] The quorum-sensing circuit is shown in **Figure 3.4**. Briefly, genes encoding the acylhomoserine lactone synthase (luxI) and the activator protein (luxR) were expressed under control of the constitutive pJ23105 promoter. To achieve positive feedback, a second copy of luxI was expressed (along with an mWasabi reporter gene) under control of a mutant pLuxI promoter to make plasmid pMTK3 (see Supplemental Discussion for details). This circuit was co-transformed with the SpyTag/SpyCatcher aggregation system to make strains sMTK1 (for the SpyTag-displaying strain) and sMTK2 (for the SpyCatcher-displaying strain).

Cultures of mixed sMTK1 and sMTK2 (at a 1:1 strain ratio) in mid-log phase (OD≈0.2) were split into three sub-cultures. The first sub-culture was induced with

0.2% arabinose, the second was left uninduced, and the third was treated with 2 mM HSL as a positive quorum sensing control; the experimental scheme is outlined in **Figure 3.4**. After 1 h, aggregates formed in the induced sub-culture, but not in the uninduced sub-culture. Aliquots of the induced and uninduced sub-cultures were imaged 75 min after induction; the results are shown in **Figure 3.5.**

In this experiment, the mCherry signal serves as a cell marker, the mWasabi signal as a marker for activation of the quorum sensing circuit. As shown in **Figure 3.5**, the aggregated sample appears to have more cells that are expressing mWasabi. To quantify this observation, we determined the Manders overlap coefficients (MOC) between mWasabi and mCherry channels, which represents the percentage of cells that have strong activation of their quorum-sensing circuits. After 75 min of induction, we find the MOC of the induced, aggregated sample to be 0.24 ± 0.018 (± SEM), while the uninduced sample MOC is 0.02 ± 0.0050. (**Figure 3.5H)**. We interpret this to mean that in the aggregated case, approximately 24% of cells display quorum-sensing behavior, whereas in the uninduced case, only 2% of the cells are quorum-sensing. This suggests that a substantial increase in quorum sensing can be observed by aggregating cells. Interestingly, the addition of exogenous HSL in figure 6H resulted in a MOC of 0.065 ± 0.0042, suggesting that aggregation resulted in a greater enhancement of quorum sensing than the addition of a large amount of exogenous inducer. This may be an artifact of image analysis (cells within aggregates are often brighter/easier to threshold) rather than the result of an actual physiological change.

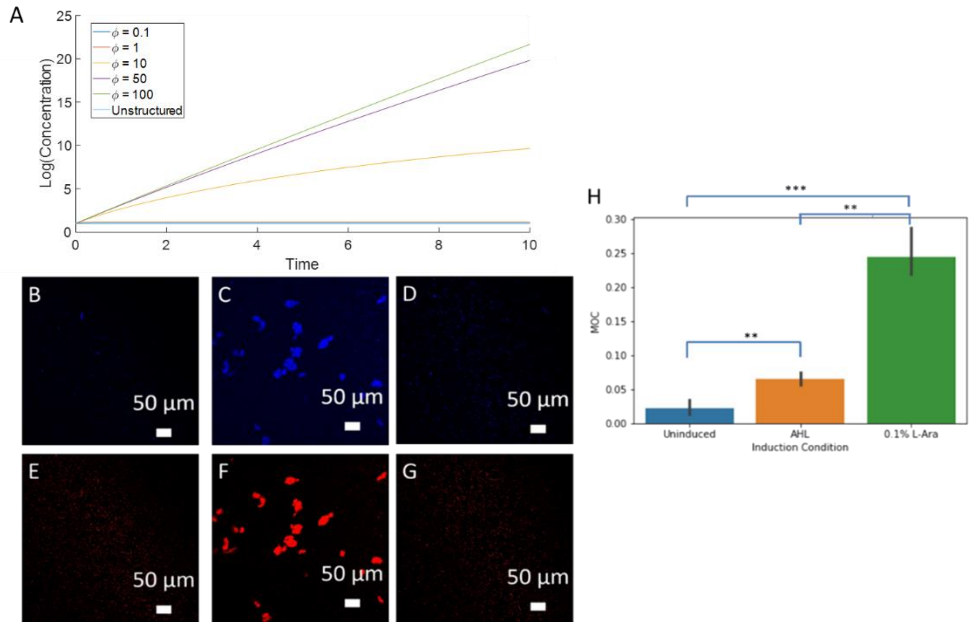


Figure 3.5: Aggregation leads to more-rapid activation of a quorum-sensing circuit **(A)** A reaction-diffusion model predicts that accumulation of AHL within aggregates is much faster than with planktonic cells; this effectiveness is a strong function of the Thiele modulus (ϕ <1 overlaps strongly with the planktonic case, as expected). **(B-G)** sMTK1 and sMTK2 cells were grown for 2 hours, then the culture was split into three different conditions. **(B,C)** Uninduced **(D,E)** Aggregation was induced with 0.1% L-arabinose. **(F,G)** Quorum sensing was induced by adding AHL. B, D, and E are the quorum sensing (mWasabi) signal, while C, E, and G are cell markers (mCherry). **(H)** Image analysis shows that 2% of uninduced cells show a quorum-sensing response, while 24% of cells that have aggregated show a quorum-sensing response, a 12-fold difference that is statistically-significant, as shown by three stars (***) above the respective bars (p=0.0065, using a two-tailed Student's t-test). Notably, in these experiments only 6.5% of cells induced with 2mM AHL show a quorum-sensing response. Three stars (***) indicate statistical significance at p<0.01 level, and two stars (**) indicate statistical significance at the p<0.02 level.

Conclusions

We have successfully demonstrated methods by which the size and architecture of bacterial aggregates can be controlled in a genetically-programmable manner. By choosing the appropriate associative protein, we can control the reversibility/irreversibility of the aggregates; while by controlling the surface expression levels of the associative proteins we can control the resulting size of aggregates. We have further demonstrated the construction of a core-shell architecture which may be useful in protecting sensitive bacteria from environmental insult or for creating complex biocatalysts. This work therefore represents an important step towards recapitulating the complex structures exhibited by natural microbial consortia, as well as a method by which cellular behavior can be made dependent on aggregation state through the use of a quorum sensing circuit.[41]

Methods developed in this work may enable the production of structured whole-cell biocatalysts, whereby multi-step reactions may be performed in series in bacterial aggregates, enabling enhanced intermediate channeling between cells in a manner complementary to previous work for substrate channeling within cells.[42-43] Particularly where biosynthetic steps may be difficult or impossible to place in the same cell, performing these steps in aggregates will enable channeling between metabolic steps without dilution into the bulk solution phase. Triggering of quorum sensing may play an important role here, as in order to minimize off-target reactivity, enzyme expression can be efficiently linked to aggregation.

We believe that the methods developed in this work are broadly applicable to other species of microbes and extendible to other types of protein-association domains. All that is required is an effective method of cell-surface display of the appropriate associative domain in the bacterial species of interest. In this manner, multi-species consortia of microbes may be established. Finally, similar techniques have recently been used for programmable surface binding[19] and may similarly be used for immobilization into protein hydrogels to form artificial biofilms.

Acknowledgements

This work was supported by Defense Advanced Research Projects Agency Biological Robustness in Complex Settings Contract HR001-15-C-0093. M.T.K. was supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program. B.R.S. is supported by NIH Training Grant 1T32GM112592. C.P.J was supported by a SURF fellowship.

Plasmid pHEA-Cm, from which the autotransporter was cloned, was a kind gift of Luis Angel Fernandez Herrero of the Centro Nacional de Biotecnologia (Madrid, Spain). Plasmid pLuxRI2, from which we cloned the quorum-sensing plasmid, was a gift of Frances Arnold of the California Institute of Technology (Pasadena, California).

pKPY680, pKPY681, and DH10B strain KY36 were a generous gift of Dr. Kai P. Yuet of Stanford University (Stanford, CA).

We thank Dr. Andres Collazo for assistance with confocal microscopy. Imaging was performed in the Biological Imaging Facility, with the support of the Caltech Beckman Institute and the Arnold and Mabel Beckman Foundation. We thank Yaron Antebi for use of EasyFlow for flow cytometry analysis. We thank Adam Silverman for a critical review of the manuscript.

Materials and Methods

General

Restriction enzymes, ligase, and Q5 DNA polymerase were purchased from New England Biolabs (Beverly, NJ). Nickel NTA was purchased from Qiagen (Hilden, Germany). DNA oligos and G-blocks were purchased from Integrated DNA Technologies (Coralville, IA).

Bacterial strains

All experiments were conducted in *E. coli* strain DH10B, obtained from Invitrogen (Carlsbad, CA). Aggregation for quorum sensing was conducted in *E. coli* strain KY36, a derivative of DH10B which contains a chromosomally-integrated mCherry under control of a leaky T5 promoter.

Plasmid Subcloning

Recombinant fusion proteins were produced by standard recombinant DNA technology. DH10b or Mach1 *Escherichia coli* were used for all cloning steps.
Genes encoding soluble Z17 and SpyCatcher proteins along with elastin solubility/stability tags have been previously cloned by our group into modified pQE-80L plasmids.[22]

Plasmids pKPY680 and pKPY681, which constitutively express mWasabi and mCherry, respectively, were constructed using mWasabi-N1/pmCherry-N1 as the template. Primers were ordered to amplify mWasabi/mCherry as well as add NsiI-J23100 promoter-SpeI-RBS-MRGS-6xHis to the 5’ end of mWasabi/mCherry, and to add HindIII to the 3’ end. This fragment was inserted into pBAD33 using NsiI and HindII sites.

To make the surface-expression constructs, the autotransporter domain downstream of the pelB leader sequence was amplified from pHEA by PCR using a PhusionII polymerase (NEB) with the addition of a 5’ XhoI and 3’ HindIII site to the autotransporter construct, which was then digested and inserted into a modified pQE-80L plasmid. Another G-block was ordered with EcoRI and XhoI sites that contained the T5 promoter, pelB, a 6xHis tag, and the protein of interest (SpyTag, SpyCatcher, SynZip17, SynZip18). A schematic of the autotransporter cassette is shown below (**Figure S3.1**).

The autotransporter constructs were also placed under the araBAD promoter to enable tighter control of the aggregation systems. The autotransporter-associative domain fusions were PCR amplified and inserted into pBAD33 using Gibson isothermal assembly.

The RBS mutant constructs were obtained from the pBAD33-based aggregation constructs by Quik-Change site-directed mutagenesis.

Plasmid pLuxRI2 was a generous gift from the lab of Prof. Frances Arnold. To make plasmid construct pMTK1, we first replaced the pLac/Ara1 promoter in that plasmid with a constitutive pJ23105 promoter. A DNA duplex containing the reverse complement of the pJ23105 promoter, and EcoRI and XhoI sites on the 5’ and 3’ ends, respectively, was ordered from IDT and inserted into pLuxRI2 following digestion with EcoRI and XhoI. The second modification required was the insertion of the quorum sensing cassette consisting of mWasabi and an additional copy of luxI synthetase under the control of the P_{luxI} promoter, as well as a p15a origin. This cassette was supplied by a gBlock gene fragment ordered from IDT, and contained restriction sites for SacI and AvrII on the 5’ and 3’ ends, respectively. Plasmid pLuxRI2 was then digested with SacI and AvrII, allowing for the insertion of the gBlock fragment.

Fluorescent proteins were chromosomally integrated using the pOSIP clonetegration system. Genes encoding mWasabi and mCherry under the control of the T5 promoter were PCR amplified and assembled into pOSIP-KO (Addgene). Z-competent E. coli MegaX DH10B T1R cells were mixed with the unpurified assembly reaction and spread on 2xYT agar plates supplemented with 35 mg/L kanamycin sulfate.

Flow cytometry

Measurement of surface expression levels was done using direct immunocytochemistry and flow cytometry. Overnight cultures of autotransporters were diluted 100x, and were grown to an optical density of approximately 0.6 prior to induction with 0.1% L-Arabinose. Expression was allowed to proceed for 90 minutes, after which the culture was centrifuged and blocked for 30 minutes with agitation (3% BSA in PBS). Cells were then centrifuged and resuspended in staining solution (5 µg/mL Anti-His conjugated Alexa-Fluor 488 Antibody (HIS.H8 Thermofisher), 1% BSA in PBS). This solution was then agitated for 1 hour, after which the cells were washed three times in PBS. Cells were strained through a 40 µm filter to remove aggregates and run on a MoFlo XDP cell sorter equipped with a 488 nm laser. Flow cytometry data were analyzed using EasyFlow.[44]

Quorum sensing

N-(B-Ketocaproyl)-DL-Homoserine Lactone (synonymous with N-(3-Oxohexanoyl)-Homoserine Lactone) was purchased from MilliporeSigma (Milwaukee, WI) and used without further modification.

Characterization of quorum sensing in bulk samples was done on a VarioSkan LUX instrument (ThermoFisher, San Diego CA). Cultures containing pMTK1, pMTK2, and pMTK3 were grown in LB medium supplemented with 35 mg/L chloramphenicol. The overnight cultures were then used to inoculate 150µL cultures at a ratio of 100:1 in flat-bottomed, clear 96-well plates with a lid (BD Falcon, Corning Inc, Corning, NY). The cultures then had varying amounts of AHL added (0-2mM), and the plate was incubated, with shaking, at 37 degrees for 18 hours. OD600 as well as mWasabi fluorescence (ex. 485 em. 515) was measured every 10 minutes. The results of this characterization are presented in the Supporting Discussion and **Figures S3.4** and **S3.5**

Expression of soluble SynZip and SpyCatcher proteins

Constructs were transformed into BL21 *E. coli* for expression. Expression was performed in Terrific Broth (12 g/L casein, 24 g/L yeast extract, 0.4% w/v glycerol, 0.017 M monobasic potassium phosphate, 0.072 M dibasic potassium phosphate). Cultures were induced at an optical density of 0.6-0.9 to a final concentration of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Expression was allowed to proceed for 5 h, after which cells were harvested by centrifugation.

For Z17 purification, cultures were resuspended in lysis buffer in denaturing lysis buffer (8 M urea, 0.1 M Na_2HPO_4, 10 mM imidazole; pH 8.0), and lysed by sonication. Lysates were cleared by centrifugation and incubated with NiNTA. The resin was washed with lysis buffer followed by wash buffer (8 M urea, 0.1 M Na_2HPO_4, 25 mM imidazole; pH 6.3). Protein was eluted with elution buffer (8 M urea, 0.1 M Na_2HPO_4, 250 mM imidazole; pH 3.5). Purity was confirmed with SDS-PAGE. Proteins were then extensively dialyzed against water and lyophilized for storage.

SpyCatcher was purified under native conditions. Cultures were resuspended in native lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, 1 mg/mL lysozyme; pH 8.0) Cells were lysed by sonication, and cleared lysates were incubated with NiNTA. The resin was washed with native wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 25 mM imidazole; pH 8.0) and eluted with native elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole; pH 8.0). Purity was confirmed with SDS-PAGE and purified SpyCatcher was dialyzed against water, and lyophilized.

Image Acquisition and Analysis

Microscopy images were taken on a Zeiss 800 LSM inverted confocal microscope (Carl Zeiss AG, Oberkochen, Germany).
All image analysis was performed using custom Matlab scripts.

Aggregate size analysis was performed similarly to what we described previously.[22] Briefly, confocal z-stacks were manually thresholded based on the intensity in each fluorescent channel. Pixels above the threshold were described as "bright." Contiguous "bright" pixels (in 3D) were identified, and the observed volume of each aggregate was determined. The volume-weighted average volume of each sample was determined using the following equation:

$$\bar{V} = \frac{\sum V_i^2}{\sum V_i}$$

where sums are taken over all of the aggregate volumes. This average represents the volume of the aggregate that the average bacterium would be found in, and is more appropriate than the number-weighted average, which is dominated by disassociated bacteria.

Core-shell fluorescence profiles were created as described previously.[22] Maximum intensity projections of the images were taken, and large aggregates were identified using thresholding. For each large aggregate, z-stacks with high levels of fluorescence were combined into a mean intensity projection. Then, starting at the centroid of the mean intensity projection, 100 radii representing equally spaced direction vectors were drawn to the edge of the aggregate, extracting the fluorescence intensities of each channel. Fluorescence intensities were scaled in each channel (with the maximum intensity in the aggregate being 1), and plotted along a radial axis where 0 represents the centroid and 1 represents the edge of the aggregate.

All image analysis code and reaction diffusion models can be downloaded at https://tirrell-lab.caltech.edu/code

Supplemental Discussion

Re-engineering of Quorum Sensing System

In preliminary experiments, we determined that the initial quorum-sensing circuit began turning on at an OD_{600} of approximately 0.4 (Figure S4). In order to emphasize the use of aggregation at turning on the circuit, we decided to systematically reduce the sensitivity of the quorum-sensing circuit to HSL, using point mutations identified in Antunes et al.[45] In particular, mutations to the quorum-sensing promoter (Lux box) C5A and C16A were made, measured, and noted to effectively reduce the quorum-sensing response even at high ODs (**Figure S3.4**). We confirmed that the C16A mutant was still responsive to high HSL (**Figure S3.5**), so this construct was used for all further experiments in this work.

Description and Derivation of Quorum Sensing Model

A model for the quorum sensing genetic circuit in bacterial aggregates was developed from the differential species balance with reaction on the autoinducer. Generically, diffusion of molecules in a dilute, non-convective, reacting system can be written as (Eq S3.1):[46]

$$\frac{\partial C_A}{\partial t} = D_A \nabla^2 C_A + R_A \qquad (S3.1)$$

where C_A is the concentration of the species (here the autoinducer), D_A is the effective diffusion coefficient of the species in the system, ∇^2 is the Laplacian operator, and R_A is the instantaneous rate of generation (or consumption) of the species. Here, we will assume that the rate of generation (production) of the autoinducer can be described as a constant term plus a Hill function to represent the positive feedback in the system (S3.2). This is a common formalism for transcriptional activation,[47] and has previously been used to characterize the LuxI promoter.[48]

$$R_A = \beta_1 + \beta_2 \frac{C_A^n}{K^n + C_A^n} \qquad (S3.2)$$

where β_1 is the zeroth order rate constant for leaky/constitutive expression, β_2 is the rate constant for the Hill function representing transcriptional activation by the autoinducer in positive feedback, and K and n are the Hill Function equilibrium constant and coefficient respectively. The full PDE that describes reaction-diffusion in the bacterial aggregates is:

$$\frac{\partial C_A}{\partial t} = D_A \nabla^2 C_A + \beta_1 + \beta_2 \frac{C_A^n}{K^n + C_A^n} \qquad (S3.3)$$

Next, we consider the boundary and initial conditions of the system. Initially, upon aggregation, we expect that the concentration of the autoinducer will be uniform throughout the aggregate and the bulk, which we will denote as C_{A0}^B. For simplicity, we assume that the aggregates are spheres, with radius R. Then, in order to retain finite concentration, the flux at the center of the sphere (r=0) must be 0. Finally, we write an interfacial mass transport equation that applies at the edge of the sphere. The boundary/initial conditions are written as S3.4-6

$$C_A(t = 0, r) = C_{A0}^B \qquad (S3.4)$$

$$\frac{\partial C_A(t, r = 0)}{\partial r} = 0 \qquad (S3.5)$$

$$D_A \nabla \mathrm{C_A}(t, r = R) = k_c(C_A^B(t) - C_A(t, r = R)) \qquad (S3.6)$$

where ∇ is the gradient operator, k_c is the interfacial mass transport coefficient, and $C_A^B(t)$ is the (time-dependent) bulk concentration of autoinducer. To further simplify the boundary condition at the surface, we assume that the mass transport coefficient is large, such that there is negligible interfacial resistance to mass transport. This is reasonable because the aggregates are being vigorously mixed, such that the mass transport resistance is likely to be dominated by the dense network of cells in the aggregate (this corresponds to Biot number>>1). Equation S3.6 then becomes

$$C_A(t, r = R) = C_A^B(t) \qquad (S3.7)$$

Equations S3.3-5 and S3.7 then are a fully defined PDE inside the aggregate. Before solution, we nondimensionalize as follows:

$$\hat{r} = \frac{r}{R} \tag{S3.8}$$

$$\widehat{C_A} = \frac{C_A}{C_{A0}^B} \tag{S3.9}$$

$$\hat{t} = \frac{C_{A0}^B}{\beta_1} \tag{S3.10}$$

If we rewrite the system with nondimensional variables, we obtain the following:

$$\frac{\partial \widehat{C_A}}{\partial \hat{t}} = \frac{D_A\, C_{A0}^B}{R^2 \beta_1}\, \nabla^2 \widehat{C_A} + 1 + \frac{\beta_2}{\beta_1} \frac{\widehat{C_A}^n}{\left(\frac{K}{C_{A0}^B}\right)^n + \widehat{C_A}^n} \tag{S3.11}$$

We will define the following dimensionless parameters:

$$\phi = R\sqrt{\frac{\beta_1}{D_A C_{A0}^B}} \tag{S3.12}$$

$$\hat{\beta} = \frac{\beta_2}{\beta_1} \tag{S3.13}$$

$$\widehat{K} = \frac{K}{C_{A0}^B} \tag{S3.14}$$

ϕ is the 0th order Thiele modulus for a sphere,[38] $\hat{\beta}$ represents the ratio of the strength of the activated promoter to constitutive/leaky expression, and $\widehat{K}$ represents the equilibrium constant for the Hill function for the promoter, expressed in units of the initial concentration. The PDE becomes:

$$\frac{\partial \widehat{C_A}}{\partial \hat{t}} = \frac{1}{\phi^2}\, \nabla^2 \widehat{C_A} + 1 + \hat{\beta}\, \frac{\widehat{C_A}^n}{\widehat{K}^n + \widehat{C_A}^n} \tag{S3.15}$$

The initial/boundary conditions are now:

$$\widehat{C_A}(\hat{t} = 0, \hat{r}) = 1 \tag{S3.16}$$

$$\frac{\partial \widehat{C_A}(\hat{t}, \hat{r} = 0)}{\partial \hat{r}} = 0 \tag{S3.17}$$

$$\widehat{C_A}(\hat{t} = 0, \hat{r}) = \widehat{C_A^B}(t) \qquad (S3.18)$$

Finally, we consider $\widehat{C_A^B}(t)$, (i.e. the time evolution of the bulk concentration). The amount of accumulation in the bulk is the total flux through the aggregate surface. Alternatively, it can be calculated as the total amount of autoinducer produced in the aggregate net of the change in the integrated concentration inside the aggregate. Writing this for a discrete time step:

$$\widehat{C_A^B}(\hat{t}) = \widehat{C_A^B}(\hat{t} - \Delta t) + \chi_v(\Delta t \int \widehat{R_A} dV - (\int \widehat{C_A}(\hat{t}) dV - \int \widehat{C_A}(\hat{t} - \Delta t) dV))$$

(S3.19)

where integrals are taken over the entire aggregate and χ_v is the volume fraction of the aggregates in solution.

Equations S3.15-S3.19 constitute a full mathematical description of the system. In order to solve them for a given parameter set, S3.15-S3.18 are first numerically solved for a constant outer bulk concentration, and then the bulk concentrations as a function of time are solved subject to the solution to the PDE (Equation S3.19). The PDE is then re-solved, and the two steps are iterated back and forth until convergence is reached. Matlab's pdepe function is used for PDE solution.

An ODE describes time evolution of the quorum sensing signal in the bulk in the absence of aggregates. By analogy, we have:

$$\frac{d\widehat{C_A}}{d\hat{t}} = \chi_v \left(1 + \hat{\beta} \frac{\widehat{C_A}^n}{\widehat{K}^n + \widehat{C_A}^n}\right) \qquad (S3.20)$$

where the volume fraction factor is required to account for dilution into the bulk. Matlab's ode45 function is used for integration. Simulation code can be found at https://tirrell-lab.caltech.edu/Code.

Supplemental Tables and Figures

Table S3.1: Plasmids/Strain Designations used in this study

Name	Backbone/origin/promoter	Purpose
pKPY680	pBAD33/p15a/pJ23100	Constitutive expression of mWasabi
pKPY681	pBAD33/p15a/pJ23100	Constitutive expression of mCherry
pAT-17	pQE80/colE1/T5	IPTG-inducible expression of SynZip 17
pAT-18	pQE80/colE1/T5	IPTG-inducible expression of SynZip 18
pAT-ST	pQE80/colE1/T5	IPTG-inducible expression of SpyTag
pAT-SC	pQE80/colE1/T5	IPTG-inducible expression of SpyCatcher
pBAT-17	pQE60/colE1/araBAD	Arabinose-inducible expression of SynZip 17 and compatibility with pMTK1-3
pBAT-18	pQE60/colE1/araBAD	Arabinose-inducible expression of SynZip 18 and compatibility with pMTK1-3
pBAT-ST	pQE60/colE1/araBAD	Arabinose-inducible expression of SpyTag and compatibility with pMTK1-3

pBAT-SC	pQE60/colE1/araBAD	Arabinose-inducible expression of SpyCatcher and compatibility with pMTK1-3
pBAD-ST	pBAD33/p15a/araBAD	Arabinose-inducible expression of SpyTag
pBAD-SC	pBAD33/p15a/araBAD	Arabinose-inducible expression of SpyCatcher
pMTK1	pHTSUB-105/p15a/luxI	“wild-type” quorum sensing
pMTK2	pHTSUB-105/p15a/luxI	Less-sensitive quorum sensing plasmid.
pMTK3	pHTSUB-105/p15a/luxI	Less-sensitive quorum sensing plasmid.
sMTK1	DH10b pMTK3/pAT-ST	Strain capable of quorum sensing and aggregating
sMTK2	DH10b pmTK3/pAT-SC	Strain capable of quorum sensing and aggregating

Table S3.2: Protein Sequences

Protein:	**Sequence**
Z17 (soluble)	MRGSHHHHHHGSVDGSGSGSGSGSGANEKEELKSKKAELRNRIEQLKQKREQLKQKIANLRKEIEAYKGSGSGSGSGSGALDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLEHHHHHHKLC
SpyCatcher (soluble)	MRGSHHHHHHGSVDGSGSGSGSGSGAAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIDGSGSGSGSGSGALDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLEHHHHHHKLC
SpyCatcher-Autotransporter	MKYLLPTAAAGLLLLAAQPAMAMRGSHHHHHHGSVDGAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIDLETPTPGPDLNVDNDLRPEAGSYIANLAAANTMFTTRLHERLGNTYYTDMVTGEQKQTTMWMRHEGGHNKWRDGSGQLKTQSNRYVLQLGGDVAQWSQNGSD

	RWHVGVMAGYGNSDSKTISSRTGYRAKASVNGYSTGLYATWYADDESRNGAYLDSWAQYSWFDNTVKGDDLQSESYKSKGFTASLEAGYKHKLAEFNGSQGTRNEWYVQPQAQVTWMGVKADKHRESNGTLVHSNGDGNVQTRLGVKTWLKSHHKMDDGKSREFQPFVEVNWLHNSKDFSTSMDGVSVTQDGARNIAEIKTGVEGQLNANLNVWGNVGVQVADRGYNDTSAMVGIKWQF
SpyTag-Autotransporter	MKYLLPTAAAGLLLLAAQPAMAMRGSHHHHHHGSVDAHIVMVDAYKPTKLDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLETPTPGPDLNVDNDLRPEAGSYIANLAAANTMFTTRLHERLGNTYYTDMVTGEQKQTTMWMRHEGGHNKWRDGSGQLKTQSNRYVLQLGGDVAQWSQNGSDRWHVGVMAGYGNSDSKTISSRTGYRAKASVNGYSTGLYATWYADDESRNGAYLDSWAQYSWFDNTVKGDDLQSESYKSKGFTASLEAGYKHKLAEFNGSQGTRNEWYVQPQAQVTWMGVKADKHRESNGTLVHSNGDGNVQTRLGVKTWLKSHHKMDDGKSREFQPFVEVNWLHNSKDFSTSMDGVSVTQDGARNIAEIKTGVEGQLNANLNVWGNVGVQVADRGYNDTSAMVGIKWQF
Z17-Autotransporter	MKYLLPTAAAGLLLLAAQPAMAMRGSHHHHHHGSVDGSGSGSGSGSGSNEKEELKSKKAELRNRIEQLKQKREQLKQKIANLRKEIEAYKGSGSGSGSGSGSLETPTPGPDLNVDNDLRPEAGSYIANLAAANTMFTTRLHERLGNTYYTDMVTGEQKQTTMWMRHEGGHNKWRDGSGQLKTQSNRYVLQLGGDVAQWSQNGSDRWHVGVMAGYGNSDS

	KTISSRTGYRAKASVNGYSTGLYATWYADDESRNGAYLDSWAQYSWFDNTVKGDDLQSESYKSKGFTASLEAGYKHKLAEFNGSQGTRNEWYVQPQAQVTWMGVKADKHRESNGTLVHSNGDGNVQTRLGVKTWLKSHHKMDDGKSREFQPFVEVNWLHNSKDFSTSMDGVSVTQDGARNIAEIKTGVEGQLNANLNVWGNVGVQVADRGYNDTSAMVGIKWQF
Z18-Autotransporter	MKYLLPTAAAGLLLLAAQPAMAMRGSHHHHHHGSVDGSGSGSGSGSGSSIAATLENDLARLENENARLEKDIANLERDLAKLEREEAYFGSGSGSGSGSGSLETPTPGPDLNVDNDLRPEAGSYIANLAAANTMFTTRLHERLGNTYYTDMVTGEQKQTTMWMRHEGGHNKWRDGSGQLKTQSNRYVLQLGGDVAQWSQNGSDRWHVGVMAGYGNSDSKTISSRTGYRAKASVNGYSTGLYATWYADDESRNGAYLDSWAQYSWFDNTVKGDDLQSESYKSKGFTASLEAGYKHKLAEFNGSQGTRNEWYVQPQAQVTWMGVKADKHRESNGTLVHSNGDGNVQTRLGVKTWLKSHHKMDDGKSREFQPFVEVNWLHNSKDFSTSMDGVSVTQDGARNIAEIKTGVEGQLNANLNVWGNVGVQVADRGYNDTSAMVGIKWQF

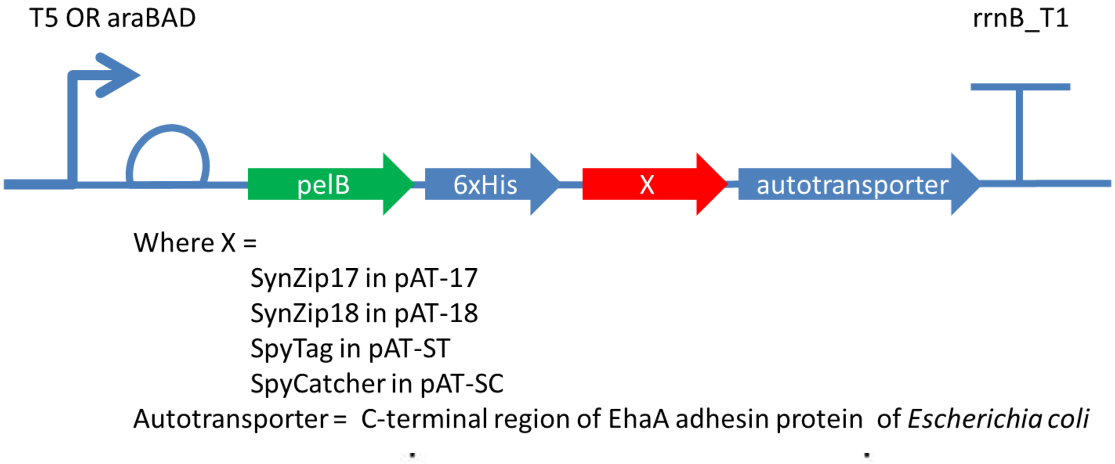


Figure S3.1: Schematic of aggregation cassette. Under the control of either a T5 promoter (in the pAT-X plasmids) or the araBAD promoter (in the pBAD-X plasmids), there is a signaling sequence pelB and a 6xHistidine tag upstream of the associative protein, followed by the C-terminal region of the EhaA adhesion protein of *Escherichia coli*.

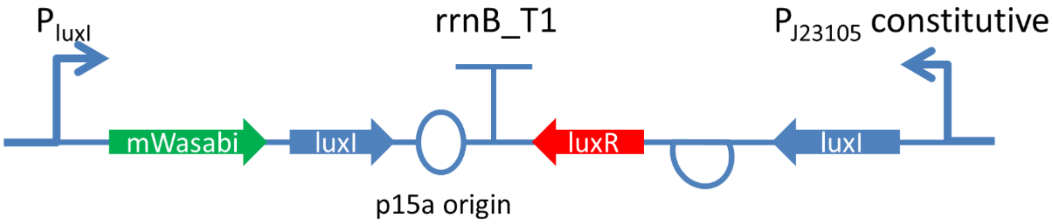


Figure S3.2: Schematic of quorum-sensing plasmid**.** On a plasmid with a p15a origin (to ensure compatibility with the aggregation plasmid), luxR activator protein and luxI AHL synthetase are under the control of a pJ23105 constitutive promoter. The LuxR protein, in the presence of a sufficient concentration of AHL, binds to the pLuxI promoter driving expression of mWasabi and an additional copy of luxI AHL synthetase giving rise to positive feedback.

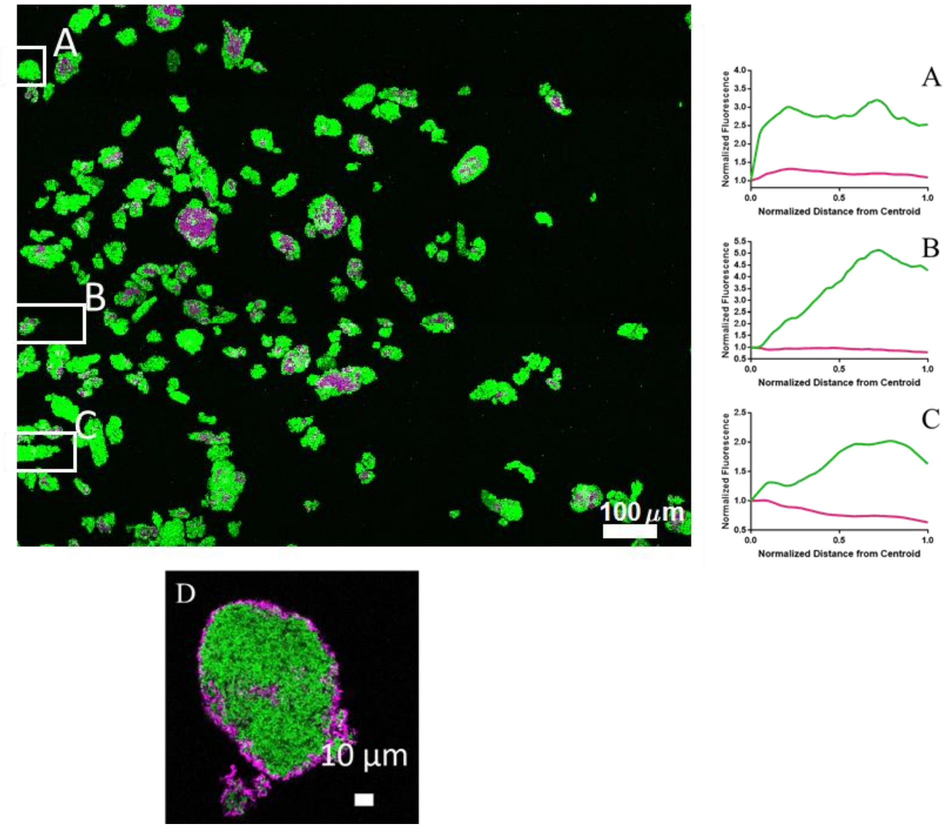


Figure S3.3: Individual line plots of core-shell structure **(A-C)** line profiles generated for individual clusters A, B, and C (as outlined in boxes in the main image, and line profiles in the right-hand panel). **(D)** Individual core-shell structure taken at 63x magnification.

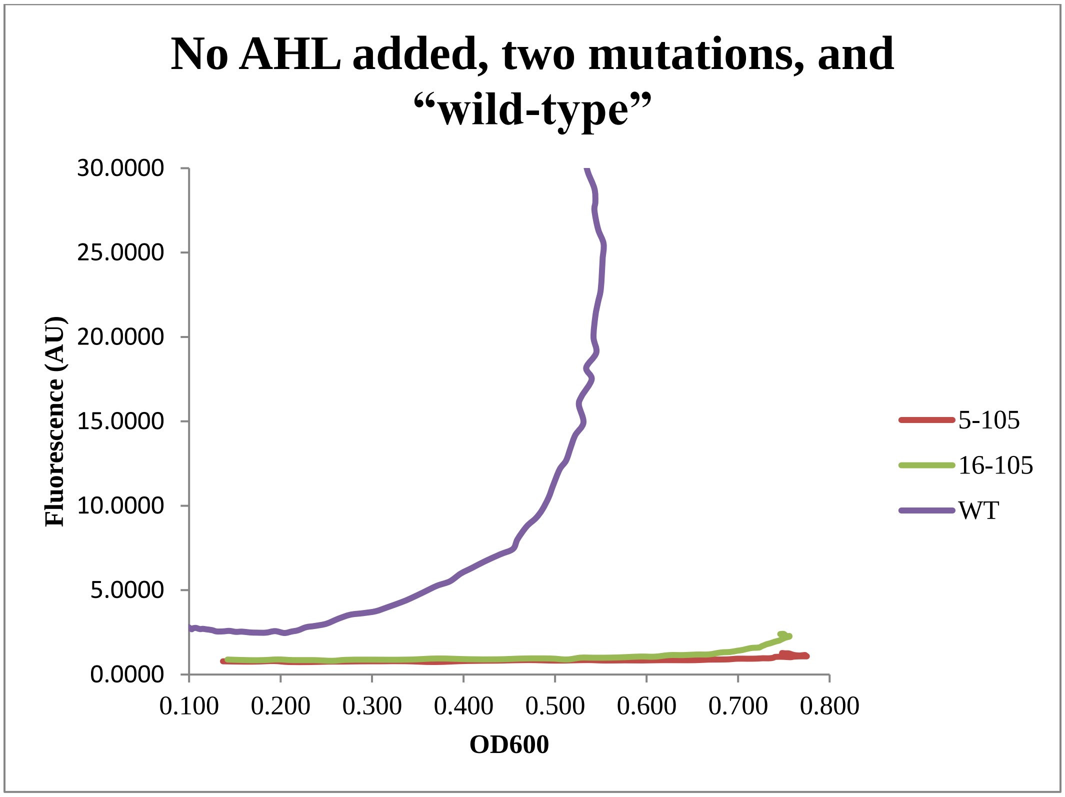


Figure S3.4: Characterization of Quorum Sensing System. The wild-type quorum sensing system turns on at an OD_{600} of approximately 0.4-0.5, making detecting aggregation differences challenging. By making C5A and C16A mutants, the circuits are much less sensitive to HSL.

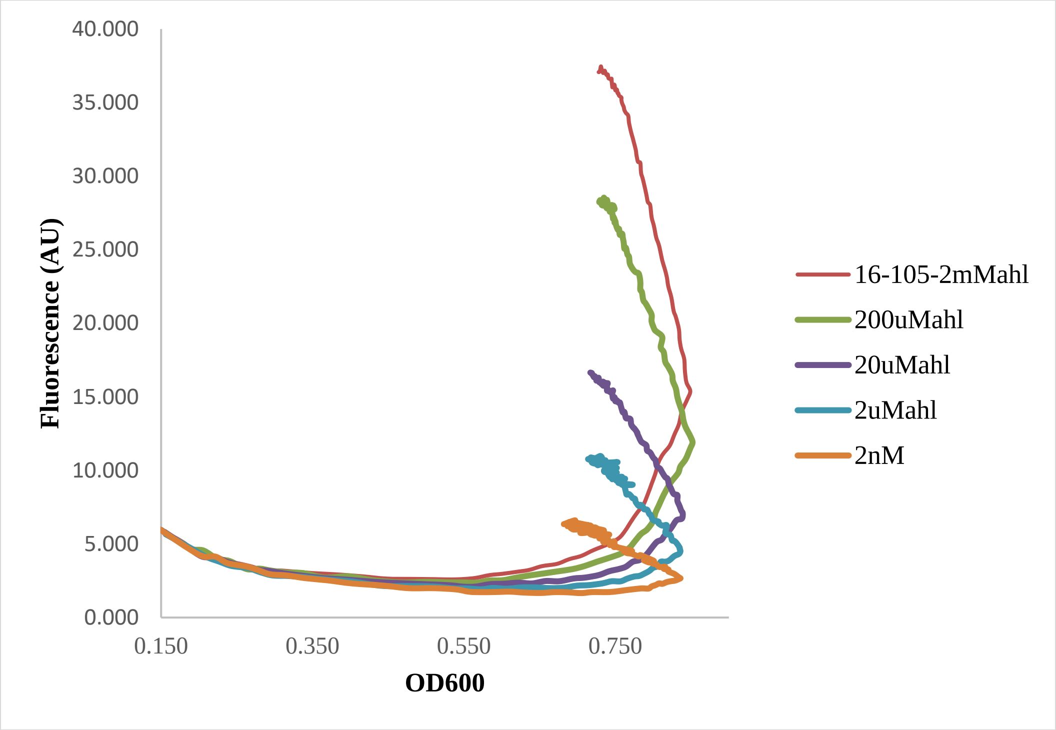


Figure S3.5: Exogenous Addition of HSL. Addition of exogenous HSL to the mutant C16A quorum sensing circuit still turns on the circuit, demonstrating that the circuit is still responsive to HSL, though at a lower level.

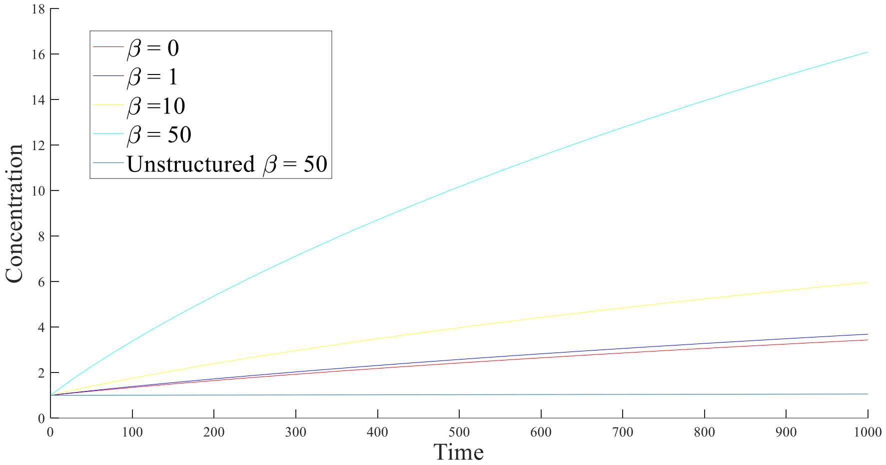


Figure S3.6: Effect of $\hat{\beta}$ on model. Over a large range of $\hat{\beta}$ (including in its absence, representing no positive feedback), accumulation of autoinducer is much faster than in the unstructured aggregate, which does not accumulate significantly on these time scales. (Other parameters: ϕ=10, $\widehat{K}$= 2, n=2)

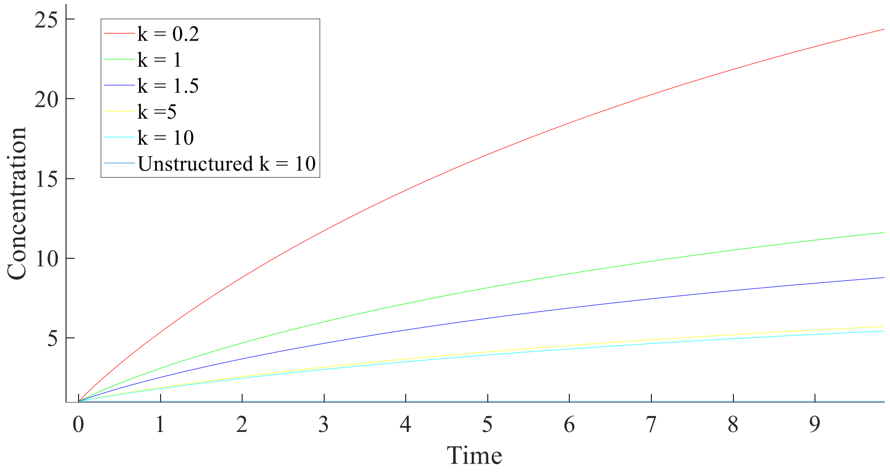


Figure S3.7: Effect of $\widehat{K}$ on model. Over a large range of $\widehat{K}$, accumulation of autoinducer is much faster than in the unstructured aggregate, which does not accumulate significantly on these time scales. (Other parameters: ϕ=10, $\widehat{B}$=5, n=2)

References

1 A. Boetius, K. Ravenschlag, C. J. Schubert, D. Rickert, F. Widdel, A. Gleseke, R. Amann, B. B. Jørgensen, U. Witte and O. Pfannkuche, A marine microbial consortium apparently mediating anaerobic oxidation methane, *Nature*, 2000, **407**, 623–626.

2 M. G. A. Van Der Heijden, R. D. Bardgett and N. M. Van Straalen, The unseen majority: Soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems, *Ecol. Lett.*, 2008, 11, 296–310.

3 M. Burmølle, J. S. Webb, D. Rao, L. H. Hansen, S. J. Sørensen and S. Kjelleberg, Enhanced biofilm formation and increased resistance to antimicrobial agents and bacterial invasion are caused by synergistic interactions in multispecies biofilms, *Appl. Environ. Microbiol.*, 2006, **72**, 3916–3923.

4 T. J. Wiles, M. Jemielita, R. P. Baker, B. H. Schlomann, S. L. Logan, J. Ganz, E. Melancon, J. S. Eisen, K. Guillemin and R. Parthasarathy, Host gut motility promotes competitive exclusion within a model intestinal microbiota, *PLoS Biol.*, 2016, **14**, 1–24.

5 A. C. Perez, B. Pang, L. B. King, L. Tan, K. A. Murrah, J. L. Reimche, J. T. Wren, S. H. Richardson, U. Ghandi and W. E. Swords, Residence of Streptococcus pneumoniae and Moraxella catarrhalis within polymicrobial biofilm promotes antibiotic resistance and bacterial persistence in vivo, *Pathog. Dis.*, 2014, **70**, 280–288.

6 Z. Liu, J. Müller, T. Li, R. M. Alvey, K. Vogl, N. U. Frigaard, N. C. Rockwell, E. S. Boyd, L. P. Tomsho, S. C. Schuster, P. Henke, M. Rohde, J. Overmann and D. A. Bryant, Genomic analysis reveals key aspects of prokaryotic symbiosis in the phototrophic consortium Chlorochromatium aggregatum, *Genome Biol.*, 2013, **14**, R127.

7 D. Cerqueda-García, L. P. Martínez-Castilla, L. I. Falcón and L. Delaye, Metabolic analysis of Chlorobium chlorochromatii CaD3 reveals clues of the

symbiosis in Chlorochromatium aggregatum, *ISME J.*, 2014, **8**, 991–998.

8 S. G. Hays, W. G. Patrick, M. Ziesack, N. Oxman and P. A. Silver, Better together: Engineering and application of microbial symbioses, *Curr. Opin. Biotechnol.*, 2015, **36**, 40–49.

9 L. M. Gieg, S. J. Fowler and C. Berdugo-Clavijo, Syntrophic biodegradation of hydrocarbon contaminants, *Curr. Opin. Biotechnol.*, 2014, **27**, 21–29.

10 E. Antoniou, S. Fodelianakis, E. Korkakaki and N. Kalogerakis, Biosurfactant production from marine hydrocarbon-degrading consortia and pure bacterial strains using crude oil as carbon source, *Front. Microbiol.*, 2015, **6.**

11 R. Du, J. Yan, S. Li, L. Zhang, S. Zhang, J. Li, G. Zhao and P. Qi, Cellulosic ethanol production by natural bacterial consortia is enhanced by Pseudoxanthomonas taiwanensis, *Biotechnol. Biofuels*, 2015, **8**, 10.

12 N. Bourdakos, E. Marsili and R. Mahadevan, A defined co-culture of Geobacter sulfurreducens and Escherichia coli in a membrane-less microbial fuel cell, *Biotechnol. Bioeng.*, 2014, **111**, 709–718.

13 K. Brenner, L. You and F. H. Arnold, Engineering microbial consortia: A new frontier in synthetic biology, *Trends Biotechnol.*, 2008, **26**, 483–9.

14 P. Y. Liao, C. W. Liu and W. Y. Liu, Bioaccumulation of mercury and polychlorinated dibenzo-p-dioxins and dibenzofurans in salty water organisms, *Environ. Monit. Assess.*, 2016, **188**, 1–15.

15 J. W. Chang, H. L. Chen, H. J. Su, P. C. Liao, H. R. Guo and C. C. Lee, Simultaneous exposure of non-diabetics to high levels of dioxins and mercury increases their risk of insulin resistance, *J. Hazard. Mater.*, 2011, **185**, 749–755.

16 H. J. Kim, W. Du and R. F. Ismagilov, Complex function by design using spatially pre-structured synthetic microbial communities: Degradation of pentachlorophenol in the presence of Hg(II), *Integr. Biol.*, 2011, **3**, 126–33.

17 I. Drachuk, R. Suntivich, R. Calabrese, S. Harbaugh, N. Kelley-Loughnane, D. L. Kaplan, M. Stone and V. V. Tsukruk, Printed dual cell arrays for

multiplexed sensing, *ACS Biomater. Sci. Eng.*, 2015, **1**, 287–294.

18 J. L. Murray, J. L. Connell, A. Stacy, K. H. Turner and M. Whiteley, Mechanisms of synergy in polymicrobial infections, *J. Microbiol.*, 2014, **52**, 188–199.

19 X. Jin and I. H. Riedel-Kruse, Biofilm lithography enables high-resolution cell patterning via optogenetic adhesin expression, *Proc. Natl. Acad. Sci.*, 2018, **115,** 3698-3703

20 D. S. Glass and I. H. Riedel-Kruse, A synthetic bacterial cell-cell adhesion toolbox for programming multicellular morphologies and patterns, *Cell*, 2018, **174**, 649-658

21 E. Veiga, V. De Lorenzo and L. A. Ferna, Autotransporters as scaffolds for novel bacterial adhesins : Surface properties of Escherichia coli cells displaying Jun / Fos dimerization domains, *J. Bacteriol.*, 2003, **185**, 5585–5590.

22 M. Obana, B. R. Silverman and D. A. Tirrell, Protein-mediated colloidal assembly, *J. Am. Chem. Soc.*, 2017, **139**, 14251–14256.

23 J. Maurer, J. Jose and T. F. Meyer, Autodisplay: One-component system for efficient surface display and release of soluble recombinant proteins from Escherichia coli, *J. Bacteriol.*, 1997, **179**, 794–804.

24 J. Jose and T. F. Meyer, The autodisplay story, from discovery to biotechnical and biomedical applications., *Microbiol. Mol. Biol. Rev.*, 2007, **71**, 600–19.

25 J. Schüürmann, P. Quehl, G. Festel and J. Jose, Bacterial whole-cell biocatalysts by surface display of enzymes: Toward industrial application, *Appl. Microbiol. Biotechnol.*, 2014, **98**, 8031–8046.

26 K. E. Thompson, C. J. Bashor, W. A. Lim and A. E. Keating, SYNZIP protein interaction toolbox: In vitro and in vivo specifications of heterospecific coiled-coil interaction domains, *ACS Synth. Biol.*, 2012, **1**, 118–129.

27 B. Zakeri and M. Howarth, Spontaneous intermolecular amide bond formation between side chains for irreversible peptide targeting, *J. Am. Chem. Soc.*,

2010, **132**, 4526–4527.

28 B. Zakeri, J. Fierer, E. Celik, E. Chittock, U. Schwarz-Linek, V. Moy and M. Howarth, Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin, *Proc. Natl. Acad. Sci.*, 2012, **109**, 690-697

29 W. Zhang, F. Sun, D. A. Tirrell and F. H. Arnold, Controlling macromolecular topology with genetically encoded SpyTag–SpyCatcher chemistry, *J. Am. Soc.*, 2013, **135**, 13988–13997.

30 F. Sun, W.-B. Zhang, A. Mahdavi, F. H. Arnold and D. A. Tirrell, Synthesis of bioactive protein hydrogels by genetically encoded SpyTag-SpyCatcher chemistry, *Proc. Natl. Acad. Sci.*, 2014, **111**, 11269–11274.

31 Z. Liu, H. Zhou, W. Wang, W. Tan, Y. X. Fu and M. Zhu, A novel method for synthetic vaccine construction based on protein assembly, *Sci. Rep.*, 2014, **4**, 1–8.

32 C. Schoene, J. O. Fierer, S. P. Bennett and P. Mark, SpyTag / SpyCatcher cyclization confers resilience to boiling on a mesophilic enzyme, *Angew. Chemie*, 2015, **53**, 6101–6104.

33 D. Na, S. Lee and D. Lee, Mathematical modeling of translation initiation for the estimation of its efficiency to computationally design mRNA sequences with desired expression levels in prokaryotes, *BMC Syst. Biol.*, 2010, **4**, 71.

34 T. Serra, J. Colomer and X. Casamitjana, Aggregation and breakup of particles in a shear flow, *J. Colloid Interface Sci.*, 1997, **187**, 466–473.

35 W. C. Fuqua, S. C. Winans and E. P. Greenberg, Quorum sensing in bacteria: The LuxR-LuxI family of cell density-responsive transcriptional regulators, *J. Bacteriol.*, 1994, 176, 269–275.

36 N. A. Whitehead, A. M. L. Barnard, H. Slater, N. J. L. Simpson and G. P. C. Salmond, Quorum-sensing in Gram-negative bacteria, *FEMS Microbiol. Rev.*, 2001, **25**, 365–404.

37 J. B. Langebrake, G. E. Dilanji, S. J. Hagen and P. De Leenheer, Traveling waves in response to a diffusing quorum sensing signal in spatially-extended

bacterial colonies, *J. Theor. Biol.*, 2014, **363**, 53–61.

38 E. W. Thiele, Relation between catalytic activity and size of particle, *Ind. Eng. Chem.*, 1939, **31**, 916–920.

39 F. St-Pierre, L. Cui, D. G. Priest, D. Endy, I. B. Dodd and K. E. Shearwin, One-step cloning and chromosomal integration of DNA, *ACS Synth. Biol.*, 2013, **2**, 537–541.

40 L. You, R. S. Cox, R. Weiss and F. H. Arnold, Programmed population control by cell-cell communication and regulated killing, *Nature*, 2004, **428**, 868–871.

41 S. Elias and E. Banin, Multi-species biofilms: Living with friendly neighbors, *FEMS Microbiol. Rev.*, 2012, **36**, 990–1004.

42 R. J. Conrado, G. C. Wu, J. T. Boock, H. Xu, S. Y. Chen, T. Lebar, J. Turnek, N. Tomšič, M. Avbelj, R. Gaber, T. Koprivnjak, J. Mori, V. Glavnik, I. Vovk, M. Beninča, V. Hodnik, G. Anderluh, J. E. Dueber, R. Jerala and M. P. Delisa, DNA-guided assembly of biosynthetic pathways promotes improved catalytic efficiency, *Nucleic Acids Res.*, 2012, **40**, 1879–1889.

43 R. J. Conrado, J. D. Varner and M. P. DeLisa, Engineering the spatial organization of metabolic enzymes: Mimicking nature's synergy, *Curr. Opin. Biotechnol.*, 2008, **19**, 492–499.

44 Y. E. Antebi, J. M. Linton, H. Klumpe, C. Su, R. Mccardell, M. B. Elowitz, Combinatorial signal perception in the BMP Pathway, *Cell*, 2017, **170**, 1184-1185

45 L. C. M. Antunes, R. B. R. Ferreira, C. P. Lostroh and E. P. Greenberg, A mutational analysis defines Vibrio fischeri LuxR binding sites, *J. Bacteriol.*, 2008, **190**, 4392–4397.

46 W. Deen, *Analysis of Transport Phenomena*, Oxford University Press, New York, 1st ed., 1998.

47 G. M. Süel, J. Garcia-Ojalvo, L. M. Liberman and M. B. Elowitz, An excitable gene regulatory circuit induces transient cellular differentiation, *Nature*, 2006, **440**, 545–550.

48 N. Rai, R. Anand, K. Ramkumar, V. Sreenivasan, S. Dabholkar, K. V. Venkatesh and M. Thattai, Prediction by promoter logic in bacterial quorum sensing, *PLoS Comput. Biol.* 2012, **8**.

Chapter 4

CONTROL OF MICROBIAL FLOCS

Abstract

Bacterial flocs are a common motif in nature and industrial applications of microbes (e.g., in water treatment). Flocculation of bacteria is driven by attractive forces, either direct cell-cell interactions or indirect colloidal forces (e.g. depletion or electrostatics). Previously, we developed a method to genetically engineer bacteria for programmable flocculation via expressing proteins on their surfaces that drive programmable assembly into aggregates. Here, using this approach, we investigate how properties of the resulting flocs can be controlled and how these relate to the underlying properties of the associative proteins and shear field. We demonstrate control of the assembly kinetics and equilibrium sizes of the resulting flocs over several orders of magnitude using different associating proteins and expression levels. This setup allows us to control the association potential between bacterial particles and investigate the flocculation of particles in this strong interaction limit. The ability to understand and control flocculation of microbes will enable their use in engineered biotechnological applications, including water treatment and biocatalysis.

Introduction

In nature, many bacteria live in close association with other microbes, in biofilms,[1] aggregates,[2] or sludges.[3] These structures lend emergent function to these (often multi-species) communities, as they are often able to resist environmental insults,[4] cross-feed essential metabolites,[5] or communicate through mechanisms such as quorum sensing[6] or electrical signaling.[7] Bacterial flocculation is also used extensively in industry to collect bacteria from a suspension, particularly in water treatment, usually by the addition of a polymeric flocculant.[8]

In any suspension of colloidal particles, flocculation occurs when particles collide and stick together. These aggregates are stable when the inter-particle attractive forces outweigh repulsive interactions. Colloidal attractive forces include depletion (excluded volume),[9] van der Waals forces,[10] electrostatic interactions,[11] or, in engineered systems, programmable molecular interactions.[12,13] Repulsive forces may include electrostatic repulsion,[14] or when aggregates are mixed, turbulent shear stresses.[15] When aggregation is reversible and dispersive forces can break the aggregates apart, an equilibrium distribution of floc size is established by the relative rates of aggregation and disassociation. This equilibrium size distribution is a complex function of the shear rate,[16] particle concentration in some cases,[17,18] and/or the composition of the floc.[19] This last factor, the nature of the particles making up the aggregate, is the least well-understood, in part because it is difficult to precisely tune inter-particle forces.

Previously, we have developed methods of programming the assembly of latex particles[13] and microbes (Chapter 3 of this thesis*)* by grafting associative proteins onto cell surfaces. Promisingly, we could modify the size distribution of the aggregates simply by controlling the amount of protein displayed on the surface of the colloidal particle. This result is important because the amount of displayed protein can easily be programmed using standard genetic engineering techniques.

Here, to better understand the impact of inter-particle forces on floc size distribution and other properties, we engineer a variety of strains of *Escherichia coli* that express varying concentrations of associative proteins on their surfaces. By measuring the equilibrium aggregate sizes and the dynamics of aggregation when the bacteria are mixed under shear, we demonstrate that floc size distributions can be understood by a set of simple empirical principles. Interestingly, we find that even in the strong interaction regime, where interaction energies far exceed *kT*, aggregate sizes are finite, likely due to very strong dispersive shear stresses as a result of turbulent mixing. Our results suggest that the mechanism of aggregate disassociation is due to fracture of smaller aggregates, rather than single particle dissociation, implying that dispersive shear stresses act on the aggregates, and not on single cells. Depending on the type of interaction programmed, the size distributions of aggregates may be under kinetic or (pseudo-)thermodynamic control. Finally, even within the same class of interactions (e.g. leucine zippers), the size and dynamic properties of the resulting aggregates may be controlled by tuning the affinity and stimuli-responsive properties of displayed proteins.

Results and Discussion

Associative Proteins

We have previously shown the ability to drive aggregation of the model bacterium *E. coli* by surface expression of associative proteins (Chapter 3 of this thesis). To do this, we fused the associative protein of interest to the N-terminus of an autotransporter protein. Upon cellular expression, the autotransporter fusion proteins traffic to the outer membrane and display their fusion partners on the surface of the cell (See **Figure 4.1A** for schematic).[20] Previously, we displayed the SpyTag/SpyCatcher domains, which form an irreversible isopeptide bond between the proteins' associative domains between a lysine residue in SpyCatcher and an aspartate residue in SpyTag. This interaction allows the "permanent" linkage of bacteria or particles displaying the cognate proteins, and has also been used to control protein topology, form protein hydrogels, enhance enzyme stability, and enable analysis of subcellular trafficking.[21–24] In this work, we also utilize the “A” designed leucine zipper protein with the purpose of accessing a different range of interaction strengths.[25] Upon association, A coils form anti-parallel tetrameric helical bundles, though dimeric interactions may also occur.[26,27] We have previously used this protein domain primarily for physically crosslinked hydrogels,[25,26,28] but we reasoned that it may also be useful for programming association of colloidal particles.

Control of Expression Levels

In our previous work, we showed that the aggregate sizes may be controlled to some extent by manipulating protein expression levels through tuning of the ribosome binding site (RBS) responsible for translation initiation. In that work (Chapter 3 of this thesis), we rationally designed a mutant RBS (here denoted as Mut2, sequence GAGCGA) that reduced the expression level of SpyTag/SpyCatcher constructs by approximately 75% as measured by immunocytochemistry against poly-histidine epitopes on the displayed proteins, followed by flow cytometry. However, further attempts at rational engineering to create more graded expression levels could not further decrease SpyTag/SpyCatcher synthesis without completely ablating expression of the associative proteins.

Instead, to create a graded set of expression constructs, we created libraries of RBSs driving translation of the autotransporter fusions. Separate libraries were created for each associative protein, as RBS strength is protein dependent due to mRNA secondary structure.[29] The entire Shine-Dalgarno (SD) sequence of the RBS was randomized using degenerate primers, and the resulting libraries were transformed into *E. coli*. We induced expression of the associative proteins in each strain, then iteratively sorted the cells by their expression levels using immunocytochemistry and fluorescence-activated cell sorting (FACS). The process was repeated several times in order to enrich the sub-libraries for intermediate expression levels (**Figure 4.1B**), and then individual clones are isolated, sequenced, and expression levels measured (**Figure 4.1C**). Overall, this process identified RBSs that span approximately two orders of magnitude in expression level—effectively limited by the dynamic range of the flow cytometer. The observed strengths are also well-correlated with the predicted expression levels from the RBS Calculator (**Figure 4.1D**).[29] All RBS sequences along with their corresponding measured strengths are found in **Table S4.2.**

We generally did not observe any RBSs stronger than the wild-type, consensus SD sequence from *E. coli* (AGGAGG), and many non-consensus sequences similarly exhibit "full" expression levels. This is likely because strong RBSs enable sufficient expression to saturate the secretion apparatus of the cell, removing the constraint of translation rate on the observed expression levels.

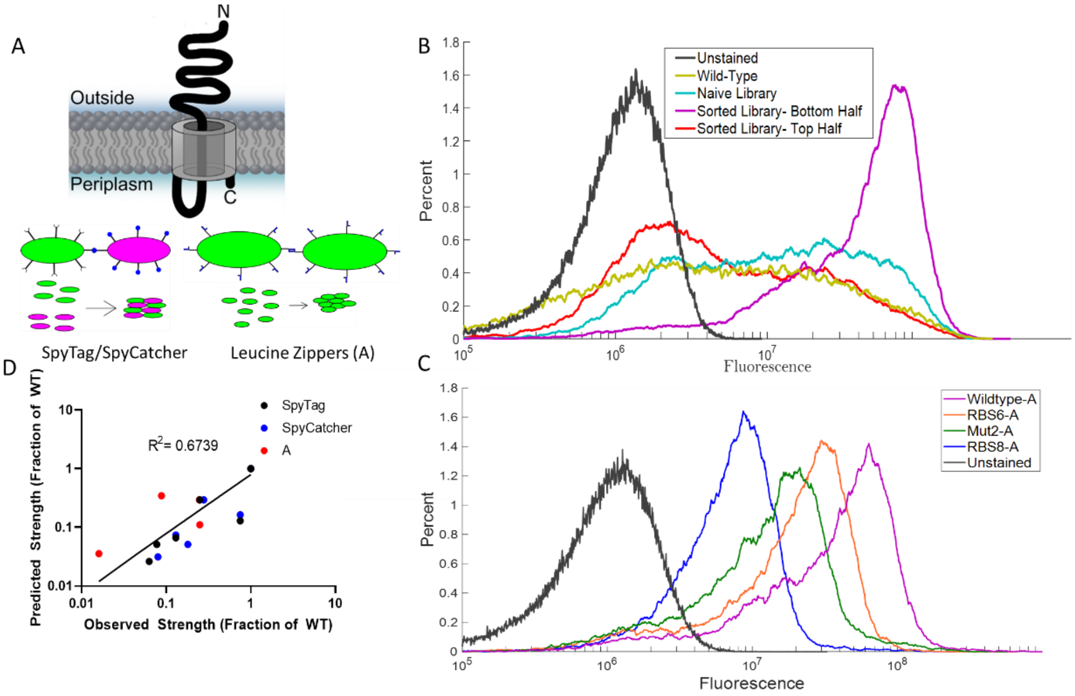


Figure 4.1. Programmable microbial assembly using self-associating surface-displayed proteins. (A) *E. coli* are engineered to express an autotransporter protein (AidA) on their surface, fused on its N-terminus to one of several target proteins in such a way as to display the target protein on the outer membrane

(exterior) of the bacterium. Multivalent expression of target proteins on the surface of bacteria programs interactions between bacterial cells. Multivalency leads to the formation of an extended network of bacterial interactions, leading to the formation of a bacterial aggregate. Either heterotypic (SpyTag/SpyCatcher) or homotypic (Leucine zippers) interactions may be used to drive aggregation. (B-C) Expression levels of target proteins can be engineered by mutation of the ribosome-binding site (RBS) upstream of the coding sequence of the gene.(B) Synthesis of a library of RBS's leads to a large range of expression levels, of which different parts can be enriched by FACS (C). Picking of single clones from the enriched libraries enables a large range of expression levels. (D) Expression levels correlate well with expected expression levels from RBS calculators.

Covalent Interactions between Cells and Kinetic Control of Aggregate Size

Cell-surface display of SpyTag and SpyCatcher results in the formation of irreversible covalent bonds between cells. We have previously found that these aggregates are resistant to disruption by soluble competitor peptide (Chapter 3 of this thesis), as opposed to dynamic, reversible systems which are disrupted by competitor binding,[30] and, in the case of polymeric particles displaying SpyTag and SpyCatcher, chemical denaturants.[13] In this limit of very strong cell-cell interactions, we aimed to measure the physical properties of the resulting flocs, using protein expression as a proxy for controlling the strength of cell-cell interaction.

We generated RBS libraries of SpyTag and SpyCatcher autotransporter fusions spanning approximately two orders of magnitude in expression levels (sequences and observed strengths in **Table S4.1**). These constructs were transformed into DH10b *E. coli* that also constitutively express fluorescent proteins (mWasabi or mCherry) to aid in fluorescence microscopy. In a typical experiment, cells were grown to mid-exponential phase (OD600 = 0.5), after which SpyTag and SpyCatcher expressing constructs were mixed, and expression of the surface-displayed proteins was induced. We generally observed visible flocculation and clearing of the culture within 30 minutes post-induction. Samples from the cultures were spotted onto glass cover slips and imaged by confocal fluorescence microscopy.

Using our library of designed RBSs to generate a smooth continuum of protein expression levels, we were able to precisely control the size of the aggregates at early time-points (**Figure 4.2A-C**). At 30 minutes, programmable protein expression level can be used to predictably control the floc volume over a range from 500 μm^3 to $3*10^5$ μm^3 (**Figure 4.2D**). The aggregates, though still irregular in shape, showed greater regularity than would be expected from the fractal pattern expected to emerge from unregulated inelastic collisions of bacteria with a growing aggregate. In that limit, we would expect our aggregates to be reminiscent of "diffusion limited aggregation," which maximizes surface area.[31] Instead, we hypothesize that some reorganizing activity for regular size and shape emerges at the system scale. This force (likely flow-induced shear stresses) penalizes strongly irregular structures, even in the covalent (very-strong force) limit.

Our ability to control the size of the aggregates diminished over time, with many distributions reaching the same final equilibrium configurations (**Figure 4.2E**). This result is consistent with classical DLVO theory.[32] In this case, assuming covalent bonds between cells are essentially irreversible once formed, as additional covalent bonds between cells are formed, aggregates will continue to grow with their kinetics controlled by the probability of successful colloidal collisions, which is proportional to the expression level.

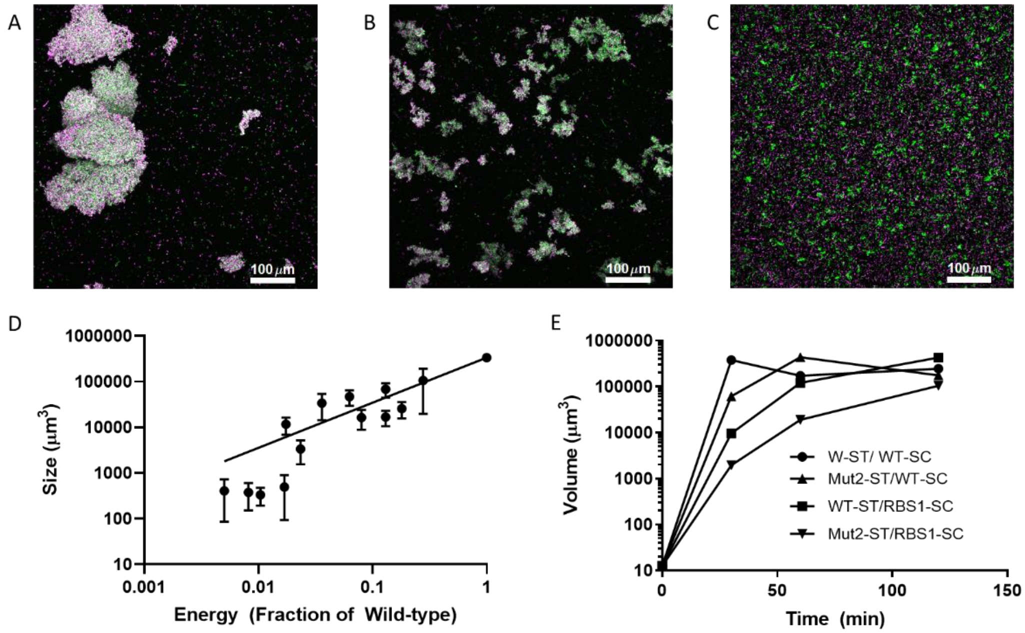


Figure 4.2. Aggregation triggered by SpyTag/SpyCatcher expression. (A-C) Surface expression of SpyTag(cells in green), SpyCatcher (cells in magenta) triggers aggregation. Aggregate size at short time points (30 minutes post-induction) is a function of expression levels (A) WT-SpyTag/WT-SpyCat, (B) Mut2-SpyTag/RBS2-SpyCat, (C) RBS2-SpyTag/RBS2-SpyCat. (D) Size of aggregates at short times are correlated with the expression levels of the adhesive proteins. Adhesive "Energy" is calculated as the product of the expression levels of each expression assuming a mass-action like expression for bridge formation. Note that the absolute values of the sizes are overestimates because fluorescence used for thresholding will often bleed into adjacent voxels, but the relative values are conserved. (E) Aggregates grow at different rates depending on their expression levels, but at long times, approach approximately the same aggregate size.

Physical Bridges Between Cells Enable Thermodynamic, Long-Time Control Over Aggregate Size

If the convergence to large aggregate sizes at long times is due to irreversible covalent bonds between cells decorated by SpyCatcher-SpyTag, we hypothesized that cell-cell aggregates formed through reversible, physical protein-protein interactions could enable longer-time control of aggregate size. To test this hypothesis, we chose to replace the SpyCatcher-SpyTag proteins with A coiled-coil domains, which form predictable and programmable non-covalent bonds. Ideally, by titrating the expression level, we could directly tune the intercellular potential, and for a sufficiently reversible system, obtain a dynamic equilibrium size distribution.

Similar to above, we generated an RBS library upstream of the A-autotransporter fusion protein, and observed expression levels spanning a large range of levels **(Table S4.2)**. These constructs were transformed into fluorescent DH10b *E. coli*, grown to mid-log phase, and aggregation was induced. Here, aggregation of the noncovalently bound cells was slower, and in most cases was not macroscopically observable.

We observed that corresponding to expectations, the size of the resulting aggregates is a strong function of the expression level (**Figure 4.3A-D**). In addition, we observe that in general the aggregates formed as a result of these reversible, physical interactions are substantially smaller than those formed by the irreversible, covalent bonds formed above (at full expression, aggregates are approximately $3*10^3$ μm^3 compared to 10^5 μm^3 in SpyTag/SpyCatcher). We speculate that this results from weaker inter-aggregate forces, which causes the shear stresses present in the culture to be sufficient to break apart aggregates larger than a critical size dependent on the expression level (and thus on the intercellular potential). Similar, or perhaps, even more so than SpyTag/SpyCatcher, A-aggregates exhibit regular shapes, further supporting the hypothesis that these aggregates are dynamic and may reorganize for hydrodynamic stability.

Additionally, if we increase the agitation speed to 350RPM, we observe that as expected, at high expression levels, the size of the aggregates is smaller at the higher agitation level (at full expression from $3*10^3$ μm^3 to $1*10^3$ μm^3) due to stronger shear stresses between aggregates and the surrounding fluid. This trend reverses as the expression level is decreased. We hypothesize that at these low expression levels, the dynamic equilibrium is between the rate of collisions and individual cell dissociation (instead of turbulent stresses being responsible for dissociation), and thus the higher agitation rate leads to more cell collisions. We note here that though that the *E. coli* used in this report are motile, the primary mechanism of collisions is convective due to vigorous mixing (corresponding to Pe_m>>1).

In line with our hypothesis, the aggregates formed using dynamic protein-protein interactions reach distinct long-time equilibria depending on their expression levels (**Figure 4.3E**). The aggregates generally reach their equilibrium distributions at approximately 90 minutes post-induction and maintain equilibrium size distributions for extended periods of time (at least until 22 hours post-induction). Since the bacteria continue to divide during this time (**Figure S4.1**), it appears that the equilibrium size distribution is a result of equilibrium between associative cell-cell forces and dispersive shear stresses, and not between collisions and sticking and dispersive forces. If instead the equilibrium was between the rate of collisions and dissociation, the equilibrium size would increase with particle number. Both of these potential equilibrium mechanisms have been observed in colloidal systems.[17]

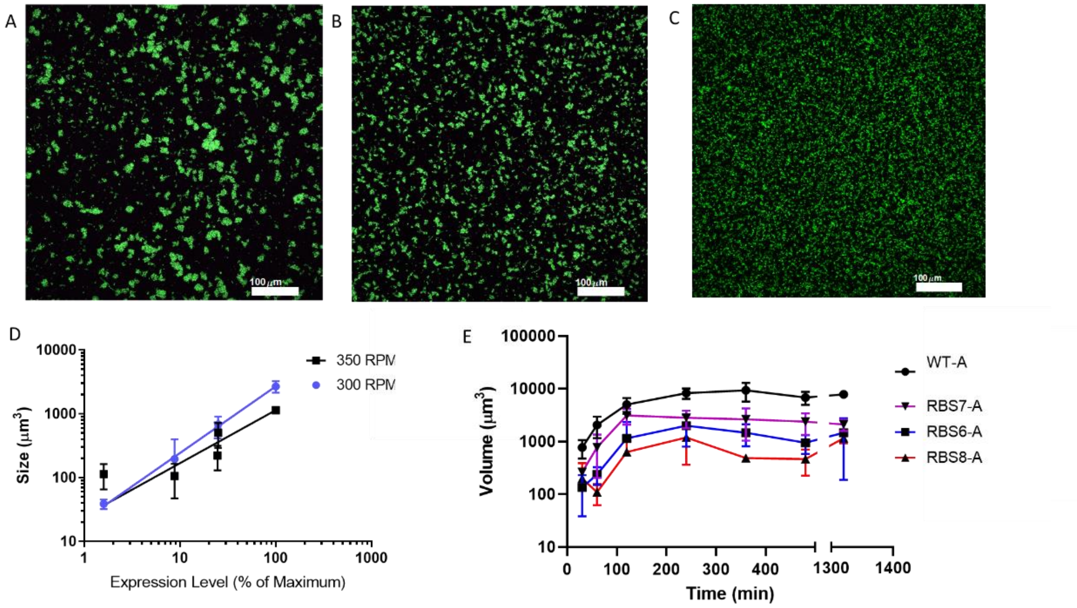


Figure 4.3: Aggregation triggered by A-coil expression. Surface expression of A triggers aggregation. (A-C) Titration of expression levels enables control of aggregate size. Micrographs are from samples taken 2 hours post-induction. (A) WT-A, (B) RBS7-A, (C) RBS8-A. (D) Aggregate size is well-correlated with expression levels (measured via flow cytometry). (E) By contrast with SpyTag/SpyCatcher aggregates, differences in aggregate sizes are maintained even at long times.

Association/Disassociation Dynamics may be Examined by Mixing Experiments

Above, we observed what appeared to be distinct dynamic properties of SpyTag/SpyCatcher and A-mediated aggregation processes. In order to distinguish these dynamics further, we carried out cross-color mixing experiments with the two different types of aggregates. To do this, we first separately aggregated bacteria expressing different fluorescent proteins, but the same associative proteins, such that we had single-color mWasabi- and mCherry-expressing SpyTag/SpyCatcher aggregates. After two hours (after the cells were expected to be fully aggregated), these single-color aggregates were mixed, and samples were then taken at defined intervals. The same experiment was also performed with A-expressing aggregates. We expect that if aggregates are dynamic, there will be significant reorganization of aggregates over the experimental time-scale of our experiment, and so we will observe mixing between the two colors, while if aggregates are more stable, we will continue to observe fully color-segregated aggregates.

As expected, the aggregates mediated by surface expression of A are much more dynamic than those mediated by SpyTag and SpyCatcher. We find that even after 9 hours post-mixing, the SpyTag and SpyCatcher aggregates are still very well-segregated (**Figure 4.4A-C,G**). By contrast, aggregates formed by bacteria expressing A are much more intermixed at moderate times (**Figure 4.4D-G**). We quantified this degree of mixing via a mixing coefficient that represents the fraction of adjacent voxels that are of different colors. This analysis confirms that the A aggregates become much better mixed with time than the SpyTag/SpyCatcher aggregates. In particular, the A aggregate mixing coefficient approaches our estimate of full mixing, 0.2 (See Materials and Methods).

The images from these experiments hint at the major mechanism for the disassociation dynamics in this system. Color mixing in both the SpyTag/SpyCatcher and A system appears to occur at large length scales at early time points, but decreases to smaller scales as time passes. This suggests that the major disruptive force on aggregates is shear stress from the turbulent flow field. If the disassociation was instead due to Brownian-type motion, we would expect to see single cells sticking onto aggregates of a different color. Instead we observe what appears to be fracture of large aggregates into smaller aggregates, which subsequently reattach to other aggregates. As time passes, this random mixing process continues, and the segregated domains decrease in size. This is especially clear for dynamic A-aggregates over the time-scale of this experiment, but also occurs in the SpyTag/SpyCatcher system.

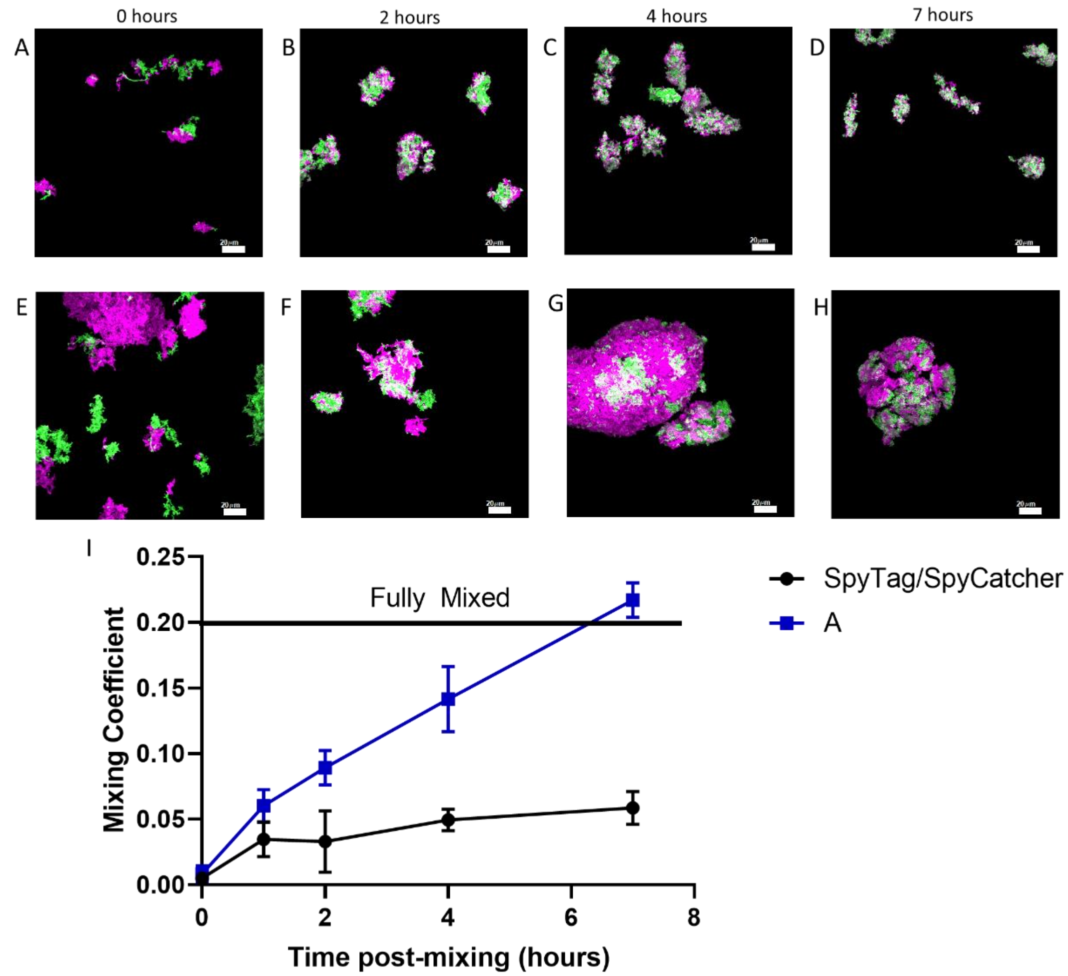


Figure 4.4. Dynamics of Aggregate Exchange. Mixing aggregates of different colors enables examination of aggregate dynamics. Aggregates appear to mix primarily by fracturing into smaller clusters followed by agglomeration of clusters. Mixing of colors is more pronounced in aggregates mediated by A coils (A-D) than in aggregates mediated by SpyTag/SpyCatcher (E-H). As time passes, more mixing takes place: (A and E) 0 hours post mixing (B and F), 2 hours post-mixing, (C and G), 4 hours post-mixing, (D and H), 7 hours post-mixing. (I) Mixing coefficient represents fraction of adjacent (6-valent) voxels that are of different colors. Leucine zipper coils have significantly better mixing, and appear to be fully mixed by the end of the experiment.

Destabilizing Mutants Exhibit Dynamic Dissociation

Having shown that we can control the equilibrium distribution of aggregate sizes driven by physical interactions by titrating the expression levels of the associating proteins, we wondered whether we could likewise change the aggregate distributions by modulating the strength of the pairing interaction, for example by modulating the effective affinity between coils. Towards this end, we made alanine mutations in the A coil at the L18 and L25 positions. These mutations are at the critical d-position of the canonical leucine zipper heptad,[33] and are thus expected to significantly change the properties of the coiled coil. Although no mutational scanning has previously been done with the A-coil, previous work by our group and others has found that these types of mutations in other coiled coils are generally destabilizing.[28,34]

When these mutant coils were surface displayed, we unexpectedly observed that the aggregates formed were noticeably larger at the two-hour time point **(Figure 4.5A-D**). Expression levels between the different coils did not appear to be significantly different (**Figure S4.2**). In order to determine whether these mutations were stabilizing the coiled-coil interactions, we expressed and purified the coils in soluble form, and used circular dichroism spectroscopy to determine the melt curves and secondary structure of the coils *in vitro* (**Figures S4.3** and **S4.4**). These experiments confirmed that the mutant coils were not thermodynamically stabilized in comparison to wild-type. To confirm the specificity of the protein-protein interactions driving aggregate assembly, we added soluble wild-type coil to previously formed aggregates, which disrupted both wild-type and mutant aggregates (**Figure S4.5**).

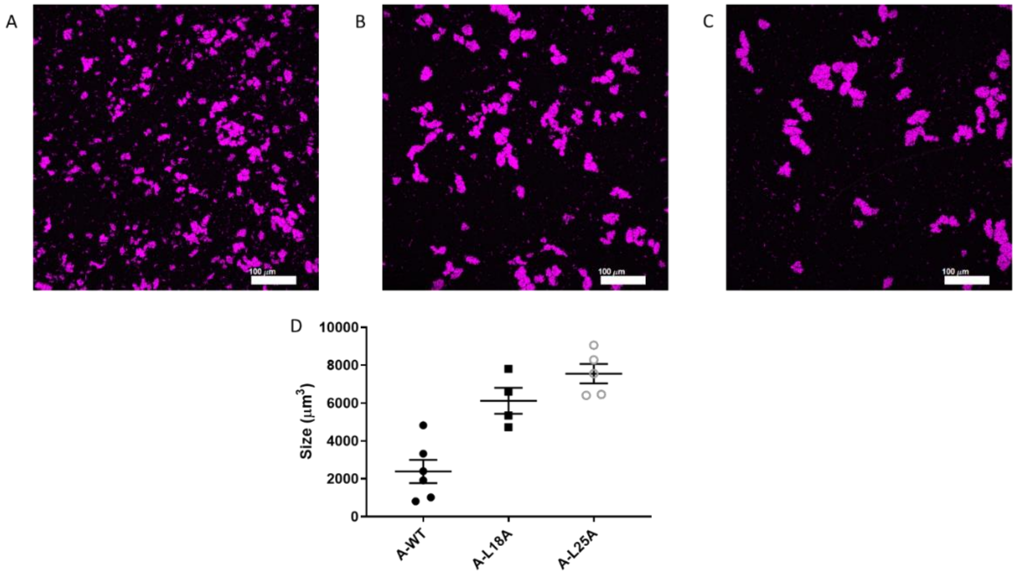


Figure 4.5. Aggregation of mutant A-coils. Mutation of L18 and L25 to alanine triggers the formation of larger aggregates, despite ostensibly being destabilizing mutations. (A-C) Micrographs of aggregates mediated by A-coils at 2 hours post-induction: (A) A-WT, (B) A-L18A, (C) A-L25A. (D) Quantification of aggregate sizes shows that mutant aggregates are significantly larger ($p<0.01$ from t-test) than the wild-type aggregates.

One possible hypothesis for the formation of larger aggregates is that the mutant coils are less rigid, and therefore could form more, albeit weaker bridges between cells. To probe this effect, we added a 150-residue elastin-like-polypeptide (ELP) linker between the coil and the autotransporter domain. These did not abolish the change in aggregate size; however, we noticed that in contrast to the wild-type coils or SpyTag/SpyCatcher, the aggregates mediated by the mutant coils began to dissociate around 4-6 hours post-induction. (**Figure 4.6**). By contrast, both wild-type A (even with the long linker) and SpyTag/SpyCatcher aggregates are stable and do not substantially dissociate even after 20 hours post-induction (see above). Indeed, upon further investigation, we noticed that even in the absence of the linker, the mutant coils will disaggregate.

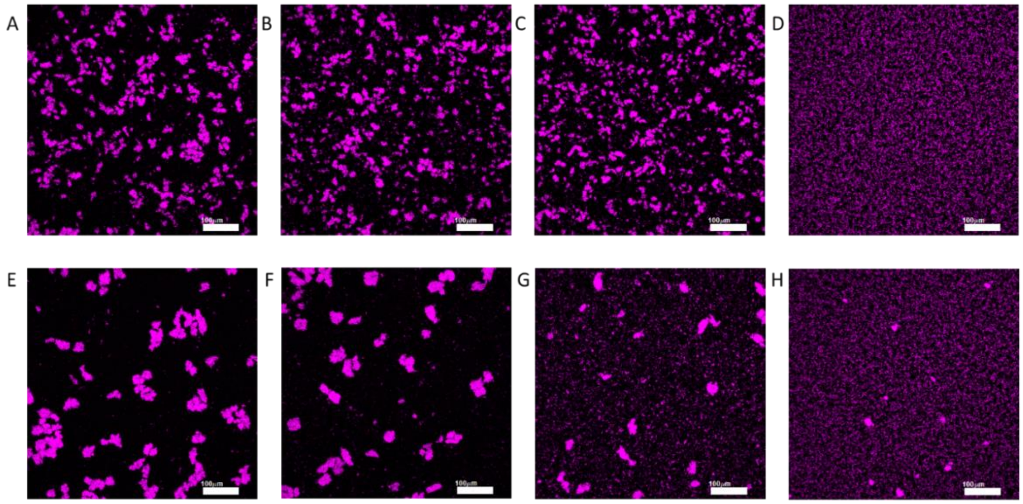


Figure 4.6 Mutant coil disassociation. Though mutant coils are initially more stable (larger), at longer time points they dissociate unlike the wild-type coils, which are stable at very long time points (See Figure 3E). (A-D) Aggregation time-course of A-L25A, (A) 2 hours post-induction, (B) 4 hours post-induction, (C) 6 hours post-induction, (D) 8 hours post-induction. When an extended linker is added to the mutant A-coil, dissociation still occurs, perhaps somewhat more quickly. (E-H) Aggregation time-course of A-L25A-ELP, (A) 2 hours post-induction, (B) 4 hours post-induction, (C) 6 hours post-induction, (D) 8 hours post-induction.

One possible explanation for the transient dissociation could be exogenous change in the properties of the media. LB media is known to become alkaline during bacterial growth,[35] and we have previously found that the A-coil is pH sensitive.[27] Our hypothesis for the onset of dissociation was then that the mutations made in the coil cause the coils to be more sensitive to high pH, triggering dissociation as the media becomes alkaline. To test this hypothesis, we found that adding base to adjust the pH of bacterial aggregates to ~8.5 causes rapid dissociation of mutant-coil aggregates, but not that of the wild-type (**Figure 4.7 A-C**). In long-time experiments of aggregation, dissociation was strongly linked to an increase in pH over 7.5 (**Figure 4.7D**). Finally, we found that using media buffered to maintain near-constant pH showed a sharp decrease in aggregate size as pH increases, but no significant changes with time (**Figure 4.7E** and **S4.6**). Though the dissociation pH in the two experiments is different (likely due to increased ionic/osmotic strength of the buffered media), both show essentially quantitative dissociation of aggregates over less than 0.5 pH units. These data demonstrate that the mutant coils act as an ultra-sensitive sensor; at neutral or somewhat acidic pH, mutant coils are able to (more) effectively form intercellular bridges, whereas as the media becomes even weakly alkaline, these bridges weaken and break. These experiments demonstrate that even simple molecular-level changes in the programmable elements may have large and complex results

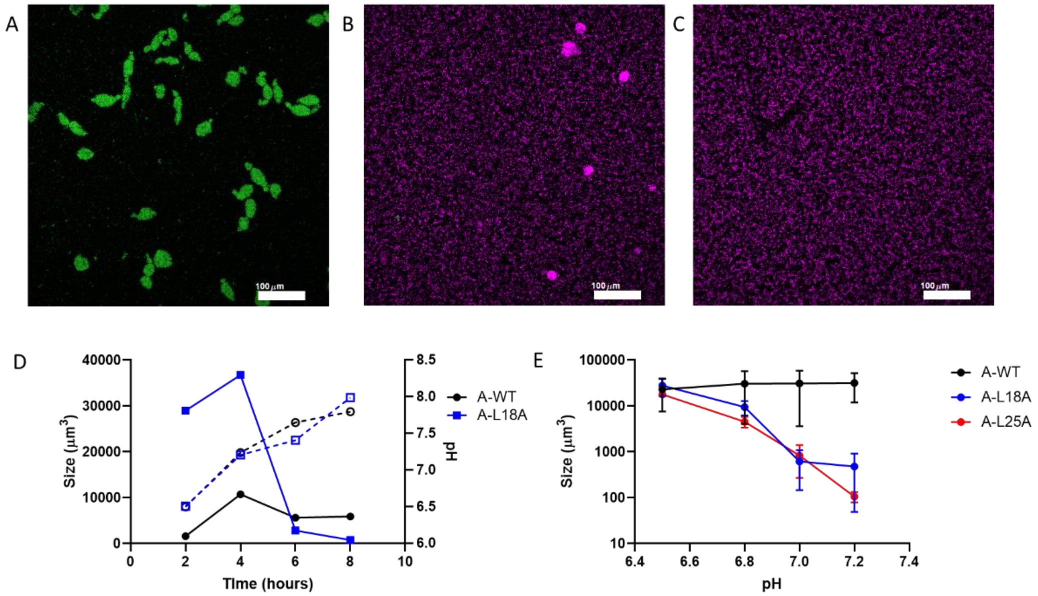


Figure 4.7. Mutant coil dissociation is a function of pH. Mutant coils dissociate when the media of the pH becomes alkaline (as happens during normal growth). A-C) When the pH of LB media is adjusted to ~8.5 by addition of base, aggregates mediated by mutant coils (B-C, A-L18A/A-L25A respectively), but not wild-type A-coil (A), dissociate. (D) During long growth, the dissociation of mutant aggregates occurs concurrently with the increase of pH above around 7. Although media containing wild-type aggregates also becomes alkaline (dotted lines represent pH), these aggregates do not dissociate (and maintain their equilibrium size). A single replicate is plotted due to inconsistency in the time course of pH changes, but replicates show similar qualitative results. (E) When the LB media is buffered by the addition of 100mM phosphate to different pH's, the sizes of the mutant aggregates but not the wild-type are strongly decreased by increases in pH. Sizes shown are after 4 hours aggregation, but trends hold over a broad range of times (see **S4.6**).

Conclusion

We have demonstrated our ability to control the properties of engineered microbial flocs using standard techniques in synthetic biology and protein engineering. We find that the properties of these flocs are related directly to the underlying intercellular interactions moderated by the engineered proteins that program the intercellular potentials. By changing the expression levels or identity of the proteins programming these interactions, we can generate a large range of floc behaviors including aggregation kinetics and equilibrium aggregate sizes. Even single amino acid changes may have outsized (and non-obvious) effects, where both aggregate sizes and dynamics may be affected. The results are likely fairly general, and may be adapted to other microbial systems with the sole requirement being the ability for cell surface display, allowing multi-species consortia to be created using this method. The work above also serves as a case study on the effects of inter-particle forces on the properties of strong colloidal flocs in a shear field which has previously been difficult to examine.

The ability to understand biological and abiotic flocs is important in a variety of contexts, including water treatment, environmental remediation, and metabolic engineering. In particular, in the latter example, the ability to create and control engineered microbial flocs may allow the creation of structured whole cell catalysts for use in biosynthetic transformations. In this case, precise control over aggregate properties (particularly size) is likely to be critical.

Acknowledgements

This work was supported by Defense Advanced Research Projects Agency Biological Robustness in Complex Settings Contract HR001-15-C-0093 and the Caltech Jacobs Institute for Molecular Engineering for Medicine. B.R.S. is supported by NIH Training Grant 1T32GM112592 and the Rosen Center for Bioengineering at Caltech. We thank Dr. Andres Collazo for assistance with confocal microscopy. Imaging was performed in the Biological Imaging Facility, with the support of the Caltech Beckman Institute and the Arnold and Mabel Beckman Foundation. We thank Professor Yaron Antebi for use of EasyFlow for flow cytometry analysis, and Adam Silverman for a critical reading of the manuscript.

Materials and Methods

General

Restriction enzymes, Q5 DNA polymerase, and Gibson assembly master mix were purchased from New England Biolabs (Beverly NJ). Nickel NTA was purchased from Qiagen (Hilden, Germany). A Zeiss LSM 800 confocal microscope was used for all imaging.

Bacterial Strains

SpyTag and SpyCatcher autotransporter fusions with the wild-type and Mut2 RBS sequences were reported previously (pBAD-ST, pBAD-SC, low-ST, and low-SC in Chapter 3).

pBAD-A was cloned as follows. The gene encoding the A-coil was ordered as a G-block (IDT, Coralville, IA), and cloned between BamHI and XhoI restriction sites in pQE-80-SpyTag-Autotransporter (removing the SpyTag, and replacing it with A). The full autotransporter gene was then cloned into an empty pBAD-33 vector using Gibson isothermal assembly and transformed into DH10b *E. coli*.

RBS libraries were made via Gibson assembly. Briefly, the autotransporter fusion was PCR amplified while also randomizing the wild-type Shine-Dalgarno sequence. These fragments were cloned back into pBAD-33 using Gibson assembly, and transformed via electroporation into DH10b *E. coli*. Generally, >20,000 clones were obtained (for a nominal library size of 16384). Different spacings (+/- 2bp) between the start codon and randomized sequence were attempted, but libraries exhibited similar activities, so only single libraries were sorted. These libraries were stored as glycerol stocks until sorting was performed.

SpyTag mutant RBS’s were cloned via Gibson assembly using an identical procedure as above for the libraries, except that rather than degenerate oligonucleotides, defined oligos were used for PCR instead.

Mutant A-coils were cloned via a modified Gibson assembly-based method. The entire pBAD-A plasmid was amplified and linearized via PCR, except that the nucleotides encoding the base pairs for the mutated residues were changed in the PCR overhang, and 25 bp overlaps were included in the overhangs on each end of the linearized plasmid. Gibson assembly was then used to re-circularize the plasmid.

Soluble A-coils were cloned by standard restriction enzyme cloning. The coils were PCR-amplified from the pBAD33-based plasmids, adding BamHI and HindIII restriction sites. The amplified genes were then digested with BamHI and HindIII and ligated into a similarly digested pQE-80L vector.

Aggregation experiments used DH10b *E. coli* constitutively expressing mWasabi or mCherry under the control of a leaky T5/Lac promoter as described previously (Chapter 3 of this thesis). Flow cytometry experiments used DH10b *E. coli* without the fluorescent cassette.

Fluorescence Activated Cell Sorting

For all flow cytometry experiments, DH10b E. coli containing the target plasmid were grown overnight in LB media to saturation, diluted 1:100 into fresh LB media, and incubated at 37°C with shaking. In mid-log phase (OD600≈0.5), expression was induced with 0.1% (wt/vol) L-arabinose, and cells were allowed to express the proteins for 90 minutes (this is sufficient for the population to reach steady state in terms of expression levels, data not shown). Cells were then harvested by centrifugation and blocked with 3% bovine serum albumin (BSA) in phosphate-buffered saline (PBS) for 30 minutes under agitation. Cells were then again harvested, and resuspended in staining solution (0.5% (v/v) fluorescently labeled Anti-His Antibody HIS.H8, 1% BSA in PBS) for 60 minutes. Cells were harvested, and then washed three times in PBS. After the last wash, cells were resuspended in PBS and strained through a 20 µm strainer to remove aggregates prior to flow cytometry.

Flow cytometry and sorting were performed with a Beckman Coulter MoFlo XDP instrument equipped with a blue laser (488 nm). Instrument alignment and calibration was performed according to manufacturer's directions. Gates for sorting and data analysis were manually set using control samples. If sorting was performed, cells were sorted into LB media without antibiotics, and allowed to grow for 1 hour, after which they were diluted into LB media with chloramphenicol (25 µg/mL) overnight, and/or plated onto LB-Cam plates. Analysis was performed with EasyFlow, a Matlab script for flow cytometry analysis written by Dr. Yaron Antebi.[36]

Soluble Protein Expression and Circular Dichroism

For soluble protein expression, BL21 *E. coli* with the desired plasmid were grown overnight to saturation, and then diluted 1:100 into 1L of Terrific Broth Media. Cells were grown to mid-log phase (OD=0.6), and induced with 1mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). After 4 hours, cells were harvested by centrifugation, and re-suspended in denaturing lysis buffer (8M Urea, 100 mM NaH_2PO_4, 10mM Tris-Cl, pH 8). Cells were then sonicated to complete lysis, and centrifuged to remove insoluble cellular debris. Proteins were then purified using NiNTA affinity chromatography according to manufacturer's directions. Protein purity was confirmed by SDS-PAGE, and purified proteins were dialyzed extensively against water and lyophilized for storage.

For Circular Dichroism spectroscopy (CD), samples were prepared at 1mg/mL in PBS, and then filtered to remove aggregates. Samples were pipetted into 1 mm width cuvettes, and CD was performed on a Model 410 Aviv Circular Dichroism Spectrophotometer. Thermal melt curves were performed at 1°C bandwidth and 1 minute equilibration times.

Supplemental Figures and Tables

Table S4.1. Amino Acid Sequences of all proteins used in this work

Protein Name	**Sequence**
SpyTag-Autotransporter	MKYLLPTAAAGLLLLAAQPAMAMRGSHHHHHHGSVDAHIVMVDAYKPTKLDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLETPTPGPDLNVDNDLRPEAGSYIANLAAANTMFTTRLHERLGNTYYTDMVTGEQKQTTMWMRHEGGHNKWRDGSGQLKTQSNRYVLQLGGDVAQWSQNGSDRWHVGVMAGYGNSDSKTISSRTGYRAKASVNGYSTGLYATWYADDESRNGAYLDSWAQYSWFDNTVKGDDLQSESYKSKGFTASLEAGYKHKLAEFNGSQGTRNEWYVQPQAQVTWMGVKADKHRESNGTLVHSNGDGNVQTRLGVKTWLKSHHKMDDGKSREFQPFVEVNWLHNSKDFSTSMDGVSVTQDGARNIAEIKTGVEGQLNANLNVWGNVGVQVADRGYNDTSAMVGIKWQF
SpyCatcher-Autotransporter	MKYLLPTAAAGLLLLAAQPAMAMRGSHHHHHHGSVDGAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIDLETPTPGPDLNVDNDLRPEAGSYIANLAAANTMFTTRLHERLGNTYYTDMVTGEQKQTTMWMRHEGGHNKWRDGSGQLKTQSNRYVLQLGGDVAQWSQNGSDRWHVGVMAGYGNSDSKTISSRTGYRAKASVNGYSTGLYATWYADDESRNGAYLDSWAQYSWFDNTVKGDDLQSESYKSKGFTASLEAGYKHKLAEFNGSQGTRNEWYVQPQAQVTWMGVKADKHRESNGTLVHSNGDGNVQTRLGVKTWLKSHHKMDDGKSREFQPFVEVNWLHNSKDFSTSMDGVSVTQDGARNIAEIKTGVEGQLNANLNVWGNVGVQVADRGYNDTSAMVGIKWQF
A-Autotransporter	MKYLLPTAAAGLLLLAAQPAMAMRGSHHHHHHGSVDGSGSGSGSGSGASGDLENEVAQLEREVRSLEDEAAELEQKVSRLKNEIEDLKAEGSGSGSGSGSGALETPTPGPDLNVDNDLRPEAGSYIANLAAANTMFTTRLHERLGNTYYTDMVTGEQKQTTMWMRHEGGHNKWRDGSGQLKTQSNRYVLQLGGDVAQWSQNGSDRWHVGVMAGYGNSDSKTISSRTGYRAKASVNGYSTGLYATWYADDESRNGAYLDSWAQYSWFDNTVKGDDLQSESYKSKGFTASLEAGYKHKLAEFNGSQGTRNEWYVQPQAQVTWMGVKADKHRES

	NGTLVHSNGDGNVQTRLGVKTWLKSHHKMDDGKSREFQPFVEVNWLHNSKDFSTSMDGVSVTQDGARNIAEIKTGVEGQLNANLNVWGNVGVQVADRGYNDTSAMVGIKWQF
A-L18A-Autotransporter	MKYLLPTAAAGLLLLAAQPAMAMRGSHHHHHHGSVDGSGSGSGSGSGASGDLENEVAQLEREVRS**A**EDEAAELEQKVSRLKNEIEDLKAEGSGSGSGSGSGALETPTPGPDLNVDNDLRPEAGSYIANLAAANTMFTTRLHERLGNTYYTDMVTGEQKQTTMWMRHEGGHNKWRDGSGQLKTQSNRYVLQLGGDVAQWSQNGSDRWHVGVMAGYGNSDSKTISSRTGYRAKASVNGYSTGLYATWYADDESRNGAYLDSWAQYSWFDNTVKGDDLQSESYKSKGFTASLEAGYKHKLAEFNGSQGTRNEWYVQPQAQVTWMGVKADKHRESNGTLVHSNGDGNVQTRLGVKTWLKSHHKMDDGKSREFQPFVEVNWLHNSKDFSTSMDGVSVTQDGARNIAEIKTGVEGQLNANLNVWGNVGVQVADRGYNDTSAMVGIKWQF
A-L25A-Autotransporter	MKYLLPTAAAGLLLLAAQPAMAMRGSHHHHHHGSVDGSGSGSGSGSGASGDLENEVAQLEREVRSLEDEAAE**A**EQKVSRLKNEIEDLKAEGSGSGSGSGSGALETPTPGPDLNVDNDLRPEAGSYIANLAAANTMFTTRLHERLGNTYYTDMVTGEQKQTTMWMRHEGGHNKWRDGSGQLKTQSNRYVLQLGGDVAQWSQNGSDRWHVGVMAGYGNSDSKTISSRTGYRAKASVNGYSTGLYATWYADDESRNGAYLDSWAQYSWFDNTVKGDDLQSESYKSKGFTASLEAGYKHKLAEFNGSQGTRNEWYVQPQAQVTWMGVKADKHRESNGTLVHSNGDGNVQTRLGVKTWLKSHHKMDDGKSREFQPFVEVNWLHNSKDFSTSMDGVSVTQDGARNIAEIKTGVEGQLNANLNVWGNVGVQVADRGYNDTSAMVGIKWQF
A-ELP-Autotransporter	MKYLLPTAAAGLLLLAAQPAMAMRGSHHHHHHGSVDGSGSGSGSGSGASGDLENEVAQLEREVRSLEDEAAELEQKVSRLKNEIEDLKAEGSGSGSGSGSGALDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLETPTPGPDLNVDNDLRPEAGSYIANLAAANTMFTTRLHERLGNTYYTDMVTGEQKQTTMWMRHEGGHNKWRDGSGQLKTQSNRYVLQLGGDVAQWSQNGSDRWHVGVMAGYGNSDSKTISSRTGYRAKASVNGYSTGLYATWYADDESRNGAYLDSWAQYSWFDNTVKGDDLQSESYKSKGFTASLEAGYKHKLAEFNGSQGTRNEWYVQPQAQVTWMGVKADKHRESNGTLVHSNGDGNVQTRLGVKTWLKSHHKMDDGKSREFQPFVEVNWLHNSKDFSTSMDGVSVTQDGAR

	NIAEIKTGVEGQLNANLNVWGNVGVQVADRGYNDTSAMVGIKWQF
A-L18A-ELP-Autotransporter	MKYLLPTAAAGLLLLAAQPAMAMRGSHHHHHHGSVDGSGSGSGSGSGASGDLENEVAQLEREVRSAEDEAAELEQKVSRLKNEIEDLKAEGSGSGSGSGSGALDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLETPTPGPDLNVDNDLRPEAGSYIANLAAANTMFTTRLHERLGNTYYTDMVTGEQKQTTMWMRHEGGHNKWRDGSGQLKTQSNRYVLQLGGDVAQWSQNGSDRWHVGVMAGYGNSDSKTISSRTGYRAKASVNGYSTGLYATWYADDESRNGAYLDSWAQYSWFDNTVKGDDLQSESYKSKGFTASLEAGYKHKLAEFNGSQGTRNEWYVQPQAQVTWMGVKADKHRESNGTLVHSNGDGNVQTRLGVKTWLKSHHKMDDGKSREFQPFVEVNWLHNSKDFSTSMDGVSVTQDGARNIAEIKTGVEGQLNANLNVWGNVGVQVADRGYNDTSAMVGIKWQF
A-L25A-ELP-Autotransporter	MKYLLPTAAAGLLLLAAQPAMAMRGSHHHHHHGSVDGSGSGSGSGSGASGDLENEVAQLEREVRSLEDEAAEAEQKVSRLKNEIEDLKAEGSGSGSGSGSGALDVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGVPGAGVPGAGVPGEGVPGAGVPGAGLETPTPGPDLNVDNDLRPEAGSYIANLAAANTMFTTRLHERLGNTYYTDMVTGEQKQTTMWMRHEGGHNKWRDGSGQLKTQSNRYVLQLGGDVAQWSQNGSDRWHVGVMAGYGNSDSKTISSRTGYRAKASVNGYSTGLYATWYADDESRNGAYLDSWAQYSWFDNTVKGDDLQSESYKSKGFTASLEAGYKHKLAEFNGSQGTRNEWYVQPQAQVTWMGVKADKHRESNGTLVHSNGDGNVQTRLGVKTWLKSHHKMDDGKSREFQPFVEVNWLHNSKDFSTSMDGVSVTQDGARNIAEIKTGVEGQLNANLNVWGNVGVQVADRGYNDTSAMVGIKWQF
A (Soluble)	MRGSHHHHHHGSVDGSGSGSGSGSGASGDLENEVAQLEREVRSLEDEAAELEQKVSRLKNEIEDLKAEGSGSGSGSGSGAKLN
A-L18A (Soluble)	MRGSHHHHHHGSVDGSGSGSGSGSGASGDLENEVAQLEREVRSAEDEAAELEQKVSRLKNEIEDLKAEGSGSGSGSGSGAKLN
A-L25A (Soluble)	MRGSHHHHHHGSVDGSGSGSGSGSGASGDLENEVAQLEREVRSLEDEAAEAEQKVSRLKNEIEDLKAEGSGSGSGSGSGAKLN

Table S4.2: Sequences and Strengths of all RBS's used in this work

RBS-Name	RBS Sequence (Putative Shine-Dalgarno sequence in red)	Observed Strength (% of WT)
WT-SpyTag	GAGGAGAAATTAACTATG	100
Mut2-SpyTag	GAGCGAGAAATTAACTATG	25
RBS2-SpyTag	CCACTGCTTAACTATG	6
RBS3-SpyTag	TAATCAGTTAACTATG	13
RBS4-SpyTag	TGCCTTGTTAACTATG	8
RBS5-SpyTag	ACGTGGTAATTAACTATG	75
WT-SpyCat	GAGGAGAAATTAACTATG	100
Mut2-SpyCat	GAGCGAGAAATTAACTATG	28
RBS2-SpyCat	CCACTGCTTAACTATG	8
RBS3-SpyCat	TAATCAGTTAACTATG	13
RBS4-SpyCat	TGCCTTGTTAACTATG	18
RBS5-SpyCat	ACGTGGTAATTAACTATG	75
WT-A	GAGGAGAAATTAACTATG	100
Mut2-A	GAGCGAGAAATTAACTATG	25
RBS6-A	GAGCGAGAAATTAACTATG	9
RBS7-A	GGGAAGCAATTAACTATG	25
RBS8-A	GTTTTAGAATTAACTATG	2

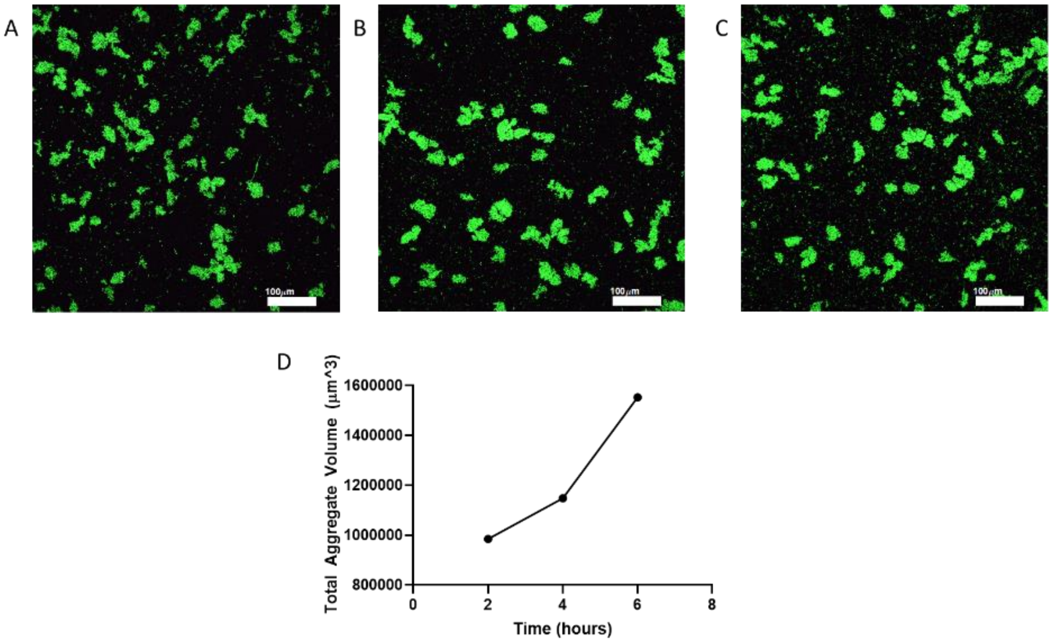


Figure S4.1: Growth during aggregation. Even after aggregation, bacterial growth may still occur. The average equilibrium size does not change (Figure 3E in main text), but the number of aggregates (and thus the total aggregate volume in the sample) increases. (A) 2 hours post-induction, (B) 4 hours post-induction, (C) 6 hours-post induction. D) Total aggregate volume across a 1.5 μL sample

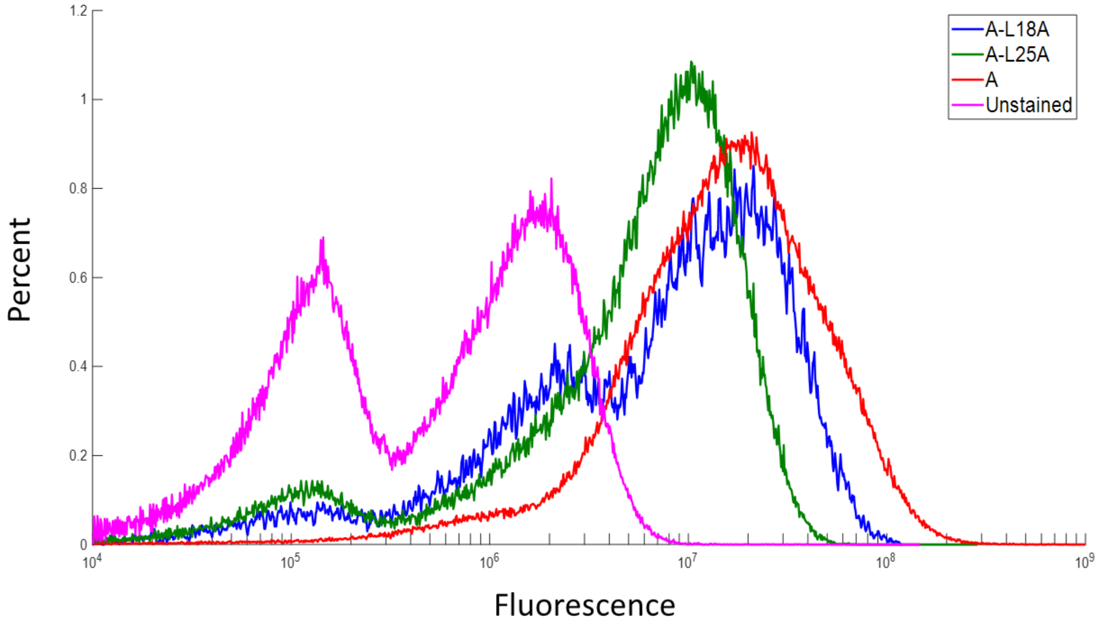


Figure S4.2: Flow cytometry of A coil mutants. Flow cytometry demonstrates that coil mutants do not express substantially differently than the wild-type coil.

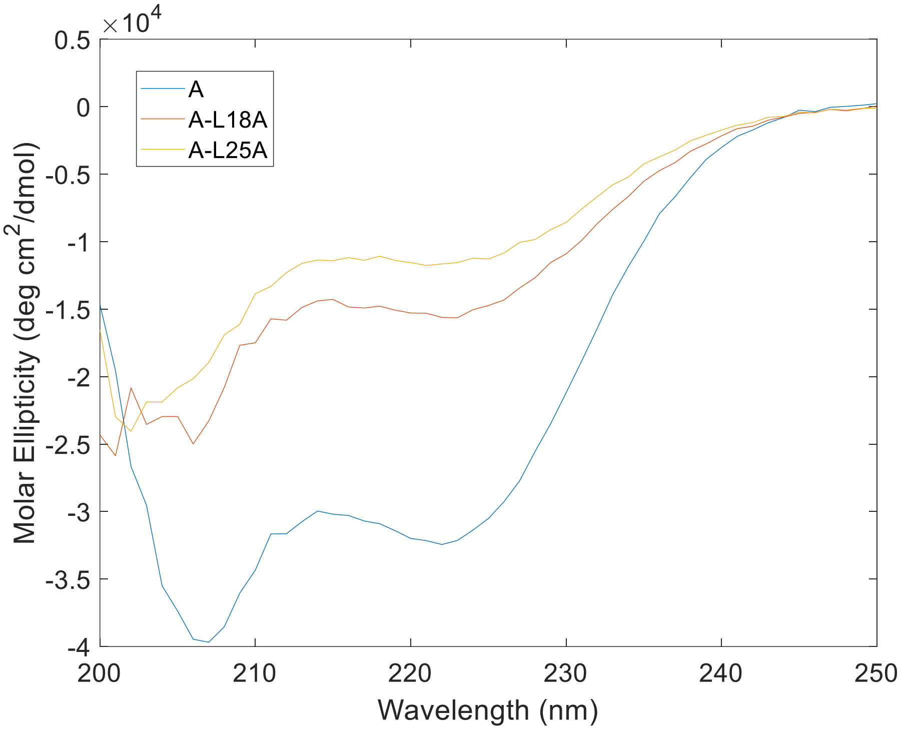


Figure S4.3: Circular dichromism spectroscopy wavelength sweep. Pronounced double peak at 222nm and 208 nm is characteristic of alpha helical structure. Wild-type exhibits substantially more helical character than either L18A or L25A mutants.

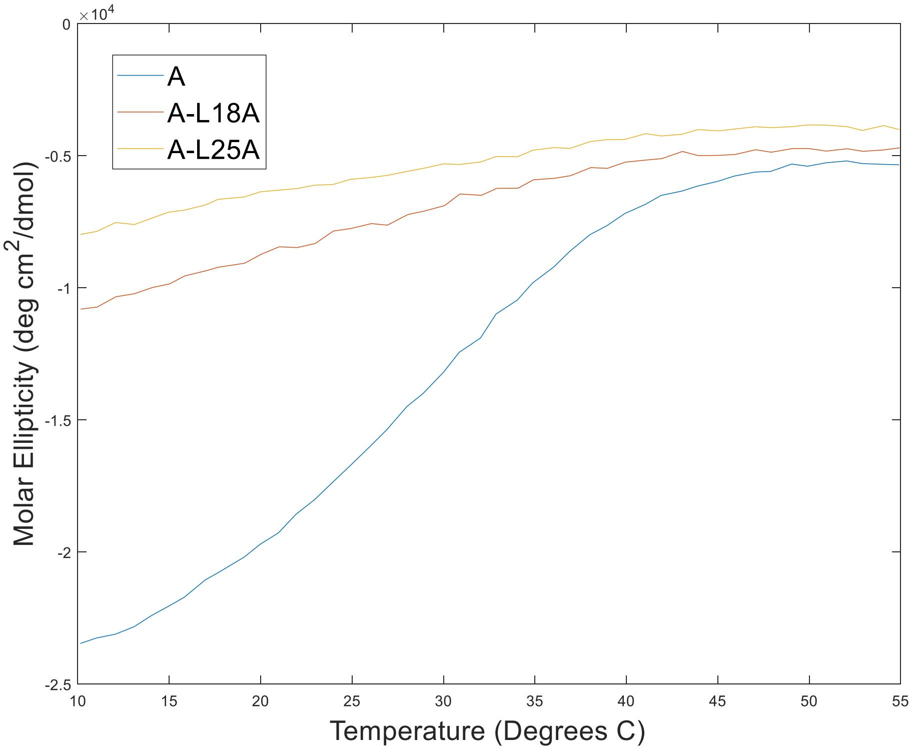


Figure S4.4: Circular dichromism temperature sweep at 222 nm. Wild-type A coil exhibits more alpha helical character at all temperatures under these conditions.

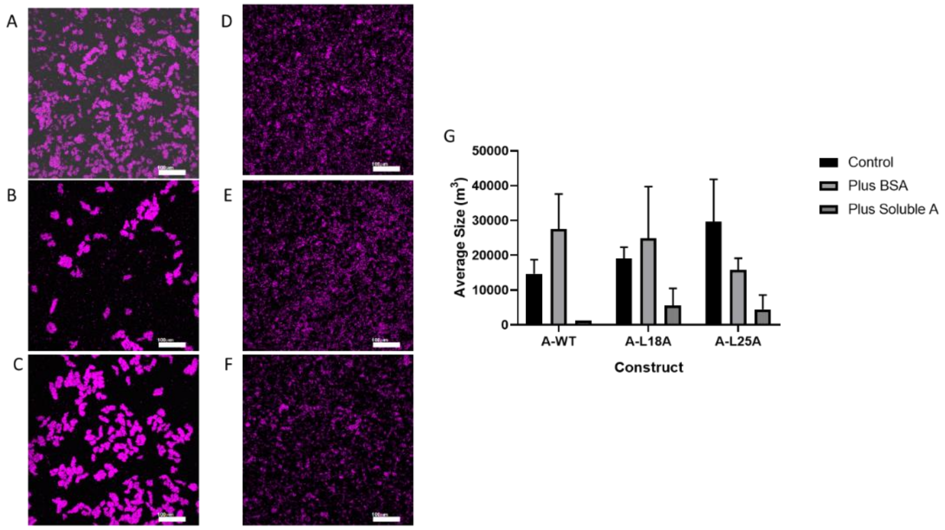


Figure S4.5: Soluble A-coil Protein triggers dissociation of Leucine Zipper Aggregates. After aggregation, soluble A protein and bovine serum albumin (BSA) were added to A-wt, A-L18A, and A-L25A aggregates. In all cases, soluble A protein, but not BSA caused aggregates to dissociate, demonstrating that aggregates are held together by specific biomolecular interactions. (A-C) Aggregates without soluble A-coil. (A) A-wt, (B) A-L18A, (C) A-L25A. (D-F) Aggregates disrupted by 1mg/mL soluble A coil. (D) A-wt, (E) A-L18A, (F) A-L25A. (G) Sizes of aggregates with added soluble A-coil or BSA.

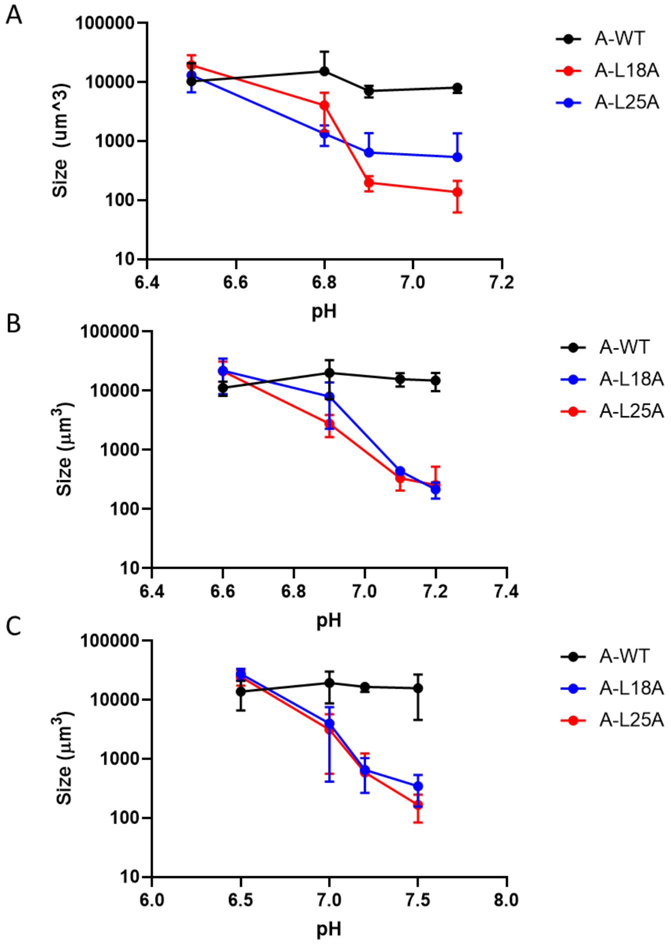


Figure S4.6: Buffered LB maintains pH values and aggregate sizes over time. LB that is buffered to different pH values maintains its pH over greater than 8 hours of bacterial culture. Aggregate sizes are broadly similar at 2 hours (A), 6 hours (B), and 8 hours (C) post-induction. In all cases, wild-type aggregates are essentially constant in size regardless of pH, while mutant aggregates rapidly decrease in size at even slightly alkaline pH's.

References

1 N. S. Jakubovics, S. R. Gill, S. E. Iobst, M. M. Vickerman and P. E. Kolenbrander, Regulation of gene expression in a mixed-genus community: Stabilized arginine biosynthesis in Streptococcus gordonii by coaggregation with Actinomyces naeslundii., *J. Bacteriol.*, 2008, **190**, 3646–57.

2 J. M. Monier and S. E. Lindow, Frequency, size, and localization of bacterial aggregates on bean leaf surfaces, *Appl. Environ. Microbiol.*, 2004, **70**, 346–355.

3 Y. Xia, X. Wen, B. Zhang and Y. Yang, Diversity and assembly patterns of activated sludge microbial communities: A review, *Biotechnol. Adv.*, 2018, **36**, 1038–1047.

4 T. Mah and G. O'Toole, Mechanisms of biofilm resistance to antimicrobial agents, *Trends Microbiol.*, 2001, **9**, 34–39.

5 S. Elias and E. Banin, Multi-species biofilms: Living with friendly neighbors., *FEMS Microbiol. Rev.*, 2012, **36**, 990–1004.

6 B. K. Hammer and B. L. Bassler, Quorum sensing controls biofilm formation in Vibrio cholerae, *Mol. Microbiol.*, 2003, **50**, 101–104.

7 A. Prindle, J. Liu, M. Asally, S. Ly, J. Garcia-Ojalvo and G. M. Süel, Ion channels enable electrical communication in bacterial communities, *Nature*, 2015, **527**, 59–63.

8 G. Eggset, E. Stenberg and J. Kjosbakken, Flocculation of a Methylomonas sp.: Possible involvement of a surface protein, *J. Gen. Microbiol.*, 1983, **129**, 3611–3617.

9 S. Asakura and F. Oosawa, On interaction between two bodies immersed in a solution of macromolecules, *J. Chem. Phys.*, 1954, **22**, 1255.

10 D. Hanaor, M. Michelazzi, C. Leonelli and C. C. Sorrell, The effects of carboxylic acids on the aqueous dispersion and electrophoretic deposition of ZrO2, *J. Eur. Ceram. Soc.*, 2012, **32**, 235–244.

11 M. E. Leunissen, C. G. Christova, A.-P. Hynninen, C. P. Royall, A. I. Campbell, A. Imhof, M. Dijkstra, R. van Roij and A. van Blaaderen, Ionic colloidal crystals of oppositely charged particles, *Nature*, 2005, **437**, 235–40.

12 C. Mirkin, R. Letsinger, R. Mucic and J. Storhoff, A DNA-based method for rationally assembling nanoparticles into macroscopic materials, *Nature*, 1996, **382,** 607-609

13 M. Obana, B. R. Silverman and D. A. Tirrell, Protein-mediated colloidal assembly*, J. Am. Chem. Soc.*, 2017, **139**, 14251–14256.

14 F. Li, D. P. Josephson and A. Stein, Colloidal assembly: The road from particles to colloidal molecules and crystals, *Angew. Chemie*, 2011, **50**, 360–88.

15 T. Serra, J. Colomer and X. Casamitjana, Aggregation and breakup of particles in a shear flow, *J. Colloid Interface Sci.*, 1997, **187**, 466–473.

16 C. A. Biggs and P. A. Lant, Activated sludge flocculation: On-line determination of floc size and the effect of shear, *Water Sci. Technol.*, 2000, **34**, 2542–2550.

17 J. A. Long, D. W. J. Osmond and B. Vincent, The equilibrium aspects of weak flocculation, *J. Colloid Interface Sci.*, 1973, **42**, 545–553.

18 T. G. Milligan and P. S. Hill, A laboratory assessment of the relative importance of turbulence, particle composition, and concentration in limiting maximal floc size and settling behaviour, *J. Sea Res.*, 1998, **39**, 227–241.

19 Y. Wang, B. Y. Gao, X. M. Xu, W. Y. Xu and G. Y. Xu, Characterization of floc size, strength and structure in various aluminum coagulants treatment, *J. Colloid Interface Sci.*, 2009, **332**, 354–359.

20 E. Veiga, V. De Lorenzo and L. A. Ferna, Autotransporters as scaffolds for novel bacterial adhesins: Surface properties of Escherichia coli cells displaying Jun / Fos dimerization domains, *J. Bacteriol.*, 2003, **185**, 5585–5590.

21 W. Zhang, F. Sun, D. A. Tirrell and F. H. Arnold, Controlling macromolecular

topology with genetically encoded SpyTag–SpyCatcher chemistry, *J. Am. Chem. Soc.*, 2013, **135**, 13988–13997.

22 F. Sun, W.-B. Zhang, A. Mahdavi, F. H. Arnold and D. A. Tirrell, Synthesis of bioactive protein hydrogels by genetically encoded SpyTag-SpyCatcher chemistry, *Proc. Natl. Acad. Sci.*, 2014, **111**, 11269–74.

23 C. Schoene, J. O. Fierer, S. P. Bennett and P. Mark, SpyTag / SpyCatcher cyclization confers resilience to boiling on a mesophilic enzyme, *Angew. Chemie*, 2015, **53**, 6101–6104.

24 C. N. Bedbrook, M. Kato, S. Ravindra Kumar, A. Lakshmanan, R. D. Nath, F. Sun, P. W. Sternberg, F. H. Arnold and V. Gradinaru, Genetically encoded Spy peptide fusion system to detect plasma membrane-localized proteins in vivo, *Chem. Biol.*, 2015, **22**, 1108–21.

25 W. Petka, J. Harden, K. McGrath, D. Wirtz and D. Tirrell, Reversible hydrogels from self-assembling artificial proteins, *Science*, 1998, **281**, 389–393.

26 W. Shen, K. Zhang, J. A. Kornfield and D. A. Tirrell, Tuning the erosion rate of artificial protein hydrogels through control of network topology, *Nat. Mater.*, 2006, **5**, 153–158.

27 W. Shen, Structure, dynamics, and properties of artificial protein hydrogels assembled through coiled-coil domains, Doctoral Thesis, California Institute of Technology, 2005.

28 L. J. Dooling and D. A. Tirrell, Engineering the dynamic properties of protein networks through sequence variation, *ACS Cent. Sci.*, 2016, **2**, 812–819.

29 H. M. Salis, E. A. Mirsky and C. A. Voigt, Automated design of synthetic ribosome binding sites to control protein expression, *Nat. Biotechnol.*, 2009, **27**, 946–950.

30 W. B. Rogers and V. N. Manoharan, Programming colloidal phase transitions with DNA strand displacement, *Science*, 2015, **347**, 639–642.

31 R. Klein and P. Meakin, Universality in colloid aggregation, *Nature*, 1989,

339, 360-362

32 R. J. Pugh and J. A. Kitchener, Theory of selective coagulation in mixed colloidal suspensions, *J. Colloid Interface Sci.*, 1971, **35**, 656–664.

33 D. Woolfson, The design of coiled-coil structures and assemblies, *Adv. Protein Chem.*, 2005, **70**, 79–112.

34 S. K. Gunasekar, M. Asnani, C. Limbad, J. S. Haghpanah, W. Hom, H. Barra, S. Nanda, M. Lu and J. K. Montclare, N-terminal aliphatic residues dictate the structure, stability, assembly, and small molecule binding of the coiled-coil region of cartilage oligomeric matrix protein, *Biochemistry*, 2009, **48**, 8559–67.

35 G. Sezonov and R. D. Ari, Escherichia coli physiology in Luria-Bertani broth, 2007, **189**, 8746–8749.

36 Y. E. Antebi, J. M. Linton, H. Klumpe, C. Su, R. Mccardell, M. B. Elowitz, Combinatorial signal perception in the BMP pathway, *Cell*, 2017, **170**, 1184-1185.e24.

Chapter 5

GENETICALLY PROGRAMMABLE MICROBIAL ASSEMBLY ENHANCES MICROBIAL PRODUCTION OF SPECIALTY CHEMICALS

Abstract

Genetic engineering of microbes has become a commonly used platform technology for the green production of fine and specialty chemicals. In particular, the over-expression of heterologous enzymes from other species enables the implementation of functional heterologous biochemical pathways for the production of desired chemicals. However, in microbes, the expression of many heterologous enzymes in a single strain may have significant genetic load on the organism, or different steps of the biochemical pathway may be incompatible and require substantial re-engineering. Many groups have instead opted to use co-culture systems, where steps of the biochemical pathway are separated into different strains of microorganisms in a division of labor; however, this requires the intermediate species to diffuse into the (dilute) bulk culture, decreasing flux through the pathway. In this work, we demonstrate that by aggregating several strains of bacteria into closely packed flocs, we can enhance fluxes through a model biosynthetic pathway by increasing the local concentration of intermediate species inside the aggregates. Further, we demonstrate that the size of aggregates is an important factor in the effectiveness of these aggregates and demonstrate that our ability to control the size of the resulting aggregates enables optimization of pathway yields. We believe that this approach is broadly generalizable to many classes of biosynthetic pathways and will enable a new tool for metabolic engineering to make specialty chemicals.

Introduction

Metabolic engineering is increasingly used to implement longer and more complex biochemical pathways into microorganisms for the production of valuable fine and specialty chemicals. Advancements in understanding of the metabolism of these organisms,[1,2] as well as significant innovations in protein engineering,[3] genome engineering,[4] and synthetic biology,[5] have led to a growing capacity for the rational engineering of microorganisms to make a broad variety of economically important compounds. However, as biochemical pathways become longer, the efficiency of these pathways may decrease for several reasons. First, as more heterologous genes are added to an organism, metabolic load on the organism's growth will also increase, due to exertion of energy in DNA replication, protein expression, and product synthesis.[6] In addition to the difficulty of simply expressing many heterologous enzymes (especially in bacteria), in some cases enzymes may require different intracellular environments[7] or compete for similar metabolites.[8] Some groups have attempted to ameliorate these problems by separating enzymes into different intracellular compartments,[7] or by carefully balancing enzyme levels with metabolite concentrations through computational modeling or combinatorial experimentation.[9,10]

Other groups have attempted to solve this problem utilizing a division of labor approach in which the enzymes are split into multiple strains of engineered microorganisms, each of which performs only a part of the entire pathway. Co-culture of these different strains then enables the reconstitution of the entire pathway in the bioreactor.[8,11]This strategy helps limit the metabolic burden on each strain, as they are now responsible for only a part of the final pathway. In addition, this strategy enables the independent engineering of each strain to optimize for its function, allowing much more flexibility in strain engineering. However, a major disadvantage in splitting enzymes between different microbial strains is the requirement for mass transfer of pathway intermediates between cells. Because a single cell is no longer able to perform the entire transformation, intermediates must be transported from the upstream strain into the bulk media, and then into the downstream strain. Many pathway intermediates, particularly CoA-conjugated and phosphorylated species are not readily membrane permeant, and are thus not suitable for this kind of engineering.[11] Suitable choice of the separation of enzymes into the different strains helps obviate this problem. However, even if intermediates are cell-permeant, pathway intermediates are still heavily diluted by secretion into the bulk media, which decreases overall flux through the biosynthetic pathway.

In nature, many microbes are found in highly spatially-organized multi-species communities of biofilms or aggregates.[12,13] Living in biofilms enables bacteria to resist environmental insults,[14] sense their environments,[15,16] and most importantly for our purposes, share metabolites through a process called metabolic cross-feeding.[12,17] In this process, one bacterial strain will synthesize and secrete a metabolite necessary for another microbe's growth. This enables a division of labor, where microbes are no longer required to contain all of the biosynthetic enzymes required for their growth, and significantly enhances the versatility of bacterial biofilms. Importantly, because biofilms are highly spatially packed with bacteria and extracellular matrix, the diffusion of metabolites is restricted, so secreted metabolites are not strongly diluted by bulk diffusion, and they can accumulate in the biofilms.[18]

Taking this as our inspiration, we hypothesized that by genetically engineering bacteria to form aggregates, we could enable them to shuttle pathway intermediates between themselves without diffusion into the bulk solution, enabling higher concentrations of intermediates, and correspondingly higher flux through the pathway. This strategy enables the advantages of modularity and limited metabolic load of co-culture systems, while restricting the slow rate of mass transfer into the bulk solution.

Our group has previously developed methods to genetically engineer *Escherichia coli* to inducibly aggregate into clusters ranging from ~5µm-100µm in diameter by the surface display of associative proteins on their surfaces (described in Chapters 3 and 4 of this thesis). By the choice of associative protein and its expression level, we may modulate the equilibrium size, aggregation kinetics, and dissociation properties of the underlying aggregate. Further, we have demonstrated that the aggregation of bacteria into aggregates may have physiological implications on the cells making up the aggregate, for example the activation of a quorum sensing circuit (Chapter 3). In this work, we demonstrate that the aggregates formed by these mechanisms may be used to enhance the yield of multi-step biosynthetic transformations compared to standard co-culture. Further, we show that the yield of these biosynthetic transformations is a function of the aggregate size, and present a mathematical model which identifies the relevant parameters for the control and optimization of these biocatalysts.

Results and Discussion

Mathematical Modeling of Biosynthetic Pathways

In order to understand the key parameters for the optimization of biosynthetic pathways in bacterial aggregates, we formulated a simple mathematical model of the process. In the model, we consider reactions to take place within aggregates and diffuse through aggregates with a characteristic diffusion coefficient (which may be distinct from the Brownian molecular diffusion coefficient). We assume that mass transport to the surface of the aggregate is not limiting, such that concentrations at the surface are equal to bulk concentrations (reasonable for closely packed aggregates and vigorous mixing). For a simple pathway where a substrate A is added to the culture and is transformed in two (first-order) steps to C, we derive the following equations (full derivation in Supplemental Discussion**, Figure 5.1A**)

$$\frac{1}{\phi^2}\nabla^2 C_A - C_A = 0 \tag{5.1}$$

$$\frac{1}{\phi^2}\nabla^2 C_B + C_A - \gamma_B C_B = 0 \tag{5.2}$$

$$\frac{1}{\phi^2}\nabla^2 C_C + \gamma_B C_B = 0 \tag{5.3}$$

$$\frac{dC_A^B}{dt} = -\frac{\int C_A dV}{\int dV} \tag{5.4}$$

$$\frac{dC_B^B}{dt} = -\frac{\int \gamma_B C_B - C_A dV}{\int dV} \tag{5.5}$$

$$\frac{dC_C^B}{dt} = \frac{\int \gamma_B C_B dV}{\int dV} \tag{5.6}$$

All variables in equations 5.1-6 are dimensionless (transformations shown in Supplemental Discussion). C_A,C_B, and C_C represent the concentrations of molecules A, B, and C, respectively, φ is the Thiele modulus (which represents the ratio of reaction rate to diffusion rate in a catalyst particle, and can be seen as a measure of the size of the aggregates),[19] γ_B represents the ratio of the rate constants for the B→C transformation, compared to the A→B transformation, ∇^2 is the Laplacian operator, and C_A^B, C_B^B, and C_C^B are the bulk concentrations of the respective species. Integrals are taken over the entire volume of an aggregate. Equations 5.1-3 represent the internal reaction-dynamics of the aggregates, while Equations 5.4-6 represent the concentration changes in the bulk fluid due to reaction in the aggregates. With suitable boundary conditions (shown in Supplemental Discussion), these equations may be numerically integrated with different parameter values for the Thiele modulus and relative reaction rate γ_B (**Figure 5.1B-D**). In this model, we do not consider increased metabolic burden or nutrient limitations due to the aggregation process, which may be important in many systems.

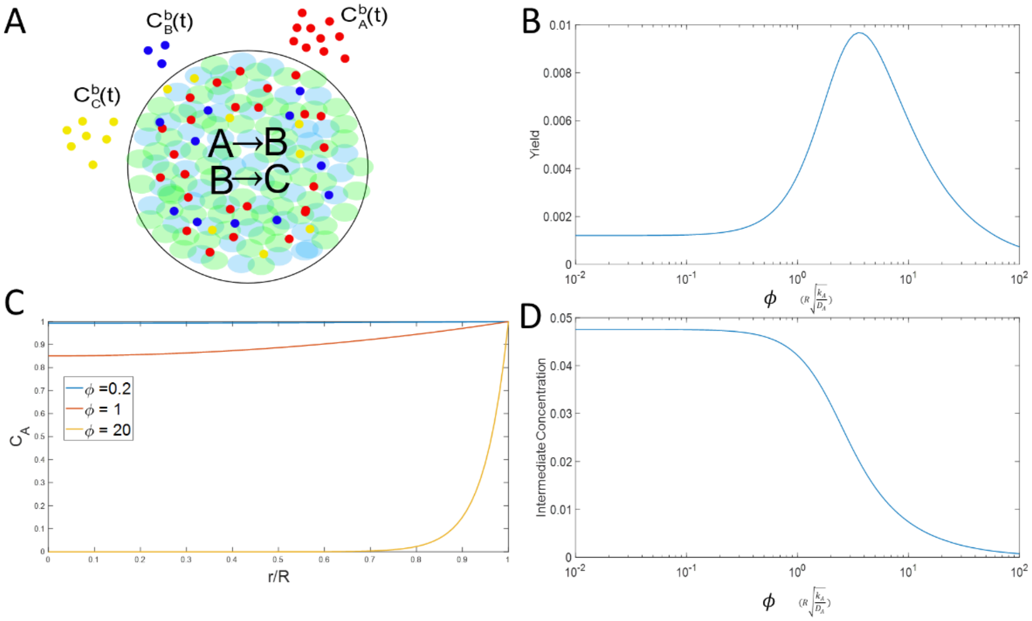


Figure 5.1: Reaction diffusion model for aggregates. A) Aggregates composed of two (or more) strains of bacteria are together able to carry out multiple biosynthetic steps, converting substrate A to product C with intermediate compound B. B) Reaction diffusion modeling demonstrates that there is an optimum Thiele modulus φ (a dimensionless size) near unity where the rates of diffusion and reaction are approximately balanced (other parameters $\gamma B = 1$, $t = 0.5$). C) Initial A concentration profiles. When aggregates are too small ($\varphi << 1$), there is no consumption of A/accumulation of B in aggregates, while when aggregates are too large ($\varphi >> 1$), much of the aggregate volume is inactive. D) Accumulation of intermediate B in aggregates is significantly diminished due to superior channeling in aggregates compared to in the bulk (other parameters $\gamma_B = 1$, $t = 0.5$).

We note that the yield of C is a non-monotonic function of the Thiele modulus, implying that there is an optimum size for aggregates in this model (**Figure 5.1B**). This can be rationalized by noting the complex interplay between reaction and diffusion into aggregates. If an aggregate is too large, then a large portion of the interior of the aggregate will not see substrate, and thus will not participate in the reactions (this is called a dead-zone in heterogeneous catalysis). If the aggregate is instead too small, the intermediate species B will diffuse quickly out of the interior of the aggregate, leading to diffusion into the bulk solution (**Figure 5.1C-D**). Thus, there is some optimum at some intermediate Thiele modulus (generally around unity, where the rates of diffusion and reaction are approximately balanced). As expected, small aggregates reduce to the planktonic, unstructured state, where intra-aggregate diffusion is not limited at all.

Flavonoid Biosynthesis in Aggregates

In order to demonstrate the use of aggregates in co-culture engineering, we require a pathway in which the intermediate is fully cell-permeant, and ideally where the strains must be optimized separately for optimum pathway yields. As an exemplar of the method, we chose to test the synthesis of flavonoids from phenylpropanoic acids. This pathway has previously been demonstrated and optimized in *Escherichia coli* co-culture by Jones and coworkers,[8,20] and so serves as a good test case for our method. We obtained the optimized flavonoid biosynthesis plasmids pETM6-At4CL-PhCHS-CmCHI (which we denote pET-Flav-US) and p168 (which we denote pET-Flav-DS). These plasmids encode the genes responsible for the upstream and downstream portions of the flavonoid biosynthesis pathway (**Figure 5.2A**). The Koffas group has previously found that the upstream strain is most effective when using the strain background BL21star™(DE3)Δ*sucC*Δ*fumC* due to an increase in intracellular malonyl-CoA productivity, so we will likewise use pET-Flav-US in this background in the remainder of this manuscript.[21] We use the standard BL21(DE3) strain with pET-Flav-DS. All strains and plasmids used in this work are described in **Table S5.1**.

As we have described earlier (Chapters 3 and 4), we are able to drive aggregation of *E. coli* into aggregates by the surface display of associative proteins fused to autotransporter membrane proteins. Here, in order to maximize the aggregate sizes, we surface display SpyCatcher and SpyTag domains, which upon association form an isopeptide bond between a lysine residue in SpyCatcher and an aspartic acid residue in SpyTag. Because these domains form an inter-cellular covalent bond, these cause the formation of very stable and relatively large aggregates (Chapter 4). In this work, we use SpyTag and SpyCatcher autotransporter fusions under the control of an arabinose-inducible P_{AraBAD} promoter (pBAD-ST and pBAD-SC, respectively). Because the enzymes are IPTG-inducible, this enables independent temporal control of enzyme and surface display expression. Dual plasmid strains containing pET-Flav-US and pBAD-ST/pBAD-SC are denoted sBRS1 and sBRS2, respectively, while strains containing pET-Flav-DS and pBAD-ST/pBAD-SC are denoted sBRS3 and sBRS4.

Initially, in order to determine the effect of aggregation on titers of flavonoids, we attempted the biosynthesis of catechin from caffeic acid (See **Figure 5.2A**). Saturated cultures of sBRS1, sBRS2, sBRS3, and sBRS4 were grown overnight in LB media with antibiotics (100 µg/mL ampicillin, 35µg/mL chloramphenicol). Thirty milliliter cultures of Andrew's Magic Medium[22] were inoculated at 1% v/v at an Upstream:Downstream strain ratio of 20:1 and a SpyTag:SpyCatcher ratio of 1:1 (Strain ratios were 20:20:1:1 of sBRS1:sBRS2:sBRS3:sBRS4). The Upstream:Downstream strain ratio was chosen based on preliminary experiments (**Figure S5.1**). Note that because the BL21star™(DE3)Δ*sucC*Δ*fumC* strain grows significantly slower than BL21(DE3) (data not shown), the strain ratio in the culture changes during the course of the experiment. Cultures were grown in a shaking incubator at 37°C 250RPM for 4.5 hours (to $OD_{600} \approx 0.55$), before aggregation was induced with 0.1% wt/vol L-arabinose. The temperature was then reduced to 30°C, and bacteria were allowed to aggregate for 2 hours prior to addition of 1mM IPTG to induce expression of the flavonoid synthesis modules. The enzymes were expressed for an additional hour, and then the caffeic acid substrate was added (100 µg/mL). Production of flavonoids took place over an additional 5 hours, and then cultures were harvested and supernatants were run in LC-MS to quantify titers (**Figure 5.2B-C**).

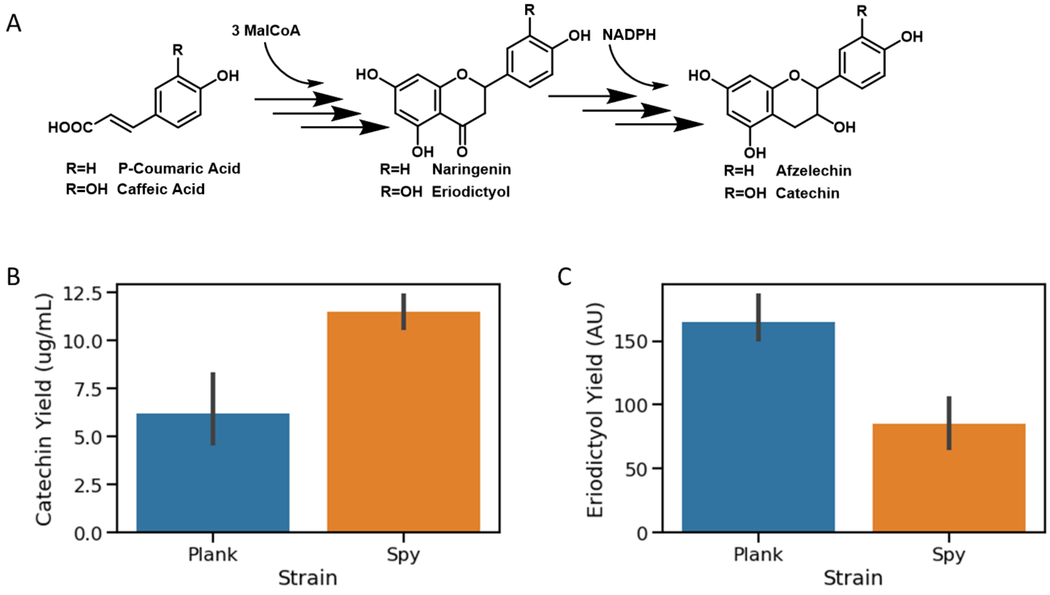


Figure 5.2: Catechin production in aggregates. A) The flavonoid synthesis pathways reacts phenylpropanoic acids to flavonoids in two modules. In the first module, 3 MalCoA molecules are condensed onto the substrate to form Naringenin (from P-Coumaric Acid) or Eriodictyol (from Caffeic Acid). In the second, a series of oxidation-reduction reactions results in the rearrangement and net reduction of the intermediate to Afzelechin or Catechin. B) Co-aggregation of the bacteria expressing the upstream and downstream halves of the flavonoid pathway leads to an enhancement in titer of the final product. (N=6, $p<0.001$, Error bars represent SEM). C) The increase in titer of the final product corresponds strongly with a decrease in the titer of the intermediate (N=6, $p<0.001$)

Upon quantification of catechin titers, we noted that cultures that had been assembled into aggregates exhibited substantially higher catechin titers (11.5 ± .5 µg/mL) than in planktonic (non-aggregated) cultures (6.2 ± 1.0 µg/mL). We note that these titers are lower than previously reported using similar strains,[8] but they were performed on much shorter time scales (5 hours vs 48 hours), and exhibit similar or better volumetric productivities. In order to better investigate the higher yields of the aggregated samples, we also quantified the titers of the accumulated intermediate species eriodictyol **(Figure 5.2C)**. As expected from the reaction diffusion model (**Figure 5.1D**), eriodictyol is significantly depleted in the aggregated samples due to its enhanced channeling inside the aggregates. We also observe that the enhancement of titers due to aggregation is higher than that caused by the metabolic defect from over-expression of membrane proteins.

Effect of Aggregate Size on Flavonoid Titers

One prediction of the reaction diffusion model described above is that the size of the aggregates (through its effect on the Thiele modulus) is an important governing parameter on the performance of the biocatalyst. Thus, we resolved to perturb the size of the aggregates to analyze the impact on the effectiveness of the aggregates. We are not able to efficiently increase the size of the aggregates in this system because we are already using strong inter-cellular covalent bonds between cells and are expressing the surface proteins at very high levels. Thus, we instead decrease the size of the aggregates. We have previously discussed methods to control the size of the aggregates including using different associative proteins and decreasing the expression levels of the surface proteins (Chapter 4); however, in preliminary experiments we found that weaker associating proteins and/or decreased expression levels did not lead to large-scale aggregation in this system. Instead, we hypothesized that we could decrease aggregate sizes using different numbers of SpyCatcher and SpyTag displaying-cells, in a method analogous to control of polymer molecular weights in condensation step growth polymerization.[23]

First, in order to enable the measurement of the sizes of the aggregates by fluorescence confocal microscopy, fluorescent protein expression cassettes were added to pET-Flav-US and pET-Flav-DS using Gibson isothermal assembly to make plasmids pET-Flav-US-mW and pET-Flav-DS-mC (expressing mWasabi and mCherry, respectively under the constitutive J23100 Anderson promoter), which were co-transformed with pBAD-ST and pBAD-SC to form sBRS5, sBRS6, sBRS7, and sBRS8. Use of these plasmids in place of the pET-Flav-US and pET-Flav-DS plasmids may decrease titers somewhat (especially when cells are aggregated, see

Figure S5.2), but enables quantification of aggregate size. By titrating the ratio of SpyCatcher to SpyTag expressing cells, we are able to control the average (volume-weighted) sizes of aggregates from $6.4*10^4$ μm^3 to $5*10^3$ μm^3, and then to 100 μm^3 in the planktonic case (**Figure 5.3A-D**). We note that the size of aggregates is positively correlated with the catechin titers, qualitatively corresponding to the results from our reaction-diffusion model (**Figure 5.3E**).

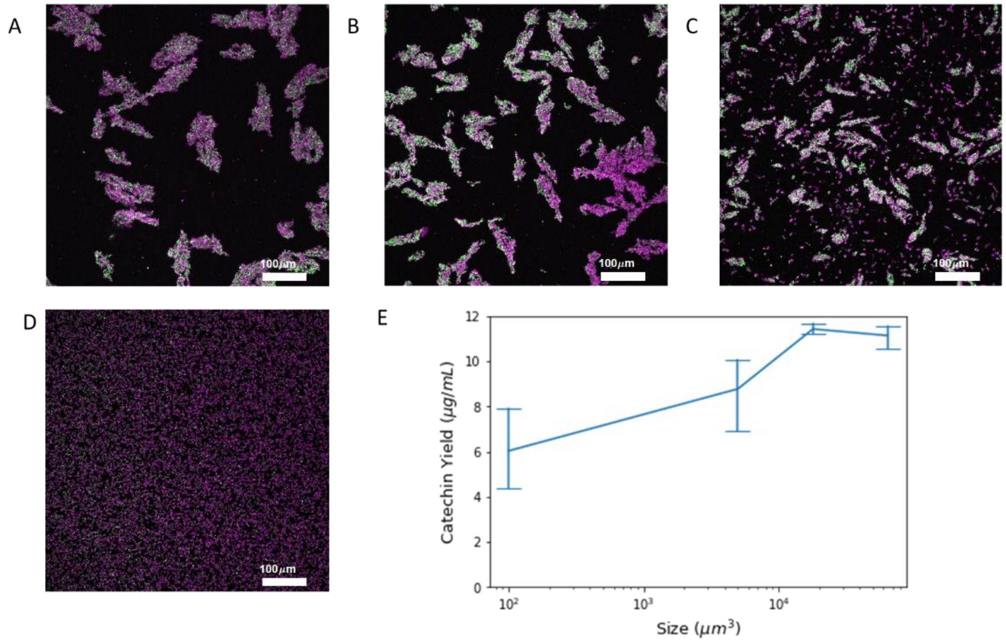


Figure 5.3: Effect of aggregate size on catechin titers. A-C) Fluorescent micrographs of aggregates at A) 1:1 SC:ST ratio, B) 3:1 SC:ST ratio, or C) 5:1 SC:ST ratio. D) Fluorescent micrograph of planktonic (unaggregated) cells. Upstream cells are shown in green, and downstream cells are shown in magenta. E) There is a positive correlation between catechin yields and the size of the aggregates (N=4, error bars represent SEM).

Three-part Synthesis of Afzlechin in Aggregates

In order to demonstrate the versatility of this method, we next chose to implement a longer pathway in which we combine the flavonoid synthesis pathway with an upstream p-coumaric acid synthesis module, enabling the synthesis of afzelechin from tyrosine (**Figure 5.4A**). A plasmid containing a tyrosine ammonia lyase (TAL) under the control of the P_{Trc} promoter was a generous gift of Prof. Guillermo Gosset.[24] pTrc-TAL was co-transformed with pBAD-ST and pBAD-SC in a BL21 background to form sBRS9 and sBRS10 strains.

We co-cultured strains containing the p-coumaric acid synthesis module, and the two halves of the flavonoid synthesis pathway along with the aggregation constructs at a strain ratio of 6:12:2 p-coumaric acid synthesis module:upstream flavonoid module: downstream flavonoid module (optimum strain ratios obtained by prior optimization, Figure S5.3). As before, we induced aggregation, followed by the enzymes, and then addition of substrate (here tyrosine). After a five-hour fermentation, we harvested samples and ran LC-MS. Upon quantification, we discovered that corresponding to our expectations, aggregation enhances yield of afzelechin by a factor of approximately 2 compared to planktonic cells (**Figure 5.4B**).

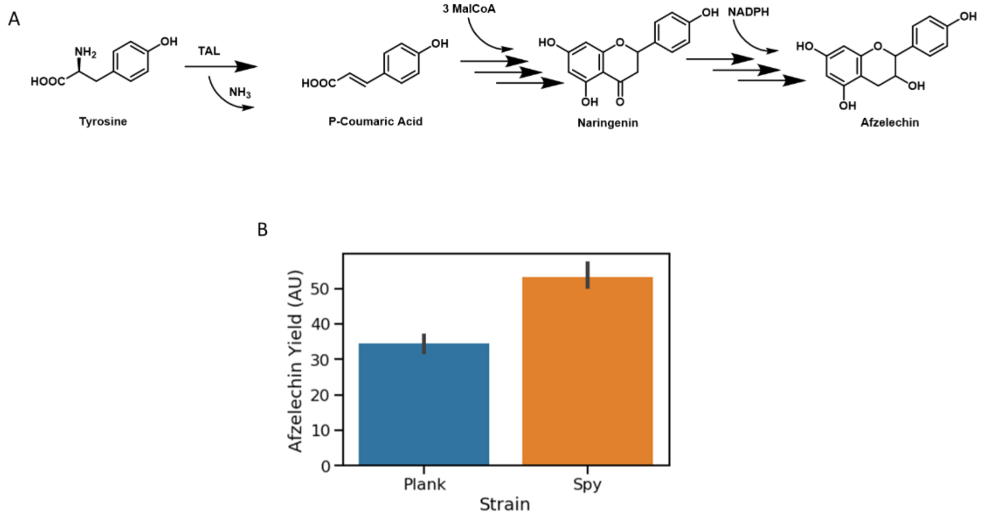


Figure 5.4: Three-part synthesis of afzelechin from tyrosine. A) Afzelechin is made biosynthetically from tyrosine in three parts. First tyrosine is transformed to p-coumaric acid by tyrosine ammonia lyase in one module, followed by condensation with 3 MalCoA to narenginin, and finally reorganization and reduction to afzelechin in the final module. B) Aggregation significantly enhances the titer of afzelechin by approximately 70% in this system (N=3, p<.001).

Conclusion

As metabolic engineers continue to increase the length and complexity of biosynthetic pathways, the ability to split these metabolic pathways between multiple bacterial strains becomes more important. We have demonstrated that by programming bacteria expressing parts of a longer pathway to associate into mesoscopic, closely packed, bacterial aggregates, we can enhance titers of the desired chemical. The primary mechanism for this increase in yields is due to the enhancement in the rate of diffusion of the product of one bacterial strain to the next due to accumulation of the intermediate in the aggregate, rather than being diluted into the bulk media. In our experience, the ability of aggregation to enhance yields takes place in situations where the accumulation of the intermediate in the bulk culture is considerable (and where it is decreased by aggregation).

We believe that the strategy that we have developed in this work is fairly general and may be applied to a variety of biosynthetic pathways and bacterial strains. The ability of this technique to enhance yields depends on several factors. First, the bacterial strain(s) of interest must be genetically tractable with the ability to surface-display target proteins. The pathway must be capable of being separated into several distinct modules, where the intermediate chemical species is cell-permeant and where its diffusion must be meaningfully decreased by encapsulation in a bacterial aggregate (i.e. $\varphi>1$). This rate of diffusion may be meaningfully affected by the properties of the bacterial aggregate and the compound of interest; species that appreciably bind to cell membranes may be especially effective here.

The techniques developed in this work are likely broadly scalable to larger culture volumes, albeit with care to control high shear stresses (bacterial aggregates are much more sensitive to high shear stresses than planktonic cultures). The use of these aggregates also allows the facile recovery of whole cell catalysts due to their rapid settling in stagnant culture (settling velocity>2mm/s). Further investigation of the physiological state of cells in these biological aggregates and the biophysical microenvironments inside the aggregates will enable the further modeling and more precise engineering of these novel biocatalysts.

Acknowledgements

This work was supported by Defense Advanced Research Projects Agency Biological Robustness in Complex Settings Contract HR001-15-C-0093 and the Caltech Jacobs Institute for Molecular Engineering for Medicine. B.R.S. is supported by NIH Training Grant 1T32GM112592 and the Rosen Center for Bioengineering at Caltech. We thank Dr. Andres Collazo for assistance with confocal microscopy and Drs. Mona Shahgholi and Joshua Baccile for assistance with LC-MS analysis. Imaging was performed in the Biological Imaging Facility, with the support of the Caltech Beckman Institute and the Arnold and Mabel Beckman Foundation. We thank Professor Guillermo Gossett for the kind gift of the pTrc-TAL plasmid, and Professor Mattheos Koffas and Abinaya Badri for the kind gift of pET-Flav-US, pET-Flav-DS, and BL21star™(DE3)Δ*sucC*Δ*fumC*.

Materials and Methods

General

Restriction enzymes, ligase, and Q5 DNA polymerase were purchased from New England Biolabs (Beverly, NJ). DNA oligos were purchased from Integrated DNA Technologies (Coralville, IA). Caffeic acid, p-coumaric acid, and naringenin were purchased from Indofine Chemical Co (Hillsborough, NJ). Catechin and eriodictyol were purchased from Sigma Aldrich (St. Louis, MO)

Bacterial Strains

As described in the main text, BL21star™(DE3)Δ*sucC*Δ*fumC Escherchia coli* was a generous gift from Mattheos Koffas and was used for the upstream flavonoid reaction steps. BL21(DE3) was used for the downstream flavonoid pathway steps, while BL21 was used for p-coumaric acid production.

pET-Flav-US was a gift from Mattheos Koffas (pETM6-At4CL-PhCHS-CmCHI (C5 mutant, pFlavo-opt) Addgene plasmid # 73404 ; http://n2t.net/addgene:73404 ; RRID:Addgene_73404).

pET-Flav-DS was a gift from Mattheos Koffas (p168, Addgene plasmid # 62618 ; http://n2t.net/addgene:62618 ; RRID:Addgene_62618)

pBAD-ST and pBAD-SC have been previously reported by our group (Chapter 3 of this thesis).

pTrc-TAL was a gift from Guillermo Gossett.

pET-Flav-US-mW and pET-Flav-DS-mC were made by Gibson isothermal assembly. pET-Flav-US and pET-Flav-DS were linearized via divergent PCR at the region between the rop and lacI loci. mWasabi and mCherry expression cassettes under the control of the constitutive J23100 Anderson promoter were obtained via PCR, while also appending overlapping overhangs with the linearized vector. Gibson assembly was then performed to re-circularize the pET vectors with the fluorescent cassettes. Construct identity was confirmed by Sanger sequencing.

Fermentation Reactions

Saturated overnight cultures of bacteria were grown in LB with ampicillin (100 µg/mL) and chloramphenicol (25 µg/mL) as appropriate. These cultures were mixed at the appropriate volume ratios (described in main text), and inoculated at a total dilution of 1:100 in 30mL Andrew's Magic Medium (AMM) with antibiotics in an unbaffled 125 mL Erlenmeyer flask.[25] These cultures were grown for 4.5 hours at 37°C with shaking, and then induced with 0.1% (v/v) L-Arabinose. The bacteria were allowed to aggregate for 2 hours at 30°C before enzymes were induced with 1mM IPTG for 1 hour. Substrate was then added (10 µg/mL caffeic acid or p-coumaric acid, 54 µg/mL tyrosine), and incubated for an additional 5 hours with shaking. Samples were then taken for LC-MS and microscopy analysis.

Liquid Chromatography Mass Spectrometry

Samples from LC-MS were immediately centrifuged and filtered to remove biomass. Supernatants were directly injected (10 µL) for LC-MS Analysis. A Waters Acquity SDS UPLC equipped with a CORTECS-C18 UPLC column maintained at 30°C was used for separation, and a LCT Premier TOF was used for detection in negative mode. The mobile phases were water (A) and acetonitrile (B). The following gradient was used at a flow rate of 0.4 mL/min: 0 min, 95% A; 0.2 min, 95% A; 7.5 min, 40% A; 8 min, 95% A; 10 min, 95% A. Extracted ion chromatograms were extracted from LC-MS traces and integrated using a custom Python script. A standard curve of catechin dissolved in a "blank" spent fermentation broth was used for absolute quantification of catechin titers.

Image Acquisition and Analysis

Microscopy images were taken on a Zeiss 800 LSM inverted confocal microscope (Carl Zeiss AG, Oberkochen, Germany). All image analysis was performed using custom Matlab scripts. Aggregate size analysis was performed similarly to what we described previously.[26] Briefly, confocal z-stacks were manually thresholded based on the intensity in each fluorescent channel. Pixels above the threshold were described as "bright." Contiguous "bright" pixels (in 3D) were identified, and the observed volume of each aggregate was determined. The volume-weighted average volume of each aggregate was determined using the following equation:

$$\bar{V} = \frac{\sum V_i^2}{\sum V_i}$$

where sums are taken over all of the aggregate volumes. This average represents the volume of the aggregate that the average bacterium would be found in, and is more appropriate than the number-weighted average, which is dominated by disassociated bacteria. The volumes determined by this method are over-estimates of the true volume due to fluorescence spilling into adjacent voxels.

Supplemental Discussion

Derivation of Reaction-Diffusion Models

Here, we provide a full derivation of Equations 1-6 in the main text. We have an aggregate of radius R, made up of bacteria that can carry out the first-order reactions A->B and B->C. Though in experimental conditions, different bacteria will carry out these steps, here we assume that the aggregates are sufficiently large that we can consider the aggregates to be a continuum of both strains of bacteria. These aggregates are found at a cell-volume fraction ε_c in a culture medium. First, we consider the interior of the aggregates by writing mass-conservation equations for each of the species.

$$\frac{\partial C_A}{\partial t} = D_A \nabla^2 C_A - k_A C_A \qquad (S5.1)$$

(S5.1)

$$\frac{\partial C_B}{\partial t} = D_B \nabla^2 C_B - k_B C_B + k_A C_A \qquad (S5.2)$$

(S5.2)

$$\frac{\partial C_C}{\partial t} = D_C \nabla^2 C_C + k_B C_B \qquad (S5.3)$$

(S5.3)

where C_i is the concentrations of species i, D_i is the diffusion coefficient of species i in the aggregate, k_i is the first order reaction rate of species i, and ∇^2 is the Laplacian operator.

Now, we consider the boundary/initial conditions for these equations. Because the PDE's are 2nd order in space and 1st order in time, we require two spatial boundary conditions and an initial condition for each species. Here, we choose to use a no-flux (Neuman) boundary condition at the center in order to maintain finite species concentrations. At the surface, we have a boundary condition where the flux out of the aggregate is equal to the flux into the aggregate from the bulk. We write these boundary conditions below:

$$\nabla C_A(t, r = 0) = 0 \qquad (S5.4)$$

$$\nabla C_B(t, r = 0) = 0 \qquad (S5.5)$$

$$\nabla C_C(t, r = 0) = 0 \qquad (S5.6)$$

$$D_A \nabla \mathrm{C_A}(t, r = R) = k_c(C_A^B(t) - C_A(t, r = R)) \qquad (S5.7)$$

$$D_A \nabla \mathrm{C_A}(t, r = R) = k_c(C_B^B(t) - C_B(t, r = R)\,) \qquad (S5.8)$$

$$D_A \nabla \mathrm{C_A}(t, r = R) = k_c(C_C^B(t) - C_C(t, r = R)) \qquad (S5.9)$$

Although, we could use this set of boundary conditions, because we are vigorously mixing the aggregates, we expect that most resistance to mass transport takes place in the aggregates (formally, that the Biot number $k_cR/D >> 1$), we will instead assume that there is no interfacial resistance to mass transfer, and that the concentration at the surface of the aggregate is equal to the bulk concentration. Equations S5.7-9 then become:

$$C_A(t, r = R) = C_A^B(t) \qquad (S5.10)$$

$$C_B(t, r = R) = C_B^B(t) \qquad (S5.11)$$

$$C_C(t, r = R) = C_C^B(t) \qquad (S5.12)$$

Now, we consider the bulk concentration. The bulk concentration will be changed only by reaction in the aggregates, as shown in the equations below:

$$\frac{dC_A^B}{dt} = -\frac{\epsilon_c k_A \int C_A dV}{\int dV} \qquad (S5.13)$$

$$\frac{dC_B^B}{dt} = \frac{\epsilon_c \int k_A C_A - k_B C_B dV}{\int dV} \qquad (S5.14)$$

$$\frac{dC_C^B}{dt} = +\frac{\epsilon_c k_B \int C_B dV}{\int dV} \qquad (S5.15)$$

The initial conditions for the PDEs S5.1-S3 and ODEs S5.13-15 are as follows, assuming an initial bolus of A:

$$C_A(t = 0, r) = C_A^B(t = 0) = \; C_A^0 \qquad (S5.16)$$

$$C_B(t=0,r) = C_B^B(t=0) = 0 \qquad (S5.17)$$

$$C_C(t=0,r) = C_C^B(t=0) = 0 \qquad (S5.18)$$

Equations S5.1-6, S5.10-12, and S5.16-S18 represent a fully defined model of the system that given parameter values could be solved. However, in order to decrease the number of parameters and understand the system more fully, we will nondimensionalize the equations.

It is straightforward to see that a natural concentration scale in this system is the initial concentration of A and that a natural length scale is the radius of the aggregate R. We nondimensionalize using these relations (leaving time for now), and rewrite the equations:

$$\widehat{C_A} = \frac{C_A}{C_A^0} \qquad (S5.19)$$

$$\widehat{C_B} = \frac{C_B}{C_A^0} \qquad (S5.20)$$

$$\widehat{C_C} = \frac{C_C}{C_A^0} \qquad (S5.21)$$

$$\hat{r} = \frac{r}{R} \qquad (S5.22)$$

$$\frac{\partial \widehat{C_A}}{\partial t} = \frac{D_A}{R^2}\nabla^2 \widehat{C_A} - k_A \widehat{C_A} \qquad (S5.23)$$

$$\frac{\partial \widehat{C_B}}{\partial t} = \frac{D_B}{R^2}\nabla^2 \widehat{C_B} + k_A \widehat{C_A} - k_B \widehat{C_B} \qquad (S5.24)$$

$$\frac{\partial \widehat{C_C}}{\partial t} = \frac{D_C}{R^2}\nabla^2 \widehat{C_C} + k_B \widehat{C_B} \qquad (S5.25)$$

$$\frac{d\widehat{C_A^B}}{dt} = -\frac{\epsilon_c k_A \int \widehat{C_A} dV}{\int dV} \qquad (S5.26)$$

$$\frac{d\widehat{C_B^B}}{dt} = \frac{\epsilon_c \int k_A \widehat{C_A} - k_B \widehat{C_B} dV}{\int dV} \qquad (S5.27)$$

$$\frac{d\widehat{C_C^B}}{dt} = +\frac{\epsilon_c k_B \int \widehat{C_B} dV}{\int dV} \qquad (S5.28)$$

$$\nabla\widehat{C_A}(t,\hat{r}=0)=0 \tag{S5.29}$$

$$\nabla\widehat{C_B}(t,\hat{r}=0)=0 \tag{S5.30}$$

$$\nabla\widehat{C_C}(t,\hat{r}=0)=0 \tag{S5.31}$$

$$\widehat{C_A}(t,\hat{r}=1)=\widehat{C_A^B}(t) \tag{S5.32}$$

$$\widehat{C_B}(t,\hat{r}=1)=\widehat{C_B^B}(t) \tag{S5.33}$$

$$\widehat{C_C}(t,\hat{r}=1)=\widehat{C_C^B}(t) \tag{S5.34}$$

Now, we consider the nondimensionalization of time. There are two time-scales in this problem: one is the time-scale of reaction inside aggregates (k_A^{-1}), and one is the time-scale of bulk concentration changes ($[\varepsilon_C k_A]^{-1}$). Because we are interested in the accumulation of products in the bulk, the more natural time scale is the latter, so we nondimensionalize using that time scale. Doing this implies that the time scale of changes in the aggregate are much faster than those outside, so we assume that the aggregates are always at a pseudo-steady state, and thus the time derivative terms in the PDE's are eliminated, and obtain the following ODE's:

$$\hat{t}=\epsilon_C k_A t \tag{S5.35}$$

$$0=\frac{D_A}{k_A R^2}\nabla^2\widehat{C_A}-\widehat{C_A} \tag{S5.36}$$

$$0=\frac{D_B}{k_A R^2}\nabla^2\widehat{C_B}+\widehat{C_A}-\frac{k_B}{k_A}\widehat{C_B} \tag{S5.37}$$

$$0=\frac{D_C}{k_A R^2}\nabla^2\widehat{C_C}+\frac{k_B}{k_A}\widehat{C_B} \tag{S5.38}$$

Defining dimensionless parameters gives:

$$0=\frac{1}{\phi^2}\nabla^2\widehat{C_A}-\widehat{C_A} \tag{S5.39}$$

$$0=\frac{\widehat{D_B}}{\phi^2}\nabla^2\widehat{C_B}+\widehat{C_A}-\gamma_B\widehat{C_B} \tag{S5.40}$$

$$0=\frac{\widehat{D_c}}{\phi^2}\nabla^2\widehat{C_C}+\frac{k_B}{k_A}\widehat{C_B} \tag{S5.41}$$

$$\phi = R\sqrt{\frac{k_A}{D_A}} \tag{S5.42}$$

$$\widehat{D_B} = \frac{D_B}{D_A} \tag{S5.43}$$

$$\widehat{D_C} = \frac{D_C}{D_A} \tag{S5.44}$$

$$\gamma_B = \frac{k_B}{k_A} \tag{S5.45}$$

Here, ϕ is the Thiele modulus for a 1st order reaction in a sphere,[2] $\widehat{D_i}$ is the ratio of the diffusion coefficient for species i to that of A, and $\gamma_B$is the ratio of the reaction rates of the second reaction to that of the first. For many reaction pathways, diffusion coefficients will not differ very much, so $D_i \approx 1$, but we keep it for completeness here. The equations for the bulk concentrations become:

$$\frac{d\widehat{C_A^B}}{d\hat{t}} = -\frac{\int \widehat{C_A} dV}{\int dV} \tag{S5.46}$$

$$\frac{d\widehat{C_B^B}}{d\hat{t}} = \frac{\int \widehat{C_A} - \gamma_B \widehat{C_B} dV}{\int dV} \tag{S5.47}$$

$$\frac{d\widehat{C_C^B}}{d\hat{t}} = \frac{\gamma_B \int \widehat{C_B} dV}{\int dV} \tag{S5.48}$$

Now, finally the boundary/initial conditions for the dimensionless system are:

$$\widehat{C_A^B}(\hat{t} = 0) = 1 \tag{S5.49}$$

$$\widehat{C_B^B}(\hat{t} = 0) = 0 \tag{S5.50}$$

$$\widehat{C_C^B}(\hat{t} = 0) = 0 \tag{S5.51}$$

$$\widehat{C_A}(\hat{r} = 1) = \widehat{C_A^B} \tag{S5.52}$$

$$\widehat{C_B}(\hat{r} = 1) = \widehat{C_B^B} \tag{S5.53}$$

$$\widehat{C_C}(\hat{r} = 1) = \widehat{C_C^B} \tag{S5.54}$$

$$\nabla\widehat{C_A}(t, \hat{r} = 0) = \nabla\widehat{C_B}(t, \hat{r} = 0) = \nabla\widehat{C_C}(t, \hat{r} = 0) = 0 \tag{S5.55}$$

Equations S5.39-S55 are a full dimensionless description of the model, and are shown in the main text (with accents dropped for brevity). Matlab was used for numerical solution of the equations. At each time-point, the concentration profiles in the aggregates were solved using Matlab's boundary value solver bvp4c. Euler's method was used to integrate Equations S5.46-S48 subject to the concentration profiles at each time-point. The Matlab scripts used for this numerical solution is available at http://tirrell-lab.caltech.edu/code.

In this framework, it is fairly simple to add additional species, different reaction kinetics, or terms related to nutrient limitation in the interior of the aggregate. However, we believe that this model does capture the essential tradeoffs between reaction and diffusion and the non-monotonicity of biosynthetic yield.

Supplemental Figures and Table

Table S5.1: Strains and Plasmids Used in this Work

Strain/Plasmid Name	Description	Source
pBAD-ST	SpyTag-Autotransporter Fusion under P_{AraBAD}	Chapter 3
pBAD-SC	SpyCatcher-Autotransporter Fusion under P_{AraBAD}	Chapter 3
pET-Flav-US	Upstream Flavonoid pathway module	(Addgene 62168)[3]
pET-Flav-DS	Downstream Flavonoid pathway module	(Addgene 73404)[4]
pET-Flav-US-mWasabi	Upstream Flavonoid pathway and constitutive mWasabi expression	This work
pET-Flav-DS-mCherry	Downstream Flavonoid pathway and constitutive mCherry expression	This work
BL21star™(DE3)ΔsucCΔfumC	Upstream Flavonoid Expression Background	[5]
BL21(DE3)	Downstream Flavonoid Expression Background	[3]
sBRS1	BL21star™(DE3)ΔsucCΔfumC with pBAD-ST and pET-Flav-US plasmids	This work
sBRS2	BL21star™(DE3)ΔsucCΔfumC with pBAD-SC and pET-Flav-US plasmids	This work

sBRS3	BL21 (DE3) with pBAD-ST and pET-Flav-DS plasmids	This work
sBRS4	BL21 (DE3) with pBAD-SC and pET-Flav-DS plasmids	This work
sBRS5	BL21star™(DE3)ΔsucCΔfumC with pBAD-ST and pET-Flav-US-mWasabi plasmids	This work
sBRS6	BL21star™(DE3)ΔsucCΔfumC with pBAD-SC and pET-Flav-US-mWasabi plasmids	This work
sBRS7	BL21 (DE3) with pBAD-ST and pET-Flav-DS-mCherry plasmids	This work
sBRS8	BL21 (DE3) with pBAD-SC and pET-Flav-DS-mCherry plasmids	This work
sBRS9	BL21 with pBAD-ST and pTrc-TAL	This work
sBRS10	BL21 with pBAD-SC and pTrc-TAL	This work

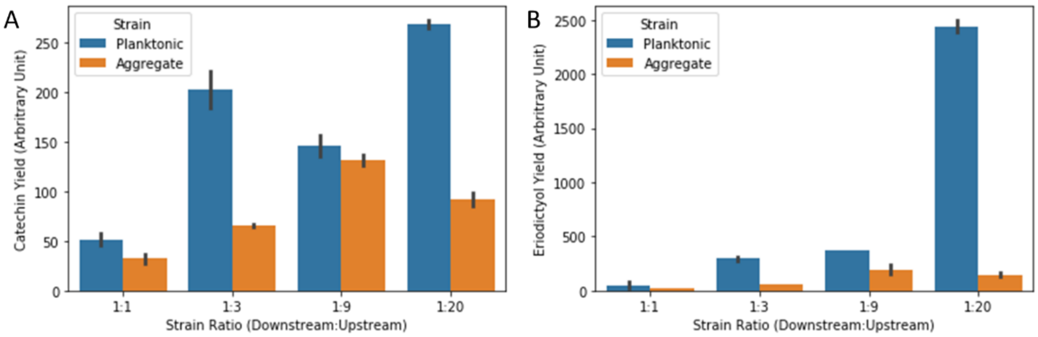


Figure S5.1: Catechin Synthesis Strain Ratio Scouting. Various ratios of Downstream:Upstream initial strain ratios were examined to find an optimum state for further experimentation. We observed that at a strain ratio of 1:20, after 16 hours fermentation, there is a strong accumulation of eriodictyol in the planktonic cells, but not in the aggregated cells. Thus, we chose this strain ratio for further investigation, while also decreasing the fermentation time to five hours.

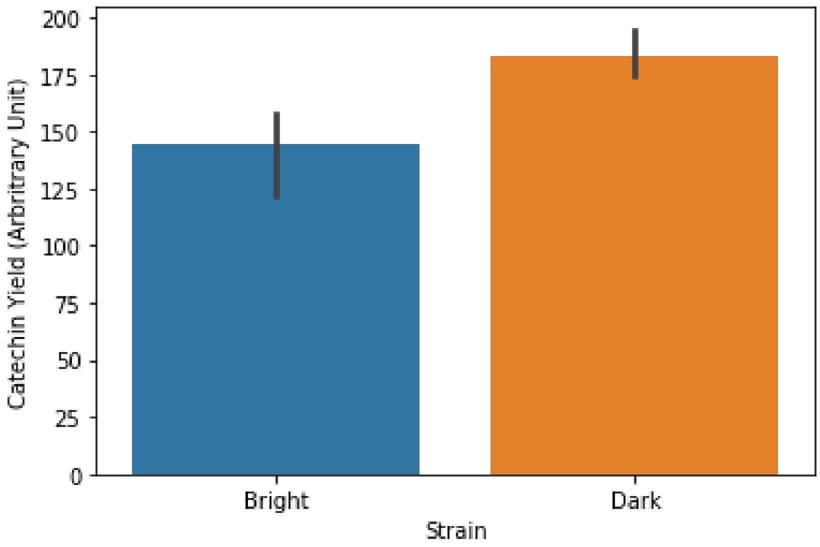


Figure S5.2: Fluorescence Effect on Catechin Titer**s**. Aggregates composed of sBRS1-4, were compared to those composed of sBRS5-8 in their ability to synthesize catechin. Due to some amount of additional metabolic stress on the cells from the synthesis of mWasabi/mCherry, the titers of “Bright” FP expressing cells (sBRS5-8) is somewhat lower than that of the “Dark” cells (sBRS1-4).

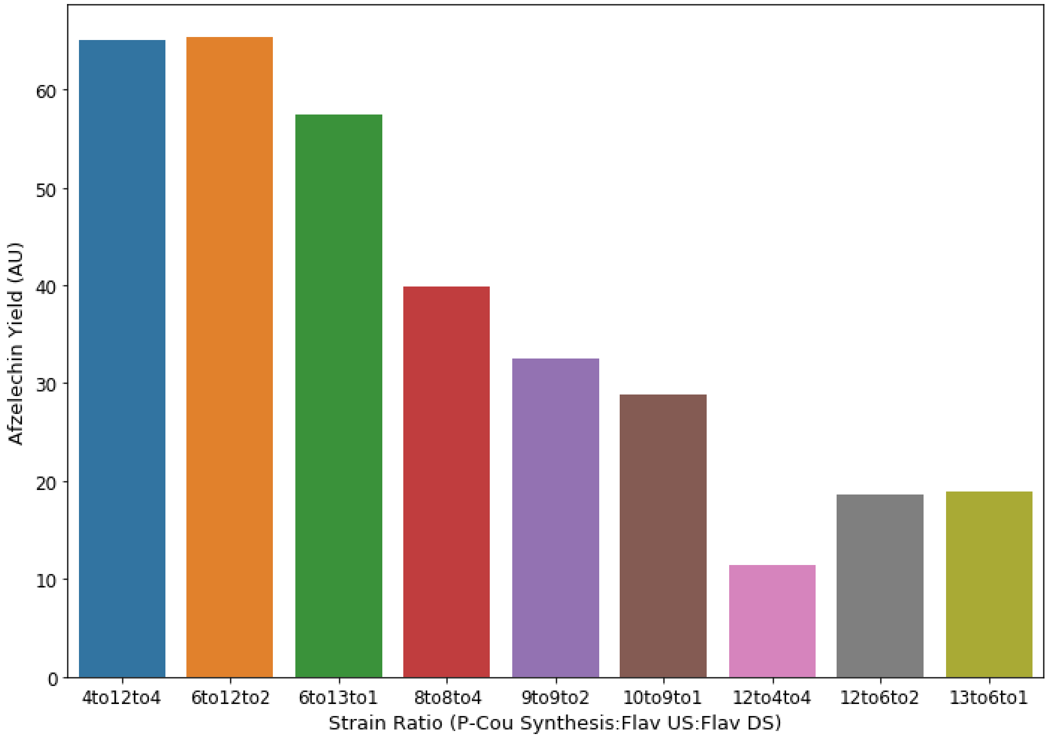


Figure S5.3: Afzelechin Synthesis Strain Ratio Scouting. Aggregates composed of sBRS1-4 and sBRS9-10 were combined in co-culture at defined ratios, aggregated as usual, and allowed to transform tyrosine to afzelechin for five hours. The amount of the p-coumaric acid synthesis strain is negatively related to afzelechin synthesis (as it is fairly efficient). We chose the 6:12:2 strain ratio for further study.

References

1 D. Machado and M. J. Herrgård, Co-evolution of strain design methods based on flux balance and elementary mode analysis, *Metab. Eng. Commun.*, 2015, **2**, 85–92.

2 M. R. Antoniewicz, Methods and advances in metabolic flux analysis: A mini-review, *J. Ind. Microbiol. Biotechnol.*, 2015, **42**, 317–325.

3 E. Leonard, P. K. Ajikumar, K. Thayer, W. H. Xiao, J. D. Mo, B. Tidor, G. Stephanopoulos and K. L. J. Prather, Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 13654–13659.

4 H. H. Wang, F. J. Isaacs, P. A. Carr, Z. Z. Sun, G. Xu, C. R. Forest and G. M. Church, Programming cells by multiplex genome engineering and accelerated evolution, *Nature*, 2009, **460**, 894–898.

5 J. D. Keasling, Synthetic biology and the development of tools for metabolic engineering, *Metab. Eng.*, 2012, **14**, 189–195.

6 G. Wu, Q. Yan, J. A. Jones, Y. J. Tang, S. S. Fong and M. A. G. Koffas, Metabolic burden: Cornerstones in synthetic biology and metabolic engineering applications, *Trends Biotechnol.*, 2016, **34**, 652–664.

7 A. H. Chen and P. A. Silver, Designing biological compartmentalization, *Trends Cell Biol.*, 2012, **22**, 662–670.

8 J. A. Jones, V. R. Vernacchio, A. L. Sinkoe, S. M. Collins, M. H. A. Ibrahim, D. M. Lachance, J. Hahn and M. A. G. Koffas, Experimental and computational optimization of an Escherichia coli co-culture for the efficient production of flavonoids, *Metab. Eng.*, 2016, **35**, 55–63.

9 M. R. Long, W. K. Ong and J. L. Reed, Computational methods in metabolic engineering for strain design, *Curr. Opin. Biotechnol.*, 2015, **34**, 135–141.

10 P. D. Matthews and E. T. Wurtzel, Metabolic engineering of carotenoid accumulation in Escherichia coli by modulation of the isoprenoid precursor

pool with expression of deoxyxylulose phosphate synthase, *Appl. Microbiol. Biotechnol.*, 2000, **53**, 396–400.

11 H. Zhang and X. Wang, Modular co-culture engineering, a new approach for metabolic engineering, *Metab. Eng.*, 2016, **37**, 114–121.

12 S. Elias and E. Banin, Multi-species biofilms: Living with friendly neighbors., *FEMS Microbiol. Rev.*, 2012, **36**, 990–1004.

13 L. Hall-Stoodley, J. W. Costerton and P. Stoodley, Bacterial biofilms: From the natural environment to infectious diseases, *Nat. Rev. Microbiol.*, 2004, **2**, 95–108.

14 T. Mah and G. O'Toole, Mechanisms of biofilm resistance to antimicrobial agents, *Trends Microbiol.*, 2001, **9**, 34–39.

15 M. Miller and B. Bassler, Quorum sensing in bacteria, *Annu. Rev. Microbiol.*

16 A. Prindle, J. Liu, M. Asally, S. Ly, J. Garcia-Ojalvo and G. M. Süel, Ion channels enable electrical communication in bacterial communities, *Nature*, 2015, **527**, 59–63.

17 A. Sakanaka, M. Kuboniwa, H. Takeuchi, E. Hashino and A. Amano, Arginine-ornithine antiporter ArcD controls arginine metabolism and interspecies biofilm development of Streptococcus gordonii, *J. Biol. Chem.*, 2015, **290**, 21185–21198.

18 P. S. Stewart, Diffusion in biofilms, *J. Bacteriol.*, 2003, **185**, 1485–1491.

19 E. W. Thiele, Relation between catalytic activity and size of particle, *Ind. Eng. Chem.*, 1939, **31**, 916–920.

20 S. Zhao, J. A. Jones, D. M. Lachance, N. Bhan, O. Khalidi, S. Venkataraman, Z. Wang and M. A. G. Koffas, Improvement of catechin production in Escherichia coli through combinatorial metabolic engineering, *Metab. Eng.*, 2015, **28**, 43–53.

21 P. Xu, S. Ranganathan, Z. L. Fowler, C. D. Maranas and M. A. G. Koffas, Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA, *Metab. Eng.*,

2011, **13**, 578–587.

22 W. He, L. Fu, G. Li, J. Andrew Jones, R. J. Linhardt and M. Koffas, Production of chondroitin in metabolically engineered E. coli, *Metab. Eng.*, 2015, **27**, 92–100.

23 P. J. Flory, Fundamental principles of condensation polymerization, *Chem. Rev.*, 1946, **39**, 137–197.

24 J. M. Camacho-Zaragoza, G. Hernández-Chávez, F. Moreno-Avitia, R. Ramírez-Iñiguez, A. Martínez, F. Bolívar and G. Gosset, Engineering of a microbial coculture of Escherichia coli strains for the biosynthesis of resveratrol, *Microb. Cell Fact.*, 2016, **15**, 163.

25 J. A. Jones, V. R. Vernacchio, D. M. Lachance, M. Lebovich, L. Fu, A. N. Shirke, V. L. Schultz, B. Cress, R. J. Linhardt and M. A. G. Koffas, EPathOptimize: A combinatorial approach for transcriptional balancing of metabolic pathways, *Sci. Rep.*, , DOI:10.1038/srep11301.

26 M. Obana, B. R. Silverman and D. A. Tirrell, Protein-mediated colloidal assembly, *J. Am. Chem. Soc.*, 2017, **139**, 14251–14256.