ABSTRACT

Title of Dissertation: BIOENGINEERED CONDUITS FOR DIRECTING DIGITIZED MOLECULAR-BASED INFORMATION

Jessica L. Terrell, Doctor of Philosophy, 2015

Directed By: Professor William E. Bentley

Fischell Department of Bioengineering

Molecular recognition is a prevalent quality in natural biological environments: molecules- small as well as macro- enable dynamic response by instilling functionality and communicating information about the system. The accession and interpretation of this rich molecular information leads to context about the system. Moreover, molecular complexity, both in terms of chemical structure and diversity, permits information to be represented with high capacity. Thus, an opportunity exists to assign molecules as chemical portrayals of natural, non-natural, and even non-biological data. Further, their associated upstream, downstream, and regulatory pathways could be commandeered for the purpose of data processing and transmission.

This thesis emphasizes molecules that serve as units of information, the processing of which elucidates context. The project first strategizes biocompatible assembly processes that integrate biological componentry in organized configurations for molecular transfer (e.g. from a cell to a receptor). Next, we have explored the use of DNA for its potential to store data in richer, digital

forms. Binary data is embedded within a gene for storage inside a cell carrier and is selectively conveyed. Successively, a catalytic relay is developed to transduce similar data from sequence-based DNA storage to a delineated chemical cue that programs cellular phenotype. Finally, these cell populations are used as mobile information processing units that independently seek and collectively categorize the information, which is fed back as 'binned' output.

Every development demonstrates a transduction process of molecular data that involves input acquisition, refinement, and output interpretation. Overall, by equipping biomimetic networks with molecular-driven performance, their interactions serve as conduits of information flow.

BIOENGINEERED CONDUITS FOR DIRECTING DIGITIZED MOLECULAR-BASED INFORMATION

By

Jessica L. Terrell

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:
Professor William E. Bentley, Chair
Professor Gregory F. Payne
Assistant Professor Christopher Jewell
Associate Professor Edward Eisenstein
Professor Herman O. Sintim, Dean's Representative

Dedication

To my family and loved ones for their unconditional support and to my earliest mentors in scientific research, Mrs. M. Turpyn and Dr. D. Shipp.

Acknowledgements

This work would not have been possible without the instrumental exchange of ideas and recommendations by my colleagues, mentors, and committee members. The people with whom I have interacted closely in pursuit of this research willingly engaged in my work, providing inspiration and enthusiasm. I have benefited from the talented people of the Bentley lab for biomolecular and metabolic engineering and the interdisciplinary Biochip Collaborative, both past and present members. In particular, Dr. Hsuan-Chen Wu, Dr. Chen Yu Tsao, Dr. Varnika Roy, and Dr. Karen Carter were exceptional role models for me; they invested time in my instruction, guidance, and inclusion on their projects and collaborations from the beginning. I am deeply grateful for this support.

My advisor, Dr. William Bentley, has provided tremendous mentorship and professional opportunities to improve my abilities as a scientist- especially to strategize projects, to recognize profound impacts, to explore concepts beyond their boundaries, and to establish vision. I have also been grateful for the mentorship of Dr. Gregory Payne, who has provided much guidance in the development of scientific skills, in particular, in experimental organization and scientific communication. This research has allowed me to pursue international collaboration- Dr. Sheref Mansy and his lab group, especially Dr. Roberta Lentini, graciously hosted me as a visiting scientist in Trento, Italy, where I learned new perspectives and techniques toward synthetic biology.

Finally, the Fischell Department of Bioengineering at the University of Maryland has provided a rich environment- due to engaged faculty, administration, and students- to learn, grow, and to contribute to this interdisciplinary field.

Table of Contents

List of Figures

List of Tables

1.1 MOTIVATION

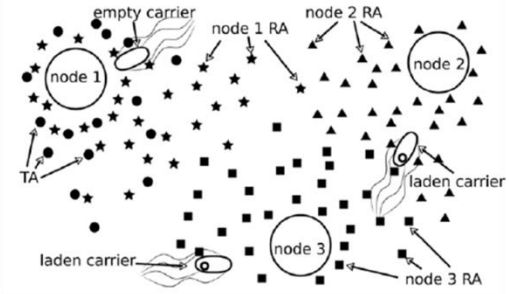
The last two decades have witnessed an advent of biosynthetic and bio-inspired research advances. Across interdisciplinary fields exists a shared vision of micro- and nano-scaled engineered systems capable of executive biological function. Diverse approaches have been explored to create 'smart' technology that features biological underpinnings: synthetic biology for the rewiring of cells [1,2], bioMEMs and microfluidic devices that simulate multicellular tissues [3], nanotechnology as automated machines [4], and materials useful for creating cell-like prototypes [5].

Central to the goal of developing 'smart' behaviors such as dynamic adaptation, operational continuity, and triggered actuation, is the necessity for components to be networked. In natural biological environments, the property of molecular recognition enables dynamic response by instilling functionality and communicating information about the system. Thus, by incorporating the same qualities into biomimetic designs, complexity can be expanded toward bio-components that are networked through the transfer of molecules.

1.1.1 Theories of biological interactions for nano-machination and computation

The conception of nano- (or micro-) machines networked through molecular communication is by in large theoretical, although some assorted demonstrations have been reported. LP Giné and IF Akyildiz (2009) recognized that bio-hybrids are strategic for the advancement of active nano-machines, considering that traditional transceiving mechanisms, which typically utilize electromagnetic signals, cannot be reduced effectively to the nano-scale [6]. Instead, molecular communication could be substituted for signal transmission, the efficacy of this approach extensively proven by natural biology at this desired length scale. Among the cited communication techniques are signaling molecules, such as pheromones, for sending

information, DNA for encoding, and fluorescent proteins for optical transduction [6]. In a more elaborate formulation, LC Cobo and IF Akyildiz described bacteria as communicative single-celled nodes, networked by the emission and response to chemotactic signals [7]. In this example, cells can act as carriers of DNA-laden data that migrate toward designated nodes, guided by a small molecule gradient (Fig. 1.1). Tangible outcomes of these models include biotechnological processes that are autonomously enacted at the molecular level, for applications such as *in vivo* maintenance of homeostasis, environmental remediation, and actuatable device-based biosensors.

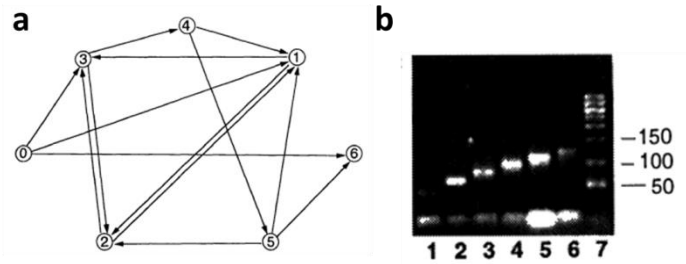


LC Cobo and IF Akyildiz. *Nano Comm Networks*, **1** (2010)

Figure 1.1 A bionetwork using nodes towards which bacterial carriers migrate. Nodes can serve as stations or destinations and bacteria are laden with genetically-stored data. This concept was described by LC Cobo and IF Akyildiz in *Nano Comm Networks* (2010) [7].

Meanwhile, molecular interactions have also been implemented for computation. Here, combinatorial solutions can be obtained to computational problems, using a physical algorithm that is dependent on a sequence of molecular interactions. LM Adleman (1994) was the first to demonstrate molecule-based computational theory [8]. He used the partial complementarity of DNA oligomers to obtain the combinatorial solution to a Hamiltonian path, or the order in which a fixed number of intricately connected states can all be accessed (Fig 1.2a). For instance, if each oligomer represents a state, then all combinations of transitioning from a specified initial state to a final state are represented by the total combination of ligated oligomers, the solution of

which can be PCRed using a forward primer for the initial state and a reverse primer for the final state. The size of the PCR product indicates how many states were included in the 'path' and the path order can be incrementally determined (Fig. 1.1b). Overall, this example represents that the recognition specificity of biological components can be used as a robust tool for the accession and interpretation of information, even that which is not biological in nature.



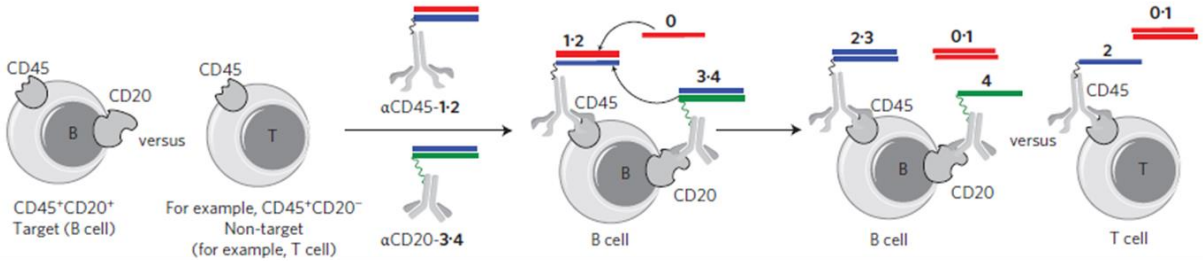
LM Adleman. *Science*, **266** (1994)

Figure 1.2 Use of non-living biological components to solve combinatorial problems. (a) A Hamiltonian path between 7 connected vertices exists such that each vertex is crossed once in numerical order. (b) Ligation products of DNA pieces that represent each vertex constitute all possible paths. The number of included vertices is reflected by the product size, and the order can be solved by analysis of each piece of the product. This work was described by LM Adleman in *Science* (1994) [8].

1.1.2 Integratable componentry for facilitating biological interactions

The pragmatic implementation of biocomponentry to address real-world challenges is emergent, especially where their integration - often in synthetic combinations- yields novel technology. Such systems need not be cell-based; in fact the construction of cell-free networks can be facile compared to synthetic cells, which must be governed by their genetic engineering. Molecular automata are abundant examples; these are molecular groupings that autonomously operate through dynamically assembling and interacting. In a recent report by M Rudchenko *et al* (2013), DNA and antibodies were conjugated to build structures with combined programmability (via DNA sequence) and specificity (via antibody-antigen affinity) features [9]. Oligo-conjugated antibodies (against cell surface receptors) were designed as molecular automata,

where partial complementarity of oligos was used to initiate DNA strand displacement cascades. This operation could be triggered by introduction of a free oligo and run in accordance with the available antibodies binding to a cell surface (Fig. 1.3); the resulting output was a unique cell surface tag representative of the cell's antigen composition [9]. The takeaway from this work is achievement of robust performance by a modular nanomachine, where each macromolecular component contributes a pre-programmed feature toward autonomous decision-making- in this case for the purpose of evaluating and classifying cell type.



M Rudchenko et al, *Nat Nano* **8** (2013)

Figure 1.3 Molecular automata as conveyors of information with biological context. Oligo-linked antibodies bind to specific cell surface proteins, and by a DNA hybridization cascade, the total receptor composition is revealed. This work was published by M Rudchenko et al in *Nature Nanotechnology* (2013) [9].

Additionally, the field of synthetic biology has been at the forefront in developing cells as engineered systems that can be controlled and guided by chemical cues [10]. Much focus has been toward the advancement of internal genetic components for switching [11], oscillation [12], and overriding state changes[13] as tunable control elements. Yet, increasingly, their application toward programming functional phenotypes has also been sought [14]. Multiple examples of programmed 'smart' cellular agents have been reported that are tasked for autonomous therapeutic assessment and intervention [2]. In work by N Saeidi et al (2011), engineered *E. coli* detected and eradicated biofilms produced by the pathogen *Pseudomonas aeruginosa* through a sense-and-destruct mechanism, where *E. coli* lysis to release a toxin against *P. aeruginosa* was

triggered by recognition of its signaling molecule [15]. In addition to fighting infection, synthetic cells have also been applied as anti-cancer treatments. HC Wu et al (2013) demonstrated that engineered cells can be preferentially targeted to and actuated at cancer cells based on biomarker recognition [16]. Specifically, an overexpressed surface protein on cancer cells was immuno-targeted for the immobilization of a "chemoattractant source" protein. In this way, cells were guided by a strong gradient that promoted their chemotaxis, and because engineered gene expression was inducible by the same molecule, cells could actuate response upon localization [16]. Thus, the application of engineered cells in processes involving their discrimination of molecular information is highly feasible.

Furthermore, the capacity of cellular machines need not be limited to what can be genetically programmed. Cells could, instead, be co-functionalized with external materials where each contribute to the overall performance of the bio-hybrid. To this point, recently reported by X Chen et al (2015) was an energy-converting engine powered by surface properties of cell spores immobilized onto micromaterials [17]. This work particularly benefitted from the integration of multiple material types because access to available molecular energy (evaporation) was enabled; through their bio-hybrid engine, the energy could be captured at the nano-scale and converted for work (object motion) at the macro-scale [17]. By coupling to biology, the ability to capitalize on molecular phenomena can either be scaled up for user control, as demonstrated, or maintained microscopically. In several instances, motile bacteria have been used as cargo-carriers, their surfaces fixed with microparticles, perhaps conjugated with a therapeutic, for targeted delivery [18], [19].

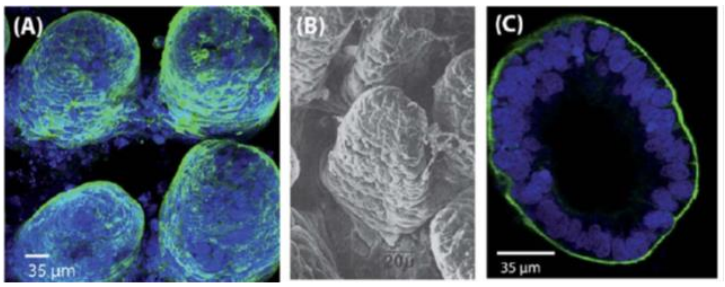
Thus, many illustrations of advanced, interactive biological systems have been demonstrated in the field of bioengineering. A common vision is autonomous computation and decision-

making at the microscale for the purpose of relaying or responding directly to obtained molecular information. Componentry may include molecules, cells, and materials, but task completion especially benefits from integrated systems.

1.2 BACKGROUND

1.2.1 Biofabrication

Biofabrication is the assembly of cells and biomolecules using biocompatible materials in a benign manner to preserve biological function [20] [21]. In particular, the use of biocompatible scaffolds has gained use for the development of synthetic tissues, either as engineered implantable tissues or for generating physiological models *in vitro* [22] [23]. Work by JH Sung et al (2011) demonstrates the utility of biofabrication towards 3D cell culture that emulates the *in vivo* environment. Here, biofabrication was used to develop a structurally accurate replica of gastrointestinal tissue, seeded with a model epithelial cell line, Caco-2 [24]. Microvilli scaffolds were molded as soft collagen-based hydrogels on which Caco-2 cells could adhere and proliferate over time to result in a confluent GI tissue monolayer (Fig. 1.4) [24]. Biofabrication techniques enabled the formation of the otherwise challenging structural features imposed by microvilli – their circumferential curvature, high aspect ratio, and necessary soft, semi-flexible length [24]. The reversible nature of many biomaterial hydrogels, such as calcium alginate in this case, allowed for a sacrificial hydrogel mold that could be dissolved to prevent damage to the collagen scaffold.



JH Sung et al , *Lab Chip* , **11** (2011)

Figure 1.4 Integration of biofabricated scaffolds and cells as an *in vivo* prototype. (A) Colocalized fluorescence, (b) scanning electron, and (c) confocal cross-sectioned image of epithelial cell adherence and confluence on biofabricated microvilli 3D scaffolds. This work was demonstrated by JH Sung et al in *Lab on a Chip* (2011) [24].

As follows, biofabrication of biocompatible environments within microfabricated and microfluidic devices can advantageously allow for micro-scale control of physiological conditions (through flow and perfusion rates, etc.) as well as *in situ* probing and monitoring. Notably, microfabrication has been well-developed as a systematic processing of inorganic materials, which has led to sophisticated devices such as bioMEMs. However, these types of devices only allow for minimal use of biological components, and particularly, in a superficial manner, not one which is fully integrated or compatible with biology. Biofabrication, on the other hand, directly addressed this challenge by developing techniques that are systematic, yet labile to incorporate biocompatible interfaces.

Chapter 2 will explore the potential of biofabrication to support molecular interaction in bio-networks that are fabricated on-chip with spatiotemporal resolution. Assembly will be implemented with modularity, where first, patterned scaffolds are established on-chip to create sites for organized, localized functionality. The particular demonstration of an integrated bio-network in Chapter 2 is representative of miniaturized bioprocessing. Many recent advances in bioprocessing have been enabled by developments in miniaturization and microfluidics. A

continuing challenge, however, is integrating multiple unit operations that require distinct spatial boundaries, especially with included labile biological components. Here, Chapter 2 suggests biofabrication as a solution, using polysaccharide films composed of chitosan and alginate, assembled on-chip by electrodeposition. These films will serve as "smart" configurable interfaces that mediate the scalable performance of a production address, where incubated cells secrete antibodies, and a capture address, where secreted antibody is retained with specificity and subsequently assayed. The antibody exchange from one electro-address to another will exemplify integrated in-film bioprocessing, facilitated by the integrated biofabrication techniques used. Overall, this work is intended to be representative of the opportunity to arrange complex combinations of diverse biological components for execution of tasks through their interaction.

1.2.2 Physical information storage in biopolymers

Concurrently with the integration of biology into fabricated systems, there is also interest in harnessing the capacity of biology to transfer information. The field of 'infobiology' has surfaced in order to tap into the potential for bio-based data transmission. Infobiology, a term coined by MA Palacios et al (2011) , is defined as data encoding, sending, and receiving through biological means, especially using living organisms as carriers of data [25]. Palacios et al represented a text cipher by matching characters to a spectrum of fluorescence-emitting cells (Fig. 1.5a). Cell colonies were arrayed, then revealed the color-coded message only in the presence of an inducing agent as the encryption 'key', or in other words, through triggered simultaneous expression of the fluorescent proteins across the array (Fig. 1.5b)[25]. A benefit from this strategy is in its simplicity; the message merely requires the physical arrangement of cell populations [26] resulting in their portrayal as alphanumeric characters, which can be visually inspected.

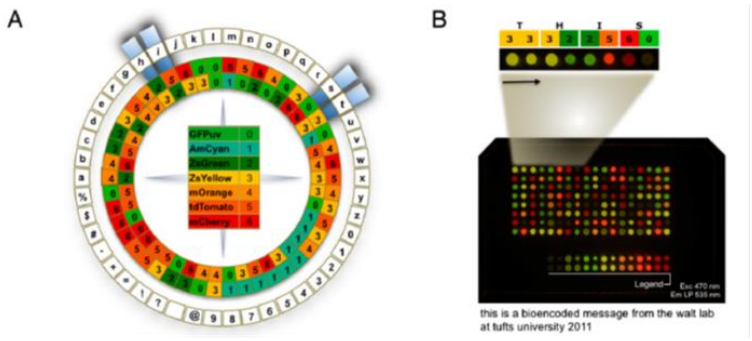


Figure 1.5 Infobiology using cells as a chromatogram of coded data. (A) Unique pairs from a 7-base color code (using cell-expressed fluorescence proteins) represent an alphanumeric character set. (b) Arrays of cells are arranged to delineate a text message, whose coded sequence is revealed upon induced protein expression. This work was demonstrated by MA Palacios et al in PNAS (2011) [25].

A different infobiology approach is to designate units of a biopolymer string to represent data, by defining and controlling the sequence order. An obvious candidate is DNA, based on the advanced synthesis and sequencing technology available for custom-design and read-through of its sequence [27]. Importantly, the input/output of these processes can be interconnected with computer-based programs for deciphering code and for further processing of the received sequence data [28]. Additionally, DNA presents a data storage option that can be preserved long-term [29] and has a greater memory capacity than cell arrays or even commercial technology- shown to be 1000-fold higher than the latter (10^{14} bits/mm^3 versus 10^{10} bits/mm^3) [30]. In this way, DNA can serve as a physical medium for the storage and maintenance of digital information. So far, DNA storage has been demonstrated either in sequences assembled de novo without any genetic relevance, or embedded within genes. The latter strategy again incorporates a cell as a carrier of the information and also provides the machinery for replicating and regenerating the DNA. Plasmid DNA, in particular, can be easily extracted, modified and reinserted. However, it is essential to retain the fidelity of the construct for proper genetic function.

In Chapter 3, we have developed a strategy of introducing subtle alterations to the DNA architecture so that it may serve a dual role for genetic function and information storage as an embedded system, yet be separated from and, importantly, selectively distinguished from a non-embedded system. Chapter 3 suggests an important strategy toward information encryption-specifically through clouding digital DNA information with presence of "empty" sequence noise, which is eliminated by user-specified clearance to yield only the stored data for decryption. Overall, the novel designation of biological components accommodates the processing of digital data by way of embedding, concealing, carrying, and ultimately exposing the data, along with its computer-automated decoding.

1.2.3 Molecular automata for robust information processing *in vitro*

Towards autonomous functionality in bio-networks, there has been an increasing effort to showcase the ability of molecular automata to processively serve executive, mechanical, and sensing roles through defining molecular interactions. For example, these functions were accomplished in work by L He et al (2010) using a DNA origami-inspired approach for the task of building an amino acid chain [31]. Synthetic DNA was designed with strategically placed annealing regions to conform a 'DNAbot' via origami. The DNAbot mimicked ribosomal peptide synthesis by "walking" along another DNA piece (serving as a template) to synthetically add amino acids to a growing chain in an assembly line fashion [31]. Here, the system included internally-programmed template DNA containing information about the amino acid order, a mechanical robot capable of self-instruction by sensing the template, and finally the autonomous performance of the task of peptide synthesis.

Chapter 4 explores a molecular automata approach toward the extracellular regulation of chemical signals as an opportunity for control of molecular input to, in turn, control cellular

response externally. The signal is first templated as preprogrammed information, then systematically transduced by an enzymatic cascade into a chemical cue of specified amplitude. Specifically, conversion to a small molecule is accomplished as a biochemical relay generated from DNA-stored code. Thus, in a similar manner to the above example by L He et al, DNA enables sequence-based information storage. The information is made accessible in molecular form through enzymatic processing. The flux of the chemical product, then, is controllable through programming the DNA sequence. The catalytic product may be transmitted as a molecular command to responder cells.

This approach works by coupling the event of DNA methylation with a fusion enzyme that generates a microbial communication molecule, autoinducer-2 (AI-2) [32]. DNA is permanently "templated" for a specific molecular dose by including methyl-accepting segments into a master sequence. The DNA is then "read" by a methyltransferase, which processively methylates each cognate site. The DNA prescript acts as a methyl-marked record of the dose, which can be accessed by endonucleases to reveal a unique digest barcode pattern. The other byproduct, S-adenosylhomocysteine (SAH), also doubles as a substrate for an AI-2 synthase. Finally, AI-2 generation drives the response of engineered bacteria. Chapter 4 will show that cellular bioluminescence output reflects the transmitted AI-2 flux. Overall, in Chapter 4, diverse biological components work synchronously to relay the encoded DNA as a specified molecular dose to tune cell output, confirmed upon execution via a DNA barcode. The root power of synthetic biology - engineering DNA – is exploited to construct an entirely new function: DNA-based programming of biochemical flux, which, in turn, is used to dynamically coordinate cell populations to meet specific demands.

1.2.4 Engineered cell populations for categorizing molecular information

Synthetic biologists have begun to recognize that engineered cells offer an opportunity, not only to be actuated by chemical cues with designer phenotypic responses, but that unique population types can be used collectively, with cooperative behavior between constituents [10] [33]. However, the task of marginalizing responses to the same chemical cue, such that unique cell outputs are obtained from a single input, is less common. One example of a successful multi-state response pattern within a consortium was demonstrated by S Basu et al (2005) [34]. In their work, reporters to an autoinducer signal were genetically tuned for triggered expression due to a precise concentration range of the autoinducer input. Excessive or dilute concentrations would lead to, in a tuned manner, inhibitory pathways, but a narrow concentration window would override the downregulatory effects for fluorescence expression. In this way, Basu and colleagues designed reporters to have "bandpass" selectivity, which was tuned for three independent cell types with varied bandpass windows. When the three cell types were plated together, centered by an autinducer plug for its radial diffusion, solitary concentric response bands were observed (Fig. 1.6). Here, the gradient created concentration contours which targeted each sub-population in order of highest threshold to lowest.

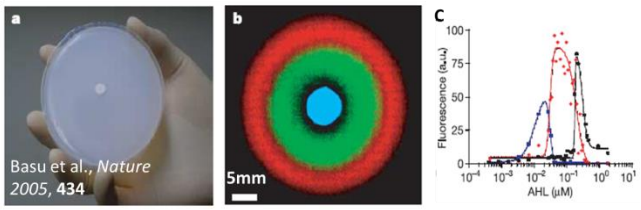


Figure 1.6 Distinguished responses of coexistant subpopulations to a single signaling molecule. (a) A plug containing the signal source (e.g. signal-producing cells) is applied to a petri dish containing a lawn of mixed reporters. (b) Demonstration of a banded fluorescence response from each of the comingled subpopulations to the radially-diffused signal. (c) Quantification of fluorescence intensity from each reporting subpopulation as a function of signal (AHL) concentration. This work was demonstrated by S Basu et al in Nature (2005) [34].

Chapter 5 explores the use of a multipartite cell population, where each subpopulation independently responds to a chemical cue, its threshold determined by their collective discrimination. A system of many unit cells, then, contributes to the relay of molecular information with context. In this way, cells provide means to sample molecular space, and in doing so, identify molecules of interest within complex environments, then actuate function. In Chapter 5, we have merged cells with nanotechnology for an integrated molecular processing network. Here, an engineered consortium autonomously generates feedback to chemical cues. Moreover, abiotic components readily assemble to cells, enabling widespread response of the dispersed system to external stimuli. Specifically, engineered populations are triggered by various threshold levels of a quorum sensing signal to express surface displayed fusions of a fluorescent marker and affinity peptide. The latter provides means for attaching magnetic nanoparticles to fluorescently-activated subpopulations for their coalescence into color-indexed output. Ultimately, quorum sensing level is assessed by probing molecular information using cells populations as a bio-litmus, conveyed as robust output via nano-coupling. By Chapter 5, we envision many such advances that will combine abiotic and biotic systems for mediating molecular communication and signal transmission.

1.3 Dissertation outline

Overall, the chapters span several relevant themes for developing interactive biological networks. The dissertation transitions from the organization of networked components, to the utilization of componentry for harboring data, a strategy for the self-directed conversion of stored data into cell-readable chemical signals, and finally, the multipopulation processing of chemical signal information to convey output.

Chapter 2: Functionalization of biofabricated scaffolds for interactive biological networks

Chapter 2 demonstrates, through the combined use of cells, biologics, and biocompatible scaffolds, integrated componentry with defined configurations to create an interactive biofunctional network capable of transmitting molecular information.

Chapter 3: Dual use of DNA for embedded data storage within a functional genetic framework

Chapter 3 develops a platform for the selective encryption and decryption of genetically-embedded digital information. By subtly changing the architecture of plasmid DNA, it can then be used as a discreet, yet distinguishable, vector of data. Moreover, a cell host can assist in the decoding process through expression of the embedded gene.

Chapter 4: Digitally-programmed information relay from DNA to a chemical signal

In Chapter 4, DNA serves as a template for biochemical synthesis as opposed to its natural function as a template for gene expression. The particular nature of this template is a permanently inscribed message (coded nucleotide sequence) that is delivered to cells as a signaling molecule through a signal synthesis process that targets the inscription.

Chapter 5: Nano-guided cell networks as conveyors of molecular communication

In Chapter 5, engineered cell populations are utilized as decentralized sensing components to probe molecular space for a signaling molecule, refine the information, and provide context based on collective output from the "living" system. By hybridization to magnetic nanoparticles as a feature of the output, the cells undergo population-wide, magnetically-triggered restructuring to assist in the feedback relay; moreover, their role is elevated beyond merely a reporting function in that the conveyed information has been discriminately categorized.

Chapter 2: Functionalization of biofabricated scaffolds for interactive biological networks.

This chapter was primarily reproduced with permission from the publication: JL Terrell et al, "Integrated biofabrication for electro-addressed in-film bioprocessing," *Biotech J,* **2012,** 7(3), 428-39, [35].

2.1 Introduction

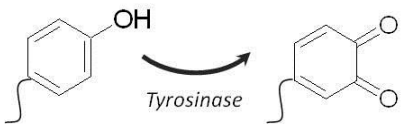
Advances in microsystems technologies have revolutionized the study of biology by allowing rapid and parallel analysis of micro and nanoliter samples; they have also enabled a reexamination of bioprocessing[3,24,36]. Virtually all bioprocess operations have been tapered to significantly smaller length scale[37,38]. Both on-chip and fluidics-based systems have appeared for a variety of tasks, many of which are now commercial products (e.g., DNA microarrays, DNA amplification, fermentation, and molecular separations) [39-45]. Systems that integrate two or more operations are less common, however, in part because the spatial arrangement of various biological components within these microsystems remains a challenge[20,23,46]. Fabrication processes must allow for spatial organization of fragile biological components while preserving functionality. In addition, systems integration concepts that ensure valving and microactuators reliably distribute various fluid components to specific locales increase complexity both at the design and fabrication stages.

We have proposed biofabrication as a strategy to address this challenge[47]. Specifically, we employ electrodeposited polysaccharide films to facilitate the spatial assembly of relevant biomolecules and cells on-chip and within fluidic microsystems[48-51]. Polysaccharide films serve as biocompatible interfaces with unique suitabilities[20]. Because they are stimuli

responsive and because the stimuli (pH, temperature) can be signaled using microelectronics, their assembly and functionality can be "programmed" with high spatial resolution[51-54]. Then, because a particular bioprocess operation of interest is associated with the polysaccharide electrodeposited at a specific site, programmed assembly of multiple bioprocess operations should be enabled. In our recent work, we assembled yeast in alginate and agarose, induced antigen-specific binding variable lymphocyte receptors (VLRs) on their outer surfaces, then demonstrated an on-chip immunoassay of the receptor-ligand binding[55]. All operations took place at a single location on a chip. In our current work, we demonstrate electroaddressing of antibody-secreting NS0 cells within alginate based hydrogels after assembling an antibody capture surface in an adjacent register. All components are surface addressed on-chip to enable rapid and simple "in-film" bioprocessing. The electroaddressability of film interfaces facilitates the proximal placement of multiple biological functionalities, conferring greater complexity in interactions by their integration. Additionally, after demonstrating the process integration, we miniaturized the system to millimeter-scaled registers while preserving function.

Central to our concept is a spatially mindful design for molecular exchange between electrode addresses. In Fig. 1A, we have envisioned a biofabricated polysaccharide-based environment for in-film bioprocessing events. Alginate has known utility as a cell-entrapping scaffold[56,57]. Here, an electrodeposited alginate film assembles and houses cells at a *production address*; metabolic products, notably monoclonal antibodies (mAbs), are secreted from this address. Adjacently, a *capture address* is biofabricated using a chitosan film functionalized with engineered protein G molecules through fusion tag tethering[49]. With protein G functionality, this address becomes receptive to the Fc portion of antibodies, and in this application, recovers those generated at the production address[49,58-60].

Fig. 1B outlines an integrated biofabrication procedure for sequential assembly of the production and capture addresses. Procedures for our work explored capture address assembly followed by production address assembly, as depicted in Fig. 1B. Our capture address is created using an aqueous solution containing chitosan pipetted onto microfabricated chips with patterned electrode pairs (one being indium tin oxide (ITO) and the other gold). Locally, at a negatively biased electrode (gold), chitosan's primary amines become deprotonated, allowing film formation at the electrode surface[52]. In our previous work, we demonstrated chitosan film conjugation by enzymatic assembly of an engineered protein G [HG3T] that is modified to contain a tyrosine rich "pro-tag"[49,61]. Enzymatic "activation" of the tag's tyrosine residues via tyrosinase, enables covalent coupling of the resultant quinone to the primary amines on deposited chitosan[62-64]. The tyrosinase activation reaction is depicted below:



HG3T is then a binder with specificity for the Fc region on freely suspended immunoglobulin G (IgG), and when coupled to chitosan, creates a capture address (Fig. 1B).

Then, shown in Fig. 1B, a solution containing NS0 cells, alginate and calcium carbonate ($CaCO_3$) is applied on-chip, where an optically-neutral positively-biased electrode (ITO) is proximal to the chitosan film. During conventional alginate gel formation, alginate chelates calcium ions to form crosslinks between chains[56]. In our strategy, the hydrolysis of water at the electrode (low pH) releases calcium ions and liberates carbon dioxide from the $CaCO_3$, yielding a cell-entrapped alginate hydrogel overlaid onto the electrode with identical dimensions[65]. The cells used here are an NS0 myeloma cell line, transfected to produce a monoclonal IgG (mAb) recognizing the vaccinia virus L1 protein. Contingent on remaining

viable, the co-deposited productive NS0 cells comprise the production address. In our previous studies, eukaryotic yeast cells were entrapped by alginate electrodeposition and remained viable[55].

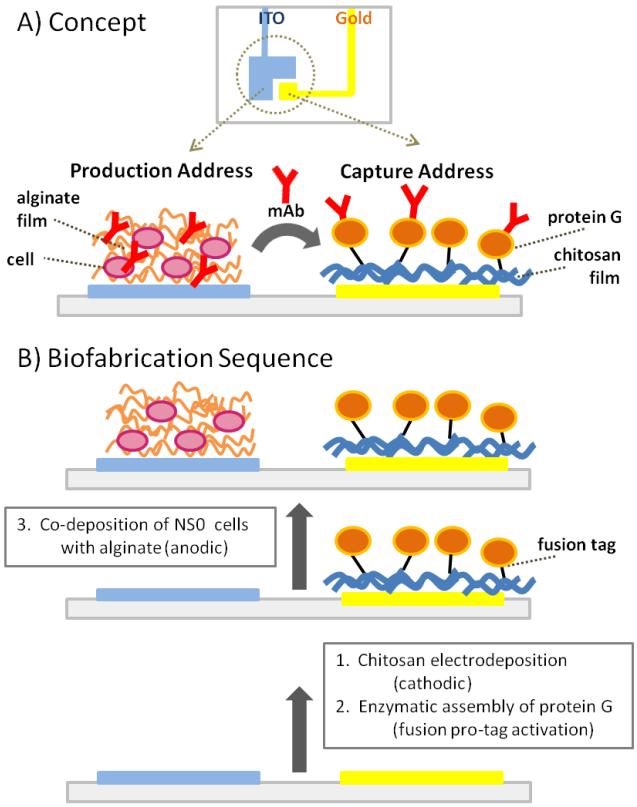


Figure 2.1 Integrated biofabrication. (A) In-film bioprocessing is facilitated by the electroaddressing of alginate and chitosan films onto adjacent ITO and gold electrodes. Film biofabrication achieves a cell-entrapped alginate film and protein G functionalized chitosan film. Once assembled at neighoboring addresses, cells secrete antibodies, which may diffuse and bind to protein G. (B) Integrated biofabrication occurs by sequential assembly. First, gold is cathodically biased to electrodeposit chitosan, followed by enzymatic assembly of protein G onto chitosan, using tyrosinase to activate a pro-tag engineered onto the protein. Then, transfectant NS0 cells are co-deposited with alginate by anodically biasing the ITO electrode in the presence of $CaCO_3$ to release calcium ions for crosslinking.

Of note, alginate hydrogel formation is reversible, subject to the presence of a calcium chelator such as citrate. That is, addition of citrate should dissolve these alginate films, presumably with minimal effect on cell metabolism[66]. Exploiting this reversible containment attribute, mAbs that were retained within the alginate film could be liberated into solution along

with cells by introducing citrate. This potentially allows for both screening production capability and subsequent recovery of assembled cells. In this manuscript, we demonstrate the creation of both addresses and their utility; we also demonstrate reversibility of cell capture for culture scale-up.

2.2 Materials and Methods

Instrumentation. A Keithley 2400 SourceMeter was used for electrodeposition and platinum foil was used as a counter electrode. Olympus imaging technologies were used: an MVX10 MacroView fluorescecence stereomicroscope for electrode and cell imaging and a DP72 camera with CellSens Standard software. Mean grey value measurements of electrode fluorescence and immunoassay signals were evaluated with ImageJ software. ELISA results were obtained with a SpectraMax M2e microplate reader and SoftMax Pro 5.3 software from Molecular Devices. FACS analyses were performed using a BD FACSCanto II flow cytometer and BioFACS Diva software for data collection by BD Biosciences.

Electrode Fabrication. Metal electrodes were coated onto a silicon wafer by thermal evaporation of 12nm of chromium (Cr) and 120nm of gold (Au). Standard photolithography and subsequent etching of Cr and Au were performed to define pairs of electrodes with 1mm separation. Each active rectangular electrode (1mm x 4mm) is connected to a contact pad via a thin contact lead (200μm in width). Electrodes were physically separated using a micro dicing saw.

To create adjacent ITO and gold electrodes (shown in Fig. 2A, 4A), photolithography and etching were performed on an ITO coated glass slide (Sigma-Aldrich) to define "L-shaped" ITO electrodes. Second, the gold electrodes were defined by additional photolithography, thermal evaporation (12nm of Cr and 120nm of Au) and lift-off. See Table 1 for electrode dimensions.

Table 2.1 Scalable electroaddress dimensions. ITO and gold electrode pairs were fabricated in three sizes, with nearly equivalent surface area ratios between the production address (on ITO) and the capture address (on gold).

Electrode Pair	Capture Area (CA) (mm^2 Gold)	Production Area (PA) (mm^2 ITO)	PA/CA
Large	9	33	3.67
Medium	4	16	4
Small	1	4	4

Solution preparation. A solution of 2% alginate was prepared by autoclaving a mixture of alginate (Sigma-Aldrich, medium viscosity) and 1% calcium carbonate (Specialty Minerals, MultiFex MM 70nm particles) suspended in deionized (DI) water to sterilize and promote dissolution, then subjecting the solution to continuous stirring. Green fluorescently labeled alginate was prepared by adding 1μL/mL of FluoroSpheres (Invitrogen, F8765) amine-modified fluorescent microspheres (1μm diameter, Ex/Em: 505/515) to the alginate solution and vortexing for 1min. The chitosan used was isolated from crab shells and at least 85% deacetylated (Sigma-Aldrich). Chitosan was prepared by dissolving at a low pH (~3) and then adjusting the pH to ~5 with NaOH. The final concentration was 1.6% and was further adjusted to 0.5% in DI water for electrodeposition. Red fluorescently labeled chitosan was prepared according to Wu et al., 2003 using 5-(and-6)-carboxyrhodamine 6G succinimidyl ester (Ex/Em: 525/560, Invitrogen)[67].

Cell culture. The cells used were provided courtesy of Darryl Sampey (Biofactura, Inc.). Briefly, an NS0 myeloma cell line had been transfected by electroporation to produce a chimeric IgG1 antibody that is monoclonal against vaccinia virus protein L1. Stably transfected cells were isolated using a proprietary metabolic selection strategy that exploits cholesterol

auxotrophy of the parental line. The stable pool of transfected cells was cloned two times by limiting dilution to isolate the high expressing subclone. Cells were incubated in Gibco CD hybridoma medium supplemented with MEM non-essential amino acids and L-glutamine (Invitrogen).

Cell Staining. A Live/Dead Viability/Cytotoxicity kit (Invitrogen Molecular Probes) was prepared according to protocol in Dulbecco's PBS (DPBS, Invitrogen) and applied to the cells for 30min. For observing stained cells in solution, the cells collected by centrifugation for 5min at 500g, resuspended in DPBS to rinse off excess staining solution, and centrifuged again prior to imaging. For staining cells entrapped within an alginate film, the film was briefly rinsed with DPBS prior to imaging. Cell counting by ImageJ analysis was performed on fluorescence images taken separately with FITC and TRITC filter sets and merged for presentation.

HG3T assembly and antibody capture. HG3T expression and purification in *E. coli* BL21 (DE3) has been described previously[49]. To prepare the chip surface for antibody capture, each electrode was electrodeposited using a solution of 0.4% chitosan in DI water. A cathodic bias of $3A/m^2$ was applied to each gold electrode for 2 minutes followed by a rinse in DI water. To enzymatically assemble HG3T, the electrodes were incubated overnight at 4°C with 0.4µM HG3T and 50U/mL tyrosinase from mushroom (Sigma-Aldrich). Next, 0.2% sodium cyanoborohydride (Sigma-Aldrich) was applied for 15min, followed by 1% sodium tri-poly phosphate (TPP) for 15min. Electrodes were blocked in 5% dry non-fat milk in PBS for 4h, then incubated with a standard dilution in 1% milk-PBS + 0.05% Tween or a sample for 2h. As a secondary label, FITC-conjugated AffiniPure $F(ab')_2$ fragment goat anti-human IgG (Jackson ImmunoResearch Laboratories) was applied in 1% milk-TPBS for 2h. Washing steps were

performed in between. For each electrode comparison, identical magnification and fluorescence exposure times were used.

Cell co-deposition and in-film bioprocessing at the production address. Cells were taken from culture and collected by centrifugation at 500g for 5min, resuspended in DPBS to remove superfluous extracellular mAbs, and centrifuged again. Samples were resuspended in DPBS at $20x10^6$ cells/mL and then an equal volume of 2% alginate and 1% $CaCO_3$ was mixed in to achieve a final concentration of 1% alginate, 0.5% $CaCO_3$, and $10x10^6$ cells/mL. Approximately 75, 150, or 250μL of the mixture was applied to each respective small, medium, or large electrode pair, and the cells were allowed to settle on the chip surface for 5min. An anodic bias of $3A/m^2$ was applied to each ITO electrode for 2 minutes followed by gentle rinsing of each film with 0.145M NaCl. 10mM $CaCl_2$ was briefly applied to the films for crosslink densification with subsequent NaCl rinses. Culturing media supplemented with 10μg/mL gentamicin (Invitrogen) and 500μM $CaCl_2$was added for incubation and samples were covered with Breathe-Easy film (Diversified Biotech). After incubating for approximately 5h at 37°C, production address films were dissolved by applying minimal volumes of 0.2M sodium citrate and gentle pipetting. Equal volumes of DPBS were added to improve isotonicity and the samples were centrifuged at 500g for 5min to recover suspended cells. Supernatants, presumably containing secreted mAbs, were applied to corresponding capture addresses for 2h, followed by additional blocking at 4°C to minimize non-specific binding, and finally application of a labeling antibody (see HG3T assembly and antibody capture methods) for fluorescence imaging. Variations of the antibody capture step were reported in our results.

ELISA Standard curve. Standards of comparison were prepared between on-chip and conventional ELISA techniques. Human IgG (Sigma-Aldrich) was prepared in serial dilutions

up to 5500ng/mL. An ELISA was performed using an anti-human IgG alkaline phosphatase conjugate (Sigma-Aldrich) as a secondary antibody and PNPP (Thermo Scientific) as the enzymatic substrate for detection.

2.3 Results and Discussion

2.3.1 Adjacent electroaddressability of chitosan and alginate films

In Fig. 1, we hypothesized that chitosan and alginate could be electrodeposited sequentially and adjacently with minimal crosstalk or interference between electrode addresses. In Fig. 2A, a bare electrode pair is depicted to demonstrate neighboring gold and ITO geometries. First, a red fluorescently labeled chitosan solution (0.5%) was electrodeposited on a gold electrode (Fig. 2B). Excess solution was rinsed from the film, and a 1% sodium tri-poly phosphate (TPP) solution was applied to electrostatically reinforce the chitosan film. Second, an alginate (green-fluorescently labeled, see Methods) and $CaCO_3$ mixture was applied and electrodeposited onto a neighboring ITO electrode. Afterward, the films were rinsed and post-deposition calcium chloride ($CaCl_2$) solution was briefly applied. A fluorescence image was taken of the resulting electrodes, showing that both films were intact (Fig. 2B). A control image is also shown in Fig. 2C where we imaged a pair of electrodes through fluorescence filters after an alginate deposition procedure in which $CaCO_3$ was absent. In this case, we demonstrate negligible background fluorescence and an inability of alginate to electrodeposit without the presence of $CaCO_3$ (Fig. 2C), thereby indicating its necessity. Also of note, this assembly is evaluated for crosstalk or electrostatic interactions that occur between chitosan and alginate due to their respective net positive and negative charges[68,69]. Direct application of dissolved alginate on a chitosan film could potentially result in the formation of an electrostatically bound alginate layer on the film's surface[70]. Our application of TPP prior to alginate served to avert such an interaction.

Because the overlaid fluorescence image did not show colocalization of the green and red films (Fig. 2B), these electrostatic interactions were minimized.

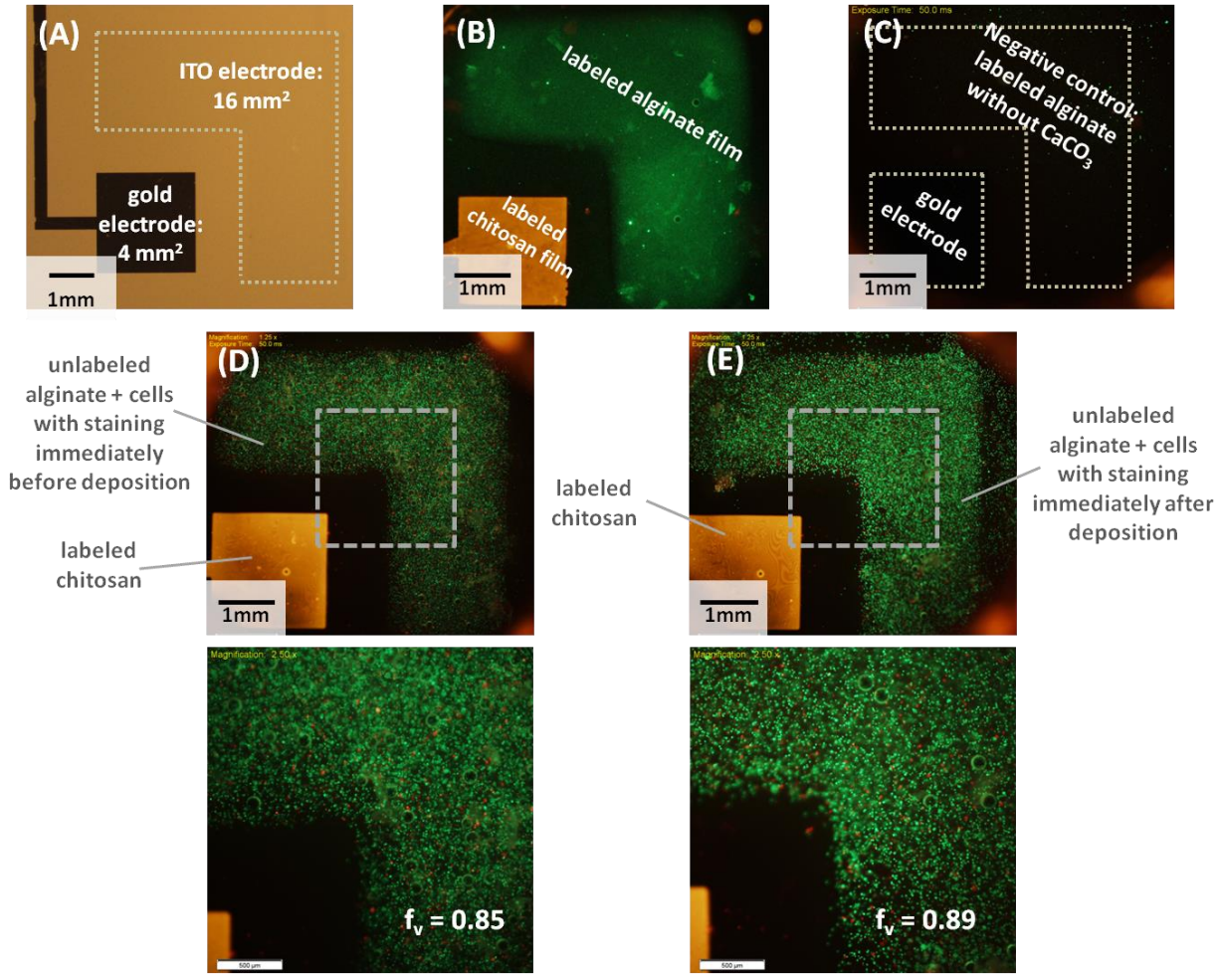


Figure 2.2 Electroaddressability of sequentially deposited and cell-entrapped films. (A) Brightfield image of patterned electrodes prior to film electrodeposition. (B) Fluorescence image of red-labeled chitosan deposited on gold followed by green-labeled alginate deposited on ITO by localized calcium release from $CaCO_3$. (C) Fluorescence image showing a gold electrode with the absence of a chitosan film and a negligible presence of labeled alginate when deposited without $CaCO_3$. (D) Cells were stained for viability (green=live; red=dead) *before* codepositing on ITO with alginate. The viable fraction (f_v) is representative of the cell culture viability, 0.85. Cells remained immobilized throughout washing steps. (E) Cells were stained for viability *after* codepositing with alginate; the f_v is 0.89.

Next, we monitored cell electroaddressing during the deposition process. In Fig. 2D, chitosan was deposited as mentioned. Cells were stained with a viability/cytotoxicity stain (Live/Dead, Invitrogen); providing visualization and viability assessment. These stained cells were then

rinsed, centrifuged and resuspended to a density of $10x10^6$ cells/mL in a solution of 1% alginate and 0.5% $CaCO_3$ in DPBS. Cells were then introduced onto the chips and co-deposited on ITO. The fluorescence image (Fig. 2D) shows that cells were successfully electroaddressed to the ITO electrode. Additionally, the cell population used, stained upon sampling from culture, appeared healthy (live cells fluoresce green while dead cells are red). By image analysis, green and red fluorescence was quantified to determine the fraction viable (f_v=0.85).

On a test ITO electrode, cells were again co-deposited with alginate as described, but were probed for viability after deposition. An identical cell sample was suspended at $10x10^6$ cells/mL in a deposition solution. Alginate deposition was performed as described, followed by application of the Live/Dead stain. The fluorescence image (Fig. 2E) again shows an electrodeposited healthy cell population (f_v = 0.89). Fig. 2E has significance in that it demonstrates co-deposition of NS0 cells without negative impact on cell viability; quantified viabilities between Fig. 2D and 2E are nearly the same (0.85 and 0.89, respectively).

2.3.2 Rapid Monitoring of mAb Production from NS0 cells

Our next set of studies demonstrates rapid and simple on-chip entrapment of cells. Moreover, because the process is simple and rapid we hypothesize these procedures may enable interrogation of cell productivity over time for a full scale production process.

2.3.2-1 Production address characterization and scalable assembly

We were aware that many parameters may affect the electrodeposited cell density in-film and sought to determine electrodeposition conditions that would ensure uniform and reproducible populations. While many factors were varied, we report here the electrodeposited cell number as a function of cell seeding density, cell "settling" time (duration between loading cell sample and biasing the electrode), film volume, and the electrode surface area. Results are depicted in Fig.

3A. For each parameter tested, data were obtained by first co-depositing cells with alginate on each electrode keeping all other conditions constant (conditions, except when varied: $10x10^6$cells/mL seeding, 5 min settling time, $4mm^2$ electrodes, $3A/mm^2$ current density for 2min. to yield films 0.38mm in thickness). To count cells, films were fixed with paraformaldehyde (2%), citrate-dissolved, then washed, resuspended, and counted via FACS. Using the electrode area and the estimated film thickness (~0.38 mm, generated at $3A/mm^2$ for 2min; ~0.52 mm, $5A/mm^2$, 2 min) [65], we estimated an in-film volumetric cell density.

The data demonstrates that "settling" time impacts in-film density; if there was no "settling" time (or an on-chip incubation period), the resultant cell density was ~ $3x10^6$ cells/mL, much lower than the initial seeding density ($10x10^6$ cells/mL). Beyond 5 min the deposited cell density was typically $7x10^6$ cells/mL. Indeed, there were many factors that influenced the deposited cell number: applied voltage and current density, cell number in the suspended sample, duration of the electrodeposition, etc. In addition to the "settling" time, we tested initial seeding density (Fig. 3A). Here, the deposited cell number was a strong function of initial seeding density, as one might expect. In seeding experiments with initial concentrations below $3x10^6$ cells/mL, we found the resultant densities were typically concentrated by a factor of nearly 2. At the highest seed level ($10x10^6$ cell/mL), the opposite was observed (~30% decrease). For our purposes, a seeding density of $10x10^6$ cells/mL and 5 min settling time were selected as the most convenient for creating sufficiently consistent production films. These values were used in all subsequent experiments.

Lastly, in Fig. 3A, we depict the in-film density as a function of the electrode area and deposited film thickness. These were of significant interest as a key focus of this work is its performance as the entire process and chip are shrunk in all geometric dimensions. That is, we

were interested as to whether the density of entrapped cells at a production address was consistent over a variety of defined electrode areas and film thicknesses. Table 1 describes the geometries of ITO and gold electrode pairs designated for production and capture addresses, respectively. By electrode patterning, the electrode dimensions were, per pair, intrinsically set to scale (see PA/CA, Table 1). The calculated in-film cell density is shown as a function of production area size in Fig. 3A. We observed that regardless of electrode size (assuming same height/width ratios), the resulting in-film densities were nearly equivalent, with no statistical difference. Also, regardless of the electrodeposited film thickness (0.38 vs. 0.52mm) the in-film density was constant. Our results demonstrate that adjusting deposition parameters provides some degree of control over the cellular in-film composition. We envision biofabrication at a yet smaller scale - electrodeposition on micron-sized electrodes[71]. In-film seeding, subject to the choice of parameters described here, could yield single cell entrapment for clone separation capabilities.

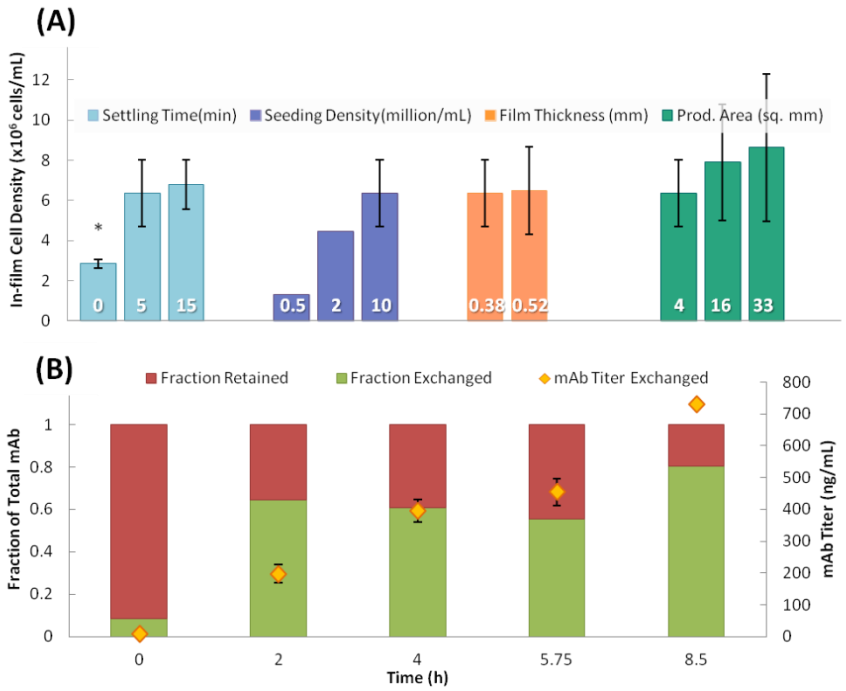


Figure 2.3 Characterization of the NS0 cell composition of a production address. (A) Electrodeposition parameters were varied to examine their effect on the resulting in-film cell density. Variation of the parameters are specified at the base of each graph and plotted against the in-film density, counted by FACS. Each variable was evaluated in at least duplicate. (*) indicates $p < 0.05$. (B) Cells were incubated in-film for time intervals ranging from 0 to 8.5h in order to measure mAb production over time. Titers were evaluated by the fraction detected within the alginate film and the fraction diffused, or "exchanged" from the film at regular time intervals. Titers were quantified by ELISA; the exchanged titer is reported in (B) in context of the fraction that it represents of total mAb produced per time.

2.3.2-2 Productivity Performance: Antibody production and release from alginate hydrogels

Electrodeposited NS0 cells were then examined for productivity when entrapped at $10x10^6$ cells/mL; mAb titers were recorded as well as their distribution (within gel or secreted from gel and in the supernatant). Factors that may contribute to this distribution include the production rate and whether it is sustained [57], the crosslink density of the hydrogel network[55], and the charge density of the mAbs relative to the hydrogel (both negatively charged)[56]. Cells were codeposited into alginate films and allowed to incubate with media for 8.5 hr; samples from parallel experiments were taken regularly by separately removing the extra-film media (~200µL per test), then dissolving the film with sodium citrate (200µL) and removing cell debris to collect the intra-film solution. The mAb levels were assayed via ELISA. In Fig. 3B, the total titer in the extra-film fluid is depicted as well as the fraction in-film and extra-film. In as little as 2 h, over 60% of the mAb was found in the extra-film fluids. This fraction was fairly constant until 8.5 h, when the bulk (>80%) of the mAb had been diffused away. Importantly, the secreted antibody titer increased uniformly over the entire experiment, reaching 700 ng/mL at 8.5h. These results indicate that mAbs generated and secreted at the production address may be available for capture within 2 h. As such, NS0 cells can be entrapped at a production address preserving their viability and productivity. – Also, the films produced are sufficiently diffuse so that produced mAbs perfuse through the film. Cells were retained within the film and are maintained in a mAb-

producing state. These results, we believe, hold promise for several applications as a means to immobilize cells and permit interactions with large molecules by diffusion into or out of the film. In addition to clonal productivity analysis, this platform could enable spatially controlled cell arrays to study metabolomics, biopsy screenings, drug targeting, or cell signaling.

2.3.2-3 In-film bioprocessing utility: Incubation of entrapped NS0 cells and recovery

Since we have observed preserved cell viability after co-deposition in an alginate film, we hypothesized that the film may serve as an adequate environment for subsequent cell incubation. Furthermore, since alginate films comprise reversible crosslinks and we have measured exceedingly high cell densities within the films, we also expected feasible cell recovery. Fig. 4A shows the smallest electrode pair having dimensions of a $1mm^2$ gold electrode (capture address) and a $4mm^2$ ITO electrode (production address) used for a cell incubation and recovery study. We monitored cell survivability after co-depostion with alginate, but especially after the cells were incubated for several hours at a high density. Cell samples concentrated to $10x10^6$ cells/mL in 1% alginate with 0.5% $CaCO_3$. After deposition, the Live/Dead stain was applied as described. The films were imaged by fluorescence microscopy and show significant viability (Fig. 4B, $f_v = 0.76$).

Equivalent cell-entrapped films were incubated in culture media at 37°C for 4h. After incubation, they were treated with 0.2M sodium citrate to dissolve the films and recover the cells. Figure 4C shows cells stained for viability after incubation within a $4mm^2$ alginate film on-chip, citrate dissolution, centrifugation and resuspension steps for recovery. The majority of cells imaged were alive ($f_v = 0.78$), suggesting that the electrodeposited alginate film provides a favorable cellular microenvironment. Other samples were rinsed and recollected by centrifugation, then incubated in a 96-well plate. Fig. 4D shows an image of cells recovered

from a single 4mm^2 ITO electrode and transferred to a single well for an additional 24h incubation after residing in the alginate film. In addition to preserving their viability (f_v = 0.63), these cells recovered from 4mm^2 addresses were eventually regrown to near confluency. These results suggest potential for subsequent scale up to larger cultures despite their electrodeposition and minimal sample size.

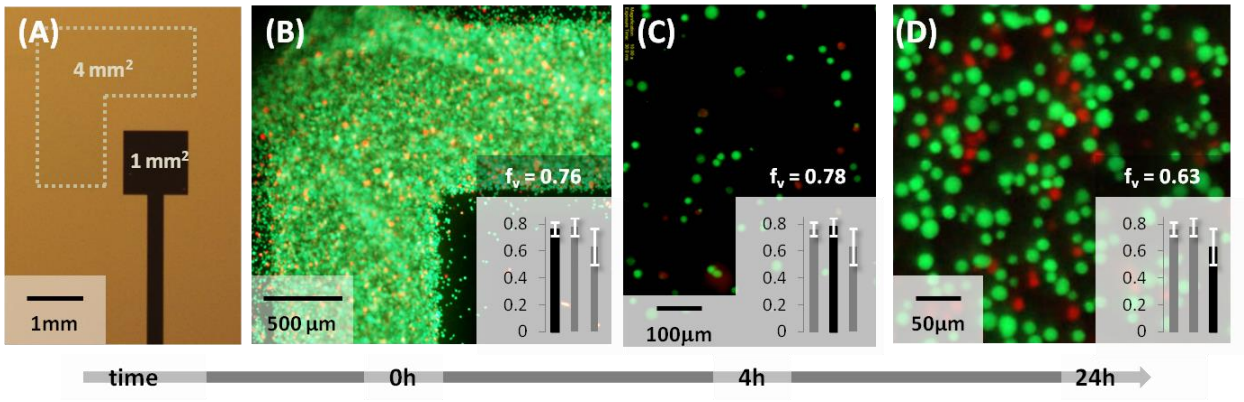


Figure 2.4 Cell viability during on-chip incubation and recovery. Codeposited cells were tested in triplicate for survival upon incubation in-film and recovery by film dissolution at the smallest electrode. (A) Electrode brightfield image. Cells stained for viability after (B) electrodeposition (f_v = 0.76), (C) 4h in-film incubation and recovery (citrate dissolution, f_v = 0.78), and (D) in-film incubation, recovery, and 24h incubation in a 96-well plate (f_v = 0.63). Each inset graph plots the viable fractions (f_v) and standard deviations for images B, C, and D, the associated data point highlighted in black.

2.3.3 HG3T functionalized chitosan films as receptive surfaces for antibody capture

In final studies, we integrated the assembly process: a production address was assembled adjacent to a capture address, placing antibody-secreting cells in proximity to antibody-receptive HG3T, thereby allowing an opportunity for mAb exchange. We evaluated capture surfaces for detectable response upon exposure to varying antibody concentrations, then their capture reliability directly from a cell sample, scalability, and semi-quantitative correlation.

2.3.3-1 Capture address characterization by fluorescence analysis

In order to ascertain whether the co-addressed cell-entrapped hydrogel would enable release and capture of antibodies (over the time span of the incubation period and after release), we first

established that area-based fluorescence measurements were correlated with standard ELISA measurements. That is, human IgG at known concentrations were prepared and assayed via alkaline phosphatase based ELISA (see Methods). For on-chip analysis, these same solutions were used as substrates for electroassembled chitosan and enzymatically assembled HG3T capture. That is, a chitosan film was first deposited on gold, and then functionalized by applying HG3T and tyrosinase (see Methods). Captured IgG was labeled, after incubation of the IgG solutions on-chip, with a fluorescently-conjugated $F(ab')_2$ secondary antibody fragment for fluorescence detection. Fluorescence images (Fig. 5A) were scanned using grey scale intensity measurements by examining equivalent areas of each electrode. In Fig. 5A, it was readily apparent that a monotonic increase in fluorescence with antibody titer was observed. This demonstrates success of several factors required for quantitative assessment. First, it demonstrates that the chitosan deposition and HG3T assembly processes were uniform within a particular electrode surface and, in results not depicted, repeated processing using parallel electrodes was equivalent, suggesting a robust and reproducible process.

The quantified results for each surface are depicted in Fig. 5B(1), and corresponding ELISA measurements in Fig. 5B(2). The ELISA standard curve was linear below 1000ng/mL ($r^2 = 0.99$). On-chip detection of the same solutions was also linear over the same range ($r^2 = 0.99$). These results suggest that antibody capture using covalently assembled HG3T and detection on-chip is feasible because differences in antibody concentration can be distinguished among samples. These results also suggest the mAb capture efficiency for the on-chip method was equivalent to ELISA methods. Finally, we selected these concentrations for study as they were estimates for mAb secretion on-chip. That is, our fluorescence standard was later used to gauge the amount of mAb directly transferred from an on-chip cell population.

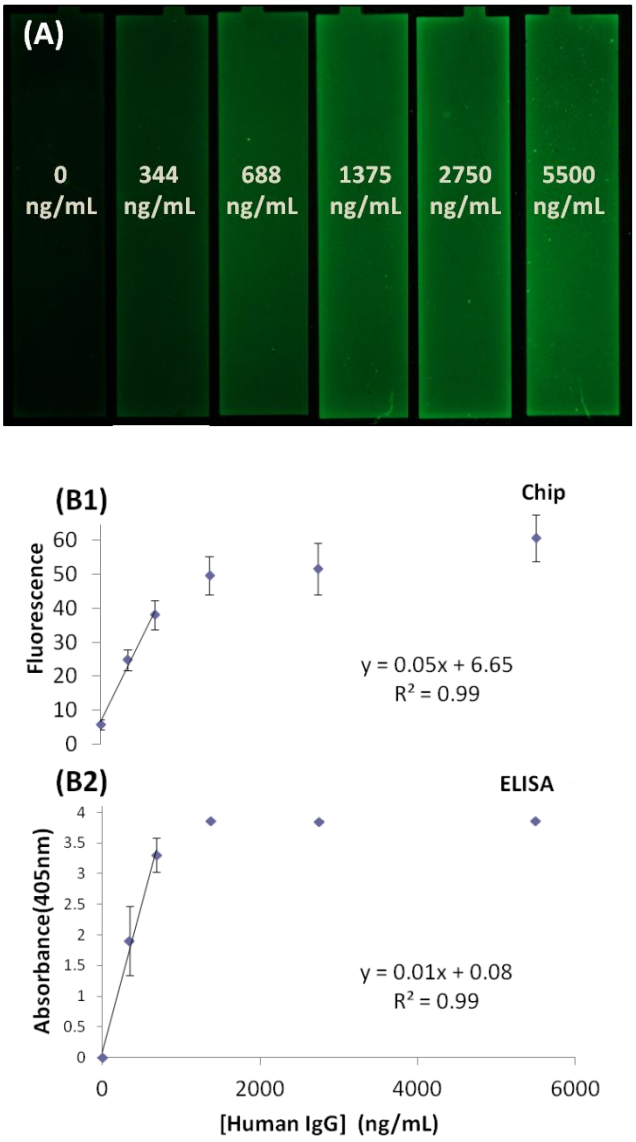


Figure 2.5 Fluorescence detection of captured antibody correlated to ELISA. Known concentrations of human IgG were detected by protein G-conjugated chitosan films and ELISA. (A) Fluorescence image of films after secondary FITC-labeling with an $F(ab')_2$ antibody. (B1) Area-based fluorescence analysis of (A), using electrode duplicates. (B2) Alkaline-phosphatase based ELISA analysis, measured in triplicate. (B1,2) show linear curves below 1000ng/mL.

2.3.3-2 Receptive Performance: Antibody exchange across addresses

Transfer of mAb produced by cells incubated at the production address directly adjacent to the capture electrode was then evaluated. In this study we examined another attribute to the system that was an unanticipated benefit. Namely, the deposition of NS0 cells were suspected to

increase productivity due to an artifact of increased cell density. Cell samples were taken, concentrated to $10x10^6$ cells/mL, and co-deposited with alginate as described above. ITO electrode addresses were each paired with gold electrodes preassembled with an HG3T-functionalized chitosan film. That is, in these studies, we first created the chitosan capture surface and then assembled the production film. After deposition, the electroaddressed cell-entrapped films were incubated at 37°C for 5.5h (see Methods). After incubation, supernatant media were removed and hydrogels were dissolved using sodium citrate. We subsequently removed the supernatants, centrifuged the cells, and reapplied the clarified supernatant to the chips. These supernatants were incubated for 2h for capture. Thus, the entire capture exposure time was 7.5h, including the cell incubation time. Fig. 6A(2-4) shows fluorescence images at capture addresses of each size (referring to Table 1). The fluorescence of the small, medium, and large electrodes were all significantly greater than that of a negative control (Fig 5A(1), normalized intensities of electrodes in A2-4 are represented by the inset graph in A1). Thus, we have demonstrated integrated in-film bioprocessing, where cells were entrapped and incubated and mAbs were secreted, collected, captured, and detected, all on-site, and all at distinctive millimeter-scaled addresses.

2.3.3-3 Scalable relationship between addresses

We completed identical capture studies on differently sized chips. We found that for the medium system, comprising a $4mm^2$ capture area and a $16mm^2$ production address, and over 5 hours incubation, sufficiently significant mAb levels were produced, captured, and imaged. We then kept nearly the same ratio of production/capture areas (Table 1) and shrunk the entire system to contain a $1mm^2$ capture address and $4mm^2$ production address. By performing studies using identical processes and times, we show in Fig. 6B(1-2) that the level of fluorescence in the small

system was nearly equivalent to that of the large system (Fig. 6B inset graph). Our results are overall consistent with the notion that the biofabricated surface construction and use is both simple and robust, but also scalable.

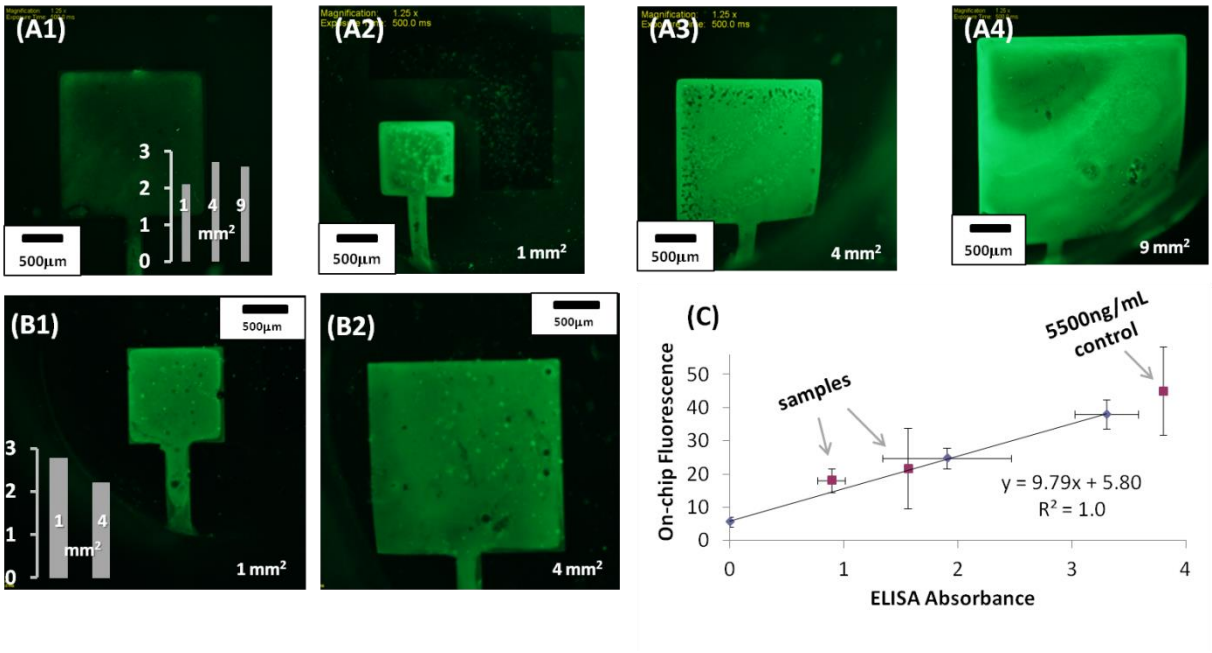


Figure 2.6 Scalable and quantifiable antibody exchange from production address to capture address. (A1) A negative control shows capture address fluorescence due to non-specific binding of the secondary FITC-conjugated $F(ab')_2$ antibody. (A2-4) Capture addresses secondarily labeled after exposure to mAbs generated at the production address during incubation. After area-based fluorescence measurements, inset graph in (A1) has been normalized to the negative control, showing fluorescence from samples as greater than two-fold above the control. (B1-2) Identical samples were fluorescently probed for mAbs at the corresponding capture addresses (B1- small, B2- medium), with equally sized negative controls (not shown). Inset graph shows nearly equivalent fluorescence measurements normalized to the controls. (C) From Fig. 5(B1-2), the on-chip standard curve was plotted against that of ELISA, and detection data of two samples (mAbs diffused from production addresses) and positive control (5500ng/mL) were semi-quantitatively measured.

2.3.3-4 In-film bioprocessing utility: Semi-quantification of antibody titer

Another test probed the extra-film supernatant for released mAb during cell culture on-chip at varied incubation periods. After incubating for 2h and 4h, we proceeded with our on-chip detection strategy. We performed grey scale quantification and also measured the supernatant

via ELISA. The grey scale quantification of the capture address' fluorescence, plotted against corresponding ELISA data, placed the antibody loading near 300ng/mL (Fig. 6C). Importantly, the quantities were both in line with the standard curve obtained via *in vitro* studies. In order to verify consistency in our on-chip detection strategy, we also remeasured a positive control of 5500ng/mL to check for correlation to standard data and found that our control closely matched (Fig. 6C). We have demonstrated success in gauging antibody presence on-chip, and have quantified its concentration.

2.4 Concluding Remarks

This work demonstrates for the first time that biofabrication strategies can be utilized in an integrated way to electroaddress multiple polysaccharide-based films for unique yet cooperative bioprocessing functions at separate yet contiguous electrode addresses. Upon cell entrapment on-chip and recovery, the functionality of an entrapped antibody-producing cell population was confirmed by on-chip detection of the secreted antibody; this is coincident with the receptive functionality maintained by HG3T after surface conjugation to capture an antibody, a metabolic product. The integration of a hydrogel-based film for non-permanent containment of a cellular component and a protein-tethered functional surface imparts new complexity in the construction of dynamic on-chip components for miniaturized bioprocesses. We suggest this approach may find the most direct utility in screening clonal populations, monitoring production runs for productivity, and for the capture of products, including other proteins, using the same technique. Furthermore, cell electrodeposition and receptor or enzyme-functionalized "biofabricated" surfaces may streamline other studies including targeting for drug delivery or signaling studies between differentiated cell populations. Ultimately, integrated biofabrication strategies may advance efforts to heighten the complexity of *in vitro* models.

Chapter 3: Dual use of DNA for embedded data storage within a functional genetic framework.

3.1 Introduction

Digital data storage and transmission has revolutionized the global availability of information by way of internetworked computers. Emerging currently is the novel concept of digital data storage in DNA [27]. This has come about, in part because DNA is widely recognized as the preferred medium by Nature for the longterm storage and passage of genetic information; in another respect, the rapidly advancing DNA synthesis and sequencing technology allows for inexpensive services to read and write DNA with unprecedented speed. Both factors have led to a sparked interest in the field of digital data storage, where innovative media with ever-higher storage capacity and smaller footprints are continually sought. The most striking realization of this mission was performed by G Church et al (2012), where 5.3×10^6 bits (encompassing the data associated with a published book [72]) were physically stored via nucleotide equivalents in DNA[30]. In doing so, the custom-synthesis of 8.7 Mb DNA (coded to include 5.3megabits of data as well as addressing and recognition sequences) demonstrated that the storage density of DNA at 5.5 petabits/mm^3 vastly surpasses that of commercially available digital storage media [30]. Their efficient use of storage space within DNA sequences further enables the technology by minimizing bit representation to a single nucleotide while additionally incorporating data processing sequence keys. These barcodes, in addition to hybridizing primers for data accession, function to organize each piece of data within the master dataset, such that data storage need not be limited to a physically linear order or even to a single DNA chain.

While other achievements of DNA-data storage have not been demonstrated to the same extent, they have progressively introduced added capabilities. Historically, PC Wong et al (2003) were among the first to conceptualize and implement the representation of a nucleic acid sequence for non-genetic code. Their work matched codons to a set of characters for the purpose of building a gene sequence ordered according to the intended message, in this case, the lyrics to "It's a Small World" [73]. From here, DNA sequencing analysis results in decryption of the message. This inspired an artist, Christian Bök, to develop a codon-based character cipher; from this he designed a transcribable coded poem, denoted Xenotext, that could be covalently fused to a fluorescent protein marker, such that it could be genetically archived within and physically 'written' by a cell (resulting in its observation through fluorescence tracking) [74]. His vision was to convey language through living means, harnessing cells as vehicles for expressing the ciphered information as unique protein structures. While closely collaborating and communicating with the scientific community to carry out the Xenotext Experiment, his visionary endeavors aimed for more of a cultural impact than a scientific one.

Since then, technologically-focused efforts have emerged to exploit DNA encryption of digital code for data security purposes [75]. The principle of stenography, or concealing one type of information within another, was applied to DNA by CT Clelland et al (1999), where an encrypted DNA segment was camouflaged amongst genomic DNA and preserved in a printed microdot of the mixed DNA, recovered by PCR targeting [76]. More recently, a watermarking approach has been realized, where data is stenographically stored, not only amongst other impertinent nucleotides, but is located within a gene. That is, the genetic code can be exploited to simultaneously harbor a protein transcript along with non-genetic information through the designated use of synonymous codons. M. Liss et al (2012) devised a cipher that represented

binary code according to the choice between alternative codons for a gene [77]. *Gfp* was demonstrated to maintain green fluorescent functionality upon watermarking.

In our work, we seek to integrate several properties of digital data storage in DNA- data encryption and watermarking- while adding a new security feature of data concealment. We also employ cells as information carriers whereby their expression of the protein alerts the genetic location of stored information, even while the code identity remains ambivalent. A purposeful link between the genetic and non-genetic aspects of a watermarked gene within a cell carrier has not been previously shown. For selective accession of the plasmid-stored embedded gene sequence, subtle architectural changes (mutational deviations from GATC sites) were designed into the vector such that its overall functionality is preserved, but can nevertheless be selectively distinguished. The plasmid is paired with an unaltered version such that they are inherently mixed within the cell; both vectors hold the gene for an identical protein, which is data-empty on the latter vector and data-embedded on the former. Enzymatic action selectively affects the plasmid architecture to access the data-embedded gene for sequencing. Finally, a computer program automates decryption from genetic sequence, its codon designation, ASCII binary code, and bit conversion to the embedded text message.

3.2 Materials and Methods

Materials and instrumentation. Polymerase chain reaction kits (Q5, Phusion, and Vent), restriction endonucleases (NcoI-HF, HindIII-HF, PluTI, EcoRV-HF, MluI, SphI-HF, and Sau3AI), T4 phosphonucleotide kinase, T4 ligase, and DNA ladders (100 bp, 2log, 1 kb) were purchased through New England Biolabs. DNA and plasmid purification kits (ZR BAC DNA mini-prep, Clean & Concentrator-5, and gel extraction) were purchased through Zymo Research.

Oligomers and Gblocks were purchased through Integrated DNA Technologies. Antibiotics were purchased through Sigma. Software included SerialCloner (Softonic) for plasmid maps and virtual cut patterns, Matlab (MathWorks) for compiling and running a decryption program, and ImageJ (National Institutes of Health) for analyzing DNA band intensity. Addtionally, a BD FACSCantoII cytometer and BioFACS Diva software was used for flow cytometry.

GATC mutations 1-2. The first mutation of pTrc-K^r (GATC1 to AATC) was generated by plasmid PCR using AATCMut1-F and AATCMut1-R as primers. The linearized plasmidΔ1 product was phosphorylated using T4 polynucleotide kinase (PNK) and then ligated with T4 ligase and transformed into NEB5α cells. The second mutation was introduced by sequential PCR of the plasmid between SphI and MluI restriction sites, first using GACCMut2-SphI-F and GACCMut2-R1 to change GATC2 to GACC, then extending the DNA fragment by another PCR with GACCMut2-SphI-F and GACCMut2-MluI-R2. The DNA fragment was inserted into pTrc-K^rΔ1 by a SphI/MluI double digest, ligation, and transformation into NEB5α.

Origin-targeted mutations 3-6. Four GATC sites within the origin of replication of pTrc-K^rΔ2 were simultaneously mutated by plasmid PCR using OriMut3to5-R and OriMut6-F. The linear product was phosphorylated, ligated, and transformed into NEB5α cells.

β-lactamase-targeted mutations 1-6 (7-12 on final plasmid). A series of silent mutations, three at a time, were introduced into the β-lactamase gene (*bla*) of pTrc-A^r. First, the plasmid was PCRed with BlaMut7to9-F and BlaMut7to9-R, phosphorylated, ligated, and transformed into Top10. Isolated pTrc-A^rΔ3 was used as a template for plasmid PCR using BlaMut10to12-F and BlaMut10-12-R. Again, phosphorylation, ligation, and Top10 transformation resulted in pTrc-A^rΔ6.

Two-step mutations of 13-15 and swapping of antibiotic resistance gene. New mutations were introduced to unique sites in pTrc-K^rΔ6 and of pTrc-A^rΔ6, and the mutated regions were then combined into a single plasmid by Gibson assembly. The gene *bla*Δ*6* was PCRed in two steps, first using Bla-F and BlaMut13-R1 to change a single remaining GATC site to AATC, then PCRing the product with Bla-F and BlaMut13-R2. The resulting *bla*Δ*7* contained 19bp on the 5' end that overlapped the sequence upstream of *nptII* for kanamycin resistance, and 36bp at the 3' end that overlap the downstream sequence of *nptII*. pTrc-K^rΔ6 was sequentially PCRed with OriMut14to15-F and BlaAssem-R, which hybridized to the sequences flanking *nptII* and introduced two GATC mutations downstream of *nptII*. The PCR product was extended to achieve 5' and 3' overlapping ends with *bla*Δ*7* by PCR with BlaAssem-F and BlaAssem-R. Gibson assembly circularized the linear pTrc-K^rΔ8 by inserting *bla*Δ*7* in place of *nptII*. Top10 transformants containing the resulting pTrcΔ15(A^r) were selected by plating on LB-agar with 100 μg/mL ampicillin.

Substitution mutations 16-17 in LacI

LacIMut16-MluI-F and LacIMut17-EcoRV-R were used to PCR a DNA fragment of *lacI* between MluI and EcoRV restriction sites and simultaneously introduce two GATC mutations within *lacI*. MluI/EcoRV double digestion removed the non-mutated section of *lacI* on pTrcΔ15, followed by ligation of the mutated insert. Top10 transformants yielded pTrcΔ17.

Generating GATCfree plasmid through final mutations 18-21. A GATC18 mutation to GTTC was introduced downstream of *lacI* between PluTI and NcoI restriction sites. The fragment was PCRed with GTTC18-F1 and GTTC18-NcoI-R to introduce the mutation, then the product was PCRed with GTTC18-PluTI-F2 and GTTC18-NcoI-R. Insertion onto pTrcΔ17 was accomplished after PluTI/NcoI double digestion and ligation. Plasmid pTrcΔ18 was isolated

from Top10 transformants, then plasmid PCRed with GATCfree21-HindIII-F and GTTC18-NcoI-R, which hybridized at the bounds of the multi-cloning site and, other than NcoI and HindIII, excluded all other available restriction sites and three GATC sites. The PCR product was phosphorylated, ligated, and transformed into Top10 to yield pGATCfree.

Table 3.1 Site-directed mutagenesis primers for the systematic deletion of GATC sites from pTrc vector. Boxed bold capital letters indicate mutations introduced to GATC sites.Boxed lower case letters indicate a restriction endonuclease site. * Indicates phosphorylated primers used for plasmid PCR. Underlines indicate complementarity to another primer. Unboxed lower case letters indicate complementarity to the plasmid for Gibson assembly, where applicable.

Oligo name	Sequence (5′—3′)*
AATCMut1-R*	GCGAATT**AATC**TGCTGCCTCGCG
AATCMut1-F*	GCGCGAAGGCGAAGCGG
GACCMut2-R1	CGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAGTT**GGTC**GGCGCGAGATTTAATCGC
GACCMut2-MluI-R2	TATATTacgcgtTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGAC
GACCMut2-SphI-F	TATATTgcatgcATTTACGTTGACACCATCGAATGGTGCAAAACC
OriMut3to5-R*	CGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAG**AATC**TCAAGAA**GAAC**CTTT**AATC**TT TTCTACGG
OriMut6-F*	CTACCAGCGGTGGTTTGTTTGCCG**GTTC**AAGAGCTACCAACTCTTTTTCCGAAG
BlaMut7to9-R*	CACTCGTGCACCCAACT**GGTC**TTCAGCATCTTT
BlaMut7to9-F*	GGTTACATCGAACTG**GACC**TCAACAGCGGTAA**GATA**CTTGAGAG
BlaMut10to12-R*	CATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCC**TATC**GTTGTCAG
BlaMut10to12-F*	GGG**GACC**ATGTAACTCGCCTT**GACC**GTTGGGAACC
Bla-F	atattgaaaaaggaagagtATGAGTATTCAACATTTCCGTGTCG
BlaMut13-R1	GTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGC**AATC**TGTCTATTTC
BlaMut13-R2*	taaagtatatatgagtaaacttggtctgacagttacCAATGCTTAATCAGTGAGGCAC
Bla-R	TAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGC
OriMut14to15-F	GATTGATTTAAAACTTCATTTTTAATTTAAAAG**GCTC**TAGGTGAA**GGTC**CTTT
BlaAssem-R	ACTCTTCCTTTTTCAATATTATTGAAG
BlaAssem-F	GGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTC
LacIMut16-MluI-F	TATTAAacgcgtCAGTGGGCT**GTTC**ATTAACTATCC
LacIMut17-EcoRV-R	AATTgatatcCGCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCT**GAAC**GTTGGC
GTTC18-F1	GCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCGCGAATT**GTTC**TGGTTTG
GTTC18-PluTI-F2	ATAAggcgccCAATACGCAAACCGCCTCTCCCCGCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCG
GTTC18-NcoI-R (*)	CTAGTccatggTTTATTCCTCCTTATTTAATCGATACATTAATATATACCT
GATCfree21-HindIII-F*	TATCTTAAGCTTGGCTGTTTTGGCGGATGAGAGAAG

Observing the presence of targeted mutations. Site-directed mutagenesis was observed comparing PCR products against equivalents treated with the restriction endonuclease Sau3AI, which cleaves at the recognition sequence GATC. Thus, GATC mutations are left uncut while intact sequences are cleaved. PCR reactions (25 μL in volume) were completed using Vent

DNA polymerase, then split in half. Half of the reaction was treated with 2.5 units of Sau3AI at 37 °C for 2 h before comparing against the untreated half via polyacrylamide gel.

FACS analysis of eGFP insertion. A green fluorescent protein gene, *egfp*, was modified with a 5' NcoI recognition site and 3' HindIII site. The gene was cloned into pTrcA^r, pTrcA^{r}Δ1, pTrcA^{r}Δ2, and pTrcA^{r}Δ17 to test for the impact of GATC mutations on the leaky expression of eGFP. Cells were grown overnight at 37 °C and 250 rpm shaking. Aliquots were then diluted four-fold in PBS and analyzed by fluorescence microscopy and flow cytometry (FACS). For FACS, a 488 nm laser and 530/30 filter set provided excitation and measurement of green fluorescence level.

Investigation of plasmid copy number. NEB5α cells transformed with either pTrcK^r or pTrcK^{r}Δ6 were inoculated from independent colonies into cultures with 50 μg/mL kanamycin and grown to an optical density of approximately 1 at 30 °C, 250 rpm shaking. Chloramphenicol (34 μg/mL) was added to a subset of triplicate cultures and incubation was continued for another 3 h. Subsequently, all samples were pelleted and ventilated for drying, after which the cell pellets were weighed. Plasmids were isolated from pellets by standard prep protocol and quantified using a NanoDrop1000 spectrophotometer (ThermoScientific); yields from untransformed NEB5α cells were subtracted from the results. Copy number was calculated based on the wet weight estimation of 1 x 10^9 *E. coli* cells per mg.

Investigation toward growth inhibition by ampicillin. Overnight cultures of Top10 cells transformed with either pTrcA^r, pTrcA^{r}Δ3, or pTrcA^{r}Δ6 were reinoculated at 1 % into LB + 500 μg/mL ampicillin and incubated in triplicate at 30 °C, 250 rpm shaking. Aliquots were taken at time intervals for absorbance measurements at 600 nm. The optical densities were averaged and reported over time with standard deviations.

EGFPtxt assignment. Watermarked *egfp* DNA was purchased through IDT. The codon sequence for *egfp* was designated based on the binary code for the embedded message, beginning with the first codon after the ATG start. Codons for methionine, cysteine, and tryptophan were not used as well as isoleucine (designated ATA by default) and aspartate codons that were followed by a C-initiated codon (these were designated GAC by default). Standard ASCII 7-bit binary code was followed for character assignment along with codon definitions in Fig. 3.7a. The embedded message was terminated with the ETX (end of text) character 0000011, which required the addition of two extra stop codons into the gene sequence to accommodate this last character. Thus, the message WEB@UMD #infobio#GATCfree48 was embedded throughout the entire gene.

Plasmid pair extraction and decryption. Top10 + pTrc-K^r-EGFP (pEGFPdec) was made chemically competent by standard methods with calcium chloride, transformed with pGATCfree-EGFPtxt (pEGFPtxt), and cultured with both kanamycin and ampicillin. The plasmid pair was extracted through standard prep protocol. Sau3AI was used to digest the plasmid mixture, and was done so at a ratio of 0.83 ng with 30 U Sau3AI overnight at 37 °C. *NptII*, *bla*, and *egfptxt* were PCR-amplified for 30 cycles for 28 pg of the plasmid mixtures as a template (-/+ Sau3AI treatment).

3.3 Results and Discussion

3.3.1 Mapping of GATC mutagenesis targets on plasmids

In order to build pGATCfree, we targeted GATC groupings on the plasmid pTrc. Site-directed mutagenesis was implemented in parallel on two similar plasmids: pTrcK^r, containing the kanamycin-resistance gene N-phosphotransferase II (*nptII*), and pTrcA^r, containing the

ampicillin resistance gene β-lactamase (*bla*). On pTrcK^r, mutations were first introduced downstream of and at the origin of replication (*oriR*, Fig. 3.1); meanwhile, on pTrcA^r, silent mutations were introduced to *bla* (Fig. 3.1). Subsequently, the mutation-containing regions of the plasmids were combined onto a single pTrc sequence while simultaneously introducing additional mutations between *oriR* and *bla* to create pTrcΔ15 with six remaining GATC pairs. Final mutations were introduced into *lacI* and the remaining GATCs were removed from the multi-cloning site (MCS), leaving a pair of restriction sites for cloning.

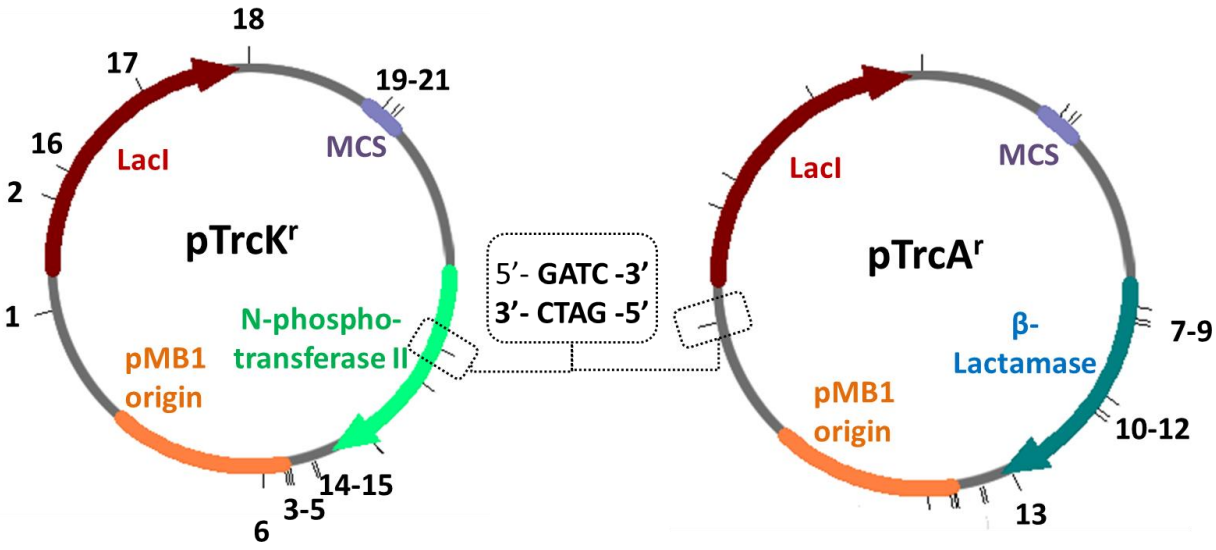


Figure 3.1 Plasmid maps of functional features and GATC sites.

3.3.2 Investigation of ΔGATC impact within the origin of replication

The alteration of GATC sites can be verified through the use of a GATC-specific restriction endonuclease such as Sau3AI. Sau3AI, derived from *Staphylococcus aureus* [78], recognizes 5'-GATC -3', cleaving across the site to create a 4 bp overhang. Unlike other GATC-specific enzymes, Sau3AI does not depend on the methylation state of the site [79]; therefore, Sau3AI cleaves any GATC-containing dsDNA alike, including methylated plasmids isolated from cells and unmethylated PCR products. However, should a GATC site be altered to a different

sequence, it becomes protected from Sau3AI activity and left intact. Therefore GATC-targeted mutations can be verified by first honing in on the mutated region of DNA through PCR, then Sau3AI-digesting the PCR product and analyzing the resulting band pattern.

First, pTrcK^{r} was mutated downstream of and throughout the origin of replication, resulting in six mutation sites, of which the mutation order is outlined in Fig. 3.2a. For pTrcK^{r}, pTrcK^{r}Δ2, and pTrcK^{r}Δ6, the presence of the intended mutations was evaluated by PCR using BlaAssem-F and GACCMut2-MluI-R2, followed by Sau3AI digest. The band patterns of the digestion reactions were resolved on a polyacrylamide gel and are shown in Fig. 3.2b. First, Fig. 3.2b shows a virtual digest on the original pTrcK^{r} sequence- the ~2000 bp region bounded by the primers- compared to a variation containing either the intended 2 or 6 total mutations. As observed by virtual cutting, an increase in the mutation number results in a reduced number of cut fragments due to mutation of the cut site, leading to larger fragment sizes that approach the size of the PCR product. From the adjacent experimental results, the band patterns of DNA treated with Sau3AI closely matches that of the virtual cut: for pTrcK^{r}, three distinct bands around and below 1200 bp are present, a major band closer to 1500 bp for pTrcK^{r}Δ1 and Δ2, and finally a large band near 1800 bp for pTrcK^{r}Δ6. Bands below 100 bp are present in the experimental results too, although difficult to resolve. The Sau3AI^{-} lanes contain PCR products that were not digested; the presence of single bands of uniform length (2000 bp) in all lanes verifies that the correct sequence is compared across lanes for all pTrc variations.

Having verified the presence of six pMB1-targeted mutations in pTrcK^{r}Δ6, we investigated the plasmid copy number carried by cells + pTrcK^{r}Δ6 in order to determine functional consequences to the plasmid's self-replication ability. Because the sequence structure of the pMB1 origin is critical for plasmid replication, we did not know whether altering its four GATC

sites would influence the resulting plasmid copy number. The pMB1 origin is similar to that of pColE1, whose two highly conserved GATC sites (denoted α and β) were previously studied for the effect of sequence, secondary structure, and hemi-methylation [80,81]. PK Patnaik et al (1990) showed that while α GATC and its downstream sequence, taken together with β GATC and its upstream sequence, constitute an inverted repeat capable of forming a putative cruciform structure, non-complementary mutations of α and β GATC, which would destabilize the cruciform, did not affect the plasmid's maintenance in a cell host. Our four mutations introduced on pTrcK^r similarly disregarded the sequence's tendency for cruciform formation; instead, GATC alterations were chosen in a way that would maximize hybridization of the triple mutation-containing primer.

Cells were grown with either pTrcK^r or pTrcK^{r}Δ6 to early stationary phase, at which point the cells were either harvested or treated with 34 μg/mL chloramphenicol (Cam). Cam treatment for cells with pMB1 vectors has been shown to augment plasmid yield because while Cam inhibits protein synthesis and protein-dependent processes such as chromosomal replication and cell division, RNAI/II-based plasmid replication is unaffected [82]. Consequently, plasmid replication becomes the sole anabolic process, which in turn, increases plasmid content relative to protein and chromosomal content per cell [82].

Fig. 3.2c shows the copy number calculated for biological triplicates of pTrcK^r and pTrcK^{r}Δ6-transformed cells without and with Cam treatment. The copy number for pMB1-based plasmids such as pTrcK^r is typically about 15-20 copies per cell. However, our lower-than-expected experimental copy number for pTrcK^r of less than 10 either indicates that culturing conditions did not promote optimal plasmid replication or that an incomplete yield was isolated from cell samples. Regardless, identical treatment of the experimental cell cultures resulted in a

statistically higher copy number of 22 for pTrcK^{r}Δ6. Thus, it seems that the origin-targeted mutations increase the copy number, perhaps by disrupting the coordination of plasmid replication with cell division. After Cam treatment, copy number was augmented to a statistically indistinguishable average of 40 for pTrcK^r and 47 for pTrcK^{r}Δ6. Copy number under Cam treatment could have been limited by the availability of free nucleotides, leading to nearly similar upper-limit levels of plasmid replication. Overall, by focusing the cell metabolism on only plasmid replication and observing a consistently increased copy number, we confirmed that the GATC mutations were inconsequential.

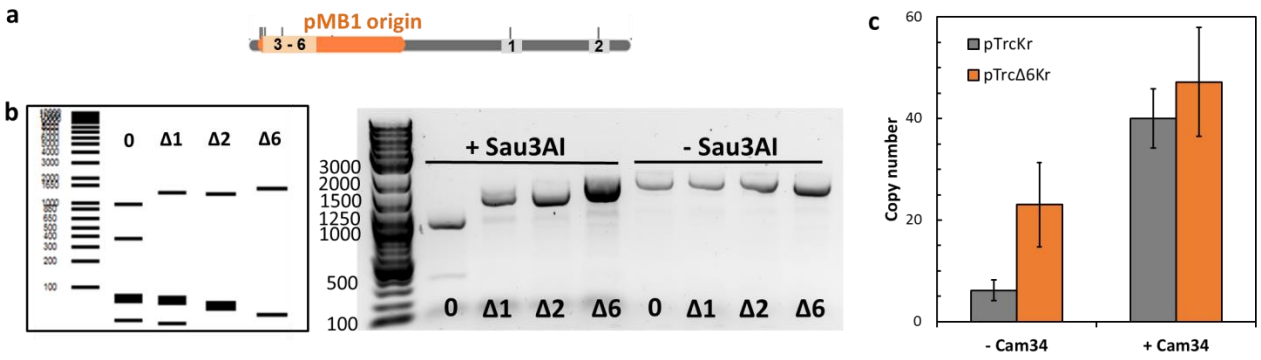


Figure 3.2 Investigating the effects of pMB1 origin mutations on plasmid copy number. (a) pMB1 map with GATC sites numbered in order of mutagenesis. (b) GATC-targeted cutting by Sau3AI on the pMB1 segment with 0 – 6 mutations. A virtual digest (left) and polyacrylamide-resolved digest (right) are shown +/- Sau3AI. (c) Copy number of plasmid hosted by NEB5α cells -/+ chloramphenicol (Cam) with pTrcK^r and pTrcK^{r}Δ6.

3.3.3 Investigation of antibiotic resistance due to ΔGATC-introduced rare codons within *β-lactamase*

Next, silent mutations were introduced into *bla* of pTrcA^r in order to maintain the property of ampicillin resistance. Seven GATC sites exist within *bla*, six of which were altered in groups of three. Within the gene, GATC occurs either as the GAT codon for aspartate (D) adjacent to a C-initiated codon, or as ATC for isoleucine (I) preceded by a G-terminated codon. To mutate these

sites silently, GAC was substituted for GAT (Fig. 3.3a), which occurs with slightly less prevalence in *E. coli* [83]. Likewise, ATA was substituted for ATC (Fig. 3.3a), noting that tRNA for ATA is significantly rarer than for ATC [84].

To verify the presence of three mutations in pTrcA^{r}Δ3 and six in pTrcA^{r}Δ6, *bla* was PCRed and treated with Sau3AI, as was done previously for pMB1. A virtual cut reference gel and experimental results are shown in Fig. 3.3b, compared to pTrcA^r. Again, the band patterns of the virtual gel closely corroborate the experimental results, where increased GATC mutation frequency has the effect of lengthening the DNA fragments due to cleavage protection. The uncut PCR products in the Sau3AI$^-$ lanes are of the same length at ~ 880 bp, which verifies parity of the compared DNA.

Sau3AI digestion confirmed the presence of six mutations within *bla* introduced in groups of three, two of which substitute common codons, and the other a rare codon substitution. We next investigated whether the inclusion of alternate codons, including rare ones, affect the property of ampicillin resistance by potentially impeding β-lactamase synthesis. We hypothesized that should β-lactamase production be inhibited, the cells' growth rate under high-ampicillin conditions would be reduced due to suboptimal availability of β-lactamase for ampicillin hydrolysis. Cells transformed with either pTrcA^r, pTrcA^{r}Δ3, or pTrcA^{r}Δ6 were grown in LB medium supplemented with 500 μg/mL ampicillin, which is 10-fold more concentrated than typically used. The optical densities (OD at 600 nm) of triplicate cultures were measured over a period of 24 h at 30 °C incubation. Growth curves are overlaid in Fig. 3.3c for cultures with each plasmid as well as a negative control culture having no ampicillin-resistance plasmid. The negative control did not grow, verifying antibiotic strength of the medium. By contrast, cultures with pTrcA^r, pTrcA^{r}Δ3, and pTrcA^{r}Δ6 all reached an OD of nearly 3, with nearly equal growth

rates throughout exponential phase. Therefore, the impact of silent codon substitutions in *bla* was negligible for antibiotic resistance at extreme ampicillin levels. Here, antibiotic resistance is a critical property for plasmid maintenance, and so this data provided confidence in the ability of the plasmid to be maintained in ampicillin conditions.

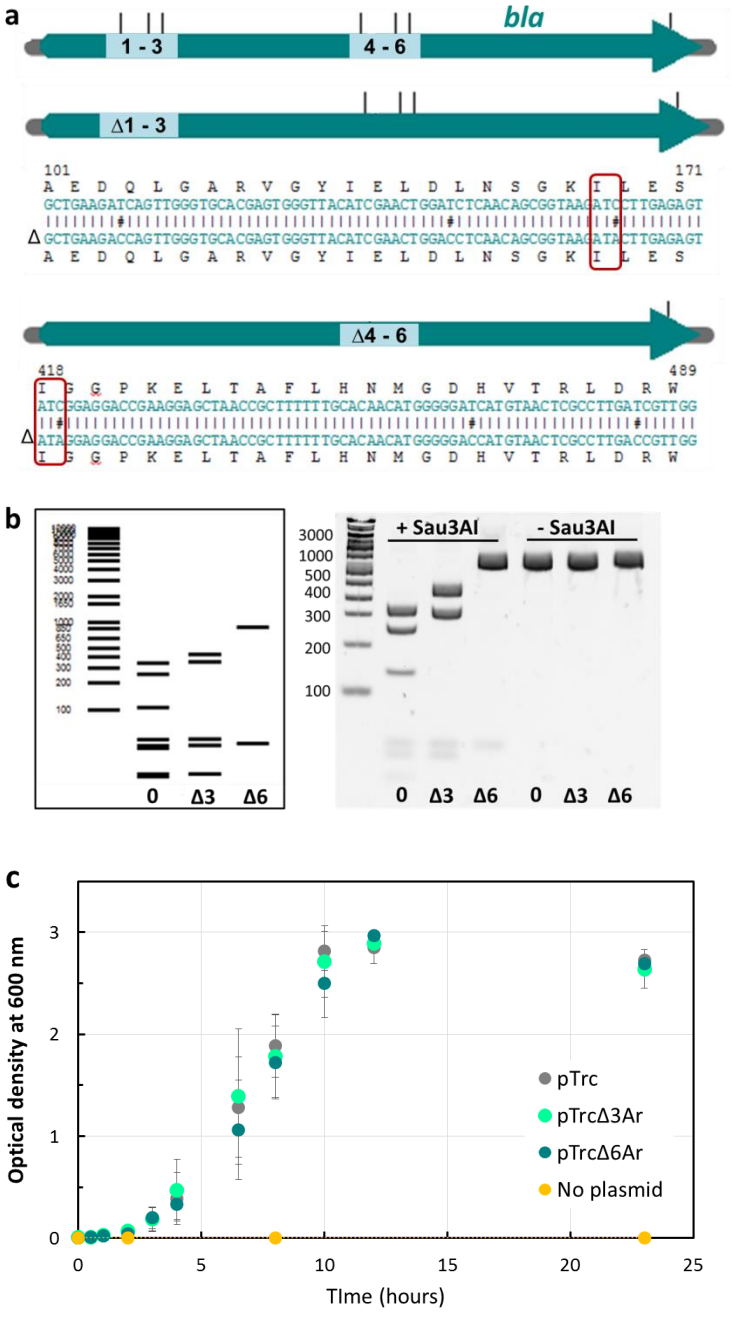


Figure 3.3 Investigating the effect of rare silent mutations in *bla* on ampicillin resistance. (a) *β-lactamase* map with GATC sites numbered in order of mutagenesis. Also shown are the mutated sequences for each target region (1-3 and 4-6) aligned with the original *bla* sequence and translation to amino acids. # indicates a point mutation, which is shown to be silent where the amino acid remains unchanged. Red circles indicate the introduction of a rare codon (abundance ratio < 0.1 relative to alternative codon options). (b) GATC-targeted cutting by Sau3AI on the *bla* segment with 0 – 6 mutations. A virtual digest (left) and polyacrylamide-resolved digest (right) are shown +/- Sau3AI. (c) Growth curves for Top10 cells in stringent antibiotic conditions (500 μg/mL ampicillin) with pTrcA^r, pTrcA^{r}Δ3, pTrcA^{r}Δ6, or no plasmid.

3.3.4 Efficient assembly of GATC-free segments onto a single plasmid

By generating sequential mutations on multiple plasmids (pTrcK^r and pTrcA^r) in parallel, mutagenesis was streamlined for efficient assembly of the mutated regions. The plasmid maps of each revealed two GATCs just upstream of the pMB1 origin on pTrcK^{r}Δ6 and a single GATC site near the 3' end of *bla* on pTrcA^{r}Δ6. The strategy of switching *nptII* on pTrcK^{r}Δ6 for *bla* (Fig. 3.4) was made such that a successful swap could be efficiently screened by plating colonies on LB-agar with ampicillin. The antibiotic resistance genes were swapped by Gibson Assembly, where each DNA fragment was first PCRed using primers that incorporated matching overlap sequences. Because the noted GATC sites were located within these priming regions, we made use of a simultaneous mutation step where the GATC sites were knocked out during PCR amplification of the DNA fragments (Fig. 3.4). Thus, once assembled, the plasmid retained 6 mutations from each fragment with the addition of 3 new ΔGATC sites, rendering pTrcΔ15.

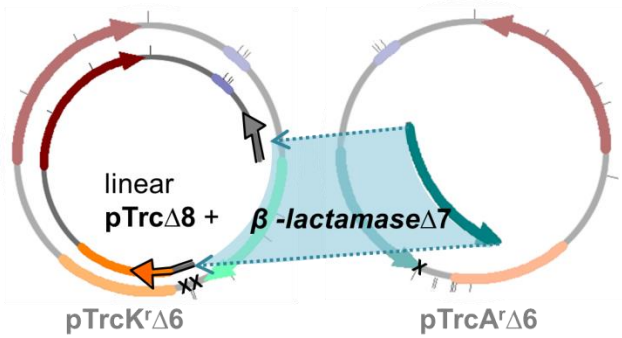


Figure 3.4 Simultaneous triple-site directed mutagenesis and swapping of antibiotic-resistance genes. Mutations targeted by PCR are indicated with an X. Resultingly, the PCR products, pTrcΔ8(*nptII*$^-$) and *β-lactamase*Δ7, were assembled into a composite plasmid.

3.3.5 Investigation of gene expression due to ΔGATC-introduced amino acid substitutions in LacI

Next, mutations were introduced into *lacI*. The GATC2 mutation added to the earlier construct created a silent codon change. Two additional GATC sites within *lacI* remained. We were interested in testing whether non-silent substitutions at these sites would alter LacI functionality. LacI is a tretrameric repressor of genes under the *trc* promoter in the pTrc construct by binding to a LacO operator site within the promoter. However, LacI is known to exhibit leaky action, resulting in high levels of *trc*-promoted protein expression without induction. We arbitrarily introduced the following amino acid substitutions: I123 for F and D247 for V, resulting in a gene for mLacI (I123F-D247V) in plasmid pTrcΔ17. To test the effect of mLacI on *trc* expression, we inserted *egfp* into the multi-cloning site of pTrc, pTrcΔ1, pTrcΔ2, and pTrcΔ17, whereby Δ1 and Δ2 were included as extra controls (ΔGATC1 was upstream of *lacI* and ΔGATC2 introduced a silent mutation into *lacI* for D). Top10 cell cultures transformed with variations of pTrc-EGFP were grown to stationary phase at 30 °C and analyzed for leaky (uninduced) expression of eGFP by fluorescence microscopy and flow cytometry, shown in Fig. 3.5. Fluorescence images using a FITC filter set captured cellular green emissions. The fluorescence distribution of each population is depicted by FACS and corroborates the visual evidence. That is, the eGFP profiles for pTrc-EGFP, pTrcΔ1-EGFP, and pTrcΔ2-EGFP–transformed cells overlay identically, indicating no shift in fluorescence distribution and no noticeable difference in the fluorescence intensity of the population. Fluorescence for pTrcΔ17-EGFP, however, is a wider distribution, peaking at a lower intensity level than that of the other cell types and also overlapping with the

non-fluorescent negative control peak. Furthermore, cells in the fluorescence image for Δ17 look dimmer, despite being imaged with the same exposure time and excitation strength of the laser. Therefore, evidence indicates that the mutations introduced to mLacI give the protein a stronger attenuation effect. While silent substitutions could have been made instead to retain the original LacI functionality, downstream attenuation applications may benefit from the use of mLacI instead.

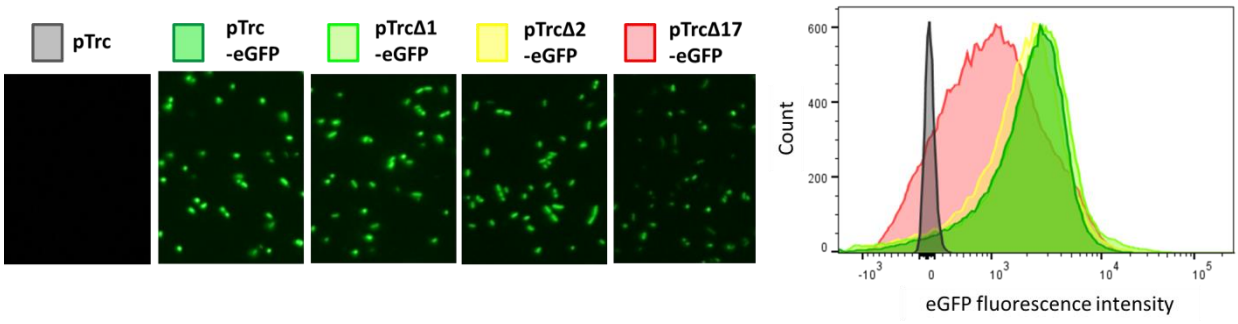


Figure 3.5 Consequences of LacI mutations on eGFP fluorescence profile of a cell population. Green fluorescence images are shown for leaky eGFP expression from Top10 cells transformed with each listed plasmid. Histogram plot depicts the relative intensity distribution of fluorescent cells in each corresponding population.

3.3.6 ΔGATC accumulation in plasmid series toward pGATCfree

Final GATC sites were either mutated or deleted- specifically, those in between NcoI and HindIII outer cut sites within the plasmid's multi-cloning region. The resulting plasmid was absent of 21 GATC pairs, replaced instead with a variety of substituting sequences (Fig. 3.6a), all of which were changed from GATC by one of the four bases (as opposed to changing or deleting the entire site). Since all mutations were targeted by mutagenesis primers, the single base substitutions promoted the highest annealing strength of the primer to the template by minimizing the extent of sequence incompatibility. In some cases, primers harbored up to 3 mutations (Table 3.1), yet all successfully altered the GATC sites, rendering the final plasmid to

be "GATC-free." The plasmid series with increasing mutation number is demonstrated by Fig. 3.6b, where all variants were treated with Sau3AI and resolved by gel electrophoresis. From the original plasmids, pTrcK^r and pTrcA^r, to the final construct pTrcΔ21, increased resistance to Sau3AI digest is observed as increased length and lack of fragmentation on the gel. Moreover, the results clearly show that pTrcΔ21 differs structurally from its original form, pTrcK^r, in the presence of Sau3AI, yet its sequence is 97 % similar; thus the effect can be interpreted as subtle by sequence analysis and drastic by enzyme action. The final construct was renamed pGATCfree for downstream applications.

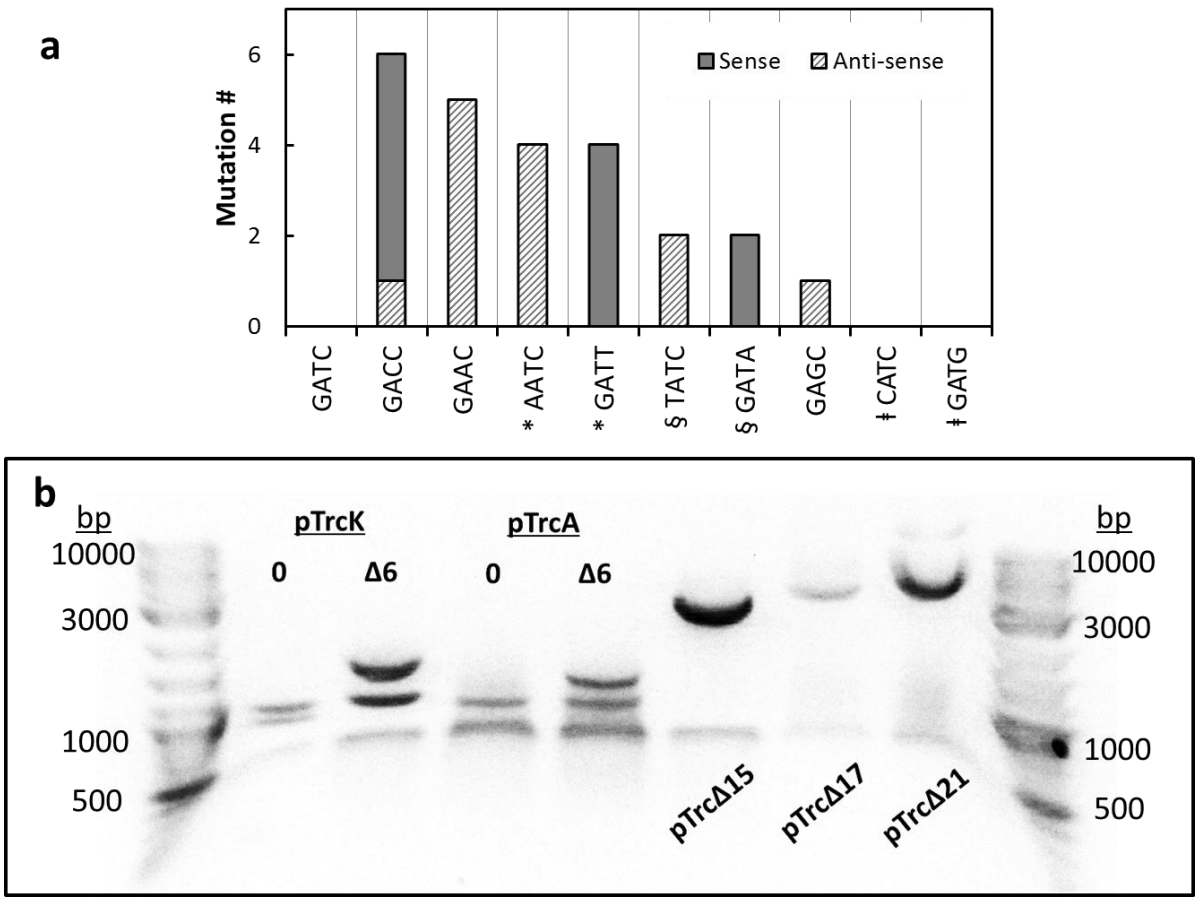


Figure 3.6 Conversion of a plasmid to Sau3AI digest-resistant pGATCfree based on differentiating plasmid architecture. (a) Frequency of the various single base mutations in pGATCfree construct. (b) Sau3AI digest patterns of pTrc series with increasing mutation number from left to right.

3.3.7 Embedding text characters as binary data in *egfp* gene

Having established the architecture of pGATCfree and demonstrating its relative indestructibility in the presence of a plasmid-chewing enzyme compared to the original sequences, pTrcK^r and pTrcA^r, the next efforts were devoted to developing a data-watermarked gene. *Egfp* was used as a model gene because its expression had been characterized in Fig. 3.5. Watermarking a gene with binary code requires codon substitutions that have been pre-designated as 0 or 1. Fig. 3.7a delineates codon representations of 0 and 1 for each amino acid. As has been demonstrated in literature, codon assignments were made in order of most to least frequently used in *E. coli*. Thus, for a given amino acid, multiple choices are available to represent a bit. While the gene sequence may change significantly, the amino acid sequence, and by reason, protein conformation and functionality should be unaffected. The bit assignment used in Fig. 3.7a expands beyond previous demonstrations in that we have made wider use of available codons, including any with at least one alternate codon, each having a frequency of at least 0.1. By definition, ATG for methionine and codons for cysteine and tryptophan were excluded. The code definition also includes bit assignments for two stop codons, TAA and TGA, allowing for a custom-length extension of the end of a gene to accommodate remaining bits. Special provisions were made for isoleucine and aspartate codons to prevent additional GATCs from being written into the encrypted sequence, discussed further below. Fig. 3.7b describes *egfp* watermarking to binary data encryption. Modern text uses 8-bit ASCII, allowing for roughly up to 30 characters to be stored in the 720 bp sequence of *egfp*. Alternatively, the Matlab embedded function char(bin2dec()) converts 7-bit code to characters, whereby representing text through 7 bits increases the capacity of *egfp* for storing 4-5 additional characters. Should the encrypted text begin with the character W, represented by 1010111 (Fig. 3.7b), the first 8 codons are used- the

first is the start codon and is therefore omitted, while the subsequent codons are adjusted to represent 1010111. In this case, bit storage is achieved for this character by only 2 codon alterations. This strategy differs from some previously published code definitions, where 0 and 1 assignments were made based on changing a codon from an established, optimized, gene sequence. By contrast, through external codon assignment, the original sequence need not be compared for data interpretation.

In Fig. 3.7c, we have written a Matlab-based program according to the assignment rules of Fig. 3.7a to decrypt the data from the sequence of a watermarked gene. The function, DNAdecrypt() requires the sequence to be input as a string. The string is analyzed by codon (n:n+2), assigned 0 or 1 accordingly, and added to an array (binary[ ]). Fig. 3.7c(i) demonstrates how the program proceeds through a list of string options for n:n+2 until a match is found and a binary digit assigned. Special consideration is required for aspartate codons (Fig. 3.7c(ii)). To maximize *egfp* data storage capacity, its GAT and GAC codons were used for 0 and 1, respectively, except when adjacent to a downstream C-initiated codon. For this exception, GAC assignment is mandated to prevent GATC cut-site formation, and so the codon is empty of data. In Fig. 3.7c(ii) the DNAdecrypt program accounts for this event by skipping the conversion of n:n+2 = GAT or GAC to 0 or 1 if n+3 = C, but adding the bit to the array otherwise. Finally, in Fig. 3.7c(iii), the binary array, analyzed as 7-bit code, is decrypted and reported as a string output.

a

AA	0	1	0	1
A	GCG	GCC	GCA	GCT
C/D	TGT(C)	TGC (C)	GAT (D)	GAC (D)
E/F	GAA (E)	GAG (E)	TTT (F)	TTC (F)
G	GGC	GGT	GGG	GGA
H/N	CAT (H)	CAC (H)	AAC (N)	AAT (N)
K/L	AAA (K)	AAG (K)	CTG (L)	TTA (L)
L (ct'd)	TTG	CTT	CTC	CTA
P	CCG	CCT	CCA	CCC
Q/R	CAG (Q)	CAA (Q)	CGT (R)	CGC (R)
S	AGC	TCT	TCC	AGT
T	ACC	ACA	ACG	ACT
V	GTG	GTT	GTC	GTA
Y/Stop	TAT (Y)	TAC (Y)	TAA (stop)	TGA (stop)

b

Character	W
7 bit code	1010111
Amino acid sequence	M V S K G E E L
eGFP native sequence	AtgGtgAgcAagGgcGagGagCtg
eGFPtxt sequence	AtgGttAgcAagGgcGagGagTta

c

```
function out = DNAdecrypt(string)
n = 1;
binary = [];
while n < length(string)
    %A = Alanine
    if string(n:n+2) == 'GCG'
        binary = [binary '0'];
    elseif string(n:n+2) == 'GCC'
        binary = [binary '1'];
    elseif string(n:n+2) == 'GCA'
        binary = [binary '0'];
    elseif string(n:n+2) == 'GCT'
        binary = [binary '1'];
    %D = Aspartate
    elseif string(n:n+2) == 'GAT'
        if string(n+3) ~= 'C'
            binary = [binary '0'];
        end
    elseif string (n:n+2) == 'GAC'
        if string(n+3) ~= 'C'
            binary = [binary '1'];
        end
                ...
k = length(binary);
p = 1;
final = [];
if rem(k,7) == 0
    while p < length(binary)
        final = [final; binary(p:p+6)];
        p = p+7;
    end
    out = char(bin2dec(final))';
end
```

i

ii

iii

Figure 3.7 Decryption of a gene-embedded text message based on specified code definitions. (a) Codon representation of binary data. For each amino acid, its codon options are designated a 0 or 1 value as defined in the chart. The following amino acids are not used: M, I, and W. (b) Demonstration of embedding the character W (based on ASCII 7-bit code) into the *egfp* at its start. Underline indicates a modification compared to the original sequence to accommodate the binary code. Red codons are data-empty (not included in the cipher). (c) Matlab program designed to decipher the embedded text message from a DNA sequence. (i) Matching a codon hit to a 0 or 1 assignment in an array. (ii) Assessment of adjacent downstream codon to that of aspartate (D). The program only considers D codons to be data-embedded when they are not followed by a C-initiated codon. (iii) ASCII-based conversion of binary array to a string of characters.

The coded sequence for *egfptxt*, watermarked with an encrypted message, and therefore differentiated from *egfp*, was written to include 30 characters (despite the necessary exclusion of

23 incompatible codons for bit storage). Additionally, the message was followed by the ETX "end of text" character termination sequence 0000011 to provide a formal "stop decryption" term that can be programmed into a future version of the DNAdecrypt() function. After the encryption of a 29 character message, *egfptxt* had six remaining codons, five of which were compatible, yet unused for data storage; these were dedicated for the ETX character along with the insertion of two data-carrying stop codon repeats prior to the final translation stop codon TAA (Table 3.2). Compared to *egfp*, the watermarked *egfptxt* shared 73 % genetic similarity and 100 % protein alignment, yet its entire sequence embedded binary data.

Table 3.2 Gene extension via stop codon to incorporate an embedded data termination character. Underline indicates a modification to the EGFP sequence to accommodate the binary code. Red codons are data-empty (not included in the cipher). (TAA) indicates the final stop codon for the transcript.

Encypted data terminator	ETX (end of text)
7-bit code	0 0 0 0 0 1 1
egfp sequence remainder	Atg Gac Gag Ctg Tac Aag (TAA)
egfptxt sequence	Atg Gat Gaa Ctg Tat Aaa TGA TGA (TAA)

3.3.8 Selective accession of embedded data by exploiting pGATCfree architecture

We next tested the capacity of information-embedded pGATCfree – denoted pEGFPtxt - to be preserved in a plasmid mixture under digestion conditions. Moreover, the other constituent plasmid – denoted pEGFPdec, with susceptibility to GATC cutting - should serve to present a decoy gene that obscures the sequence. Both vectors should be carried together in a cell host such that they are inherently blended. That is, the embedded *egfp* gene should be exposed only via GATC-selective digest that eradicates only the decoy vector (Fig. 3.8a). To fulfill this objective, Top10 cells were transformed with both pTrcK^{r}-EGFPdec and pGATCfree(A^{r})-

EGFPtxt; both plasmids were maintained by culturing the cells with kanamycin and ampicillin. Due to leaky *trc* expression, the cells constitutively expressed eGFP (Fig. 3.8b), presumably a combination of both the native gene *egfpdec* and the embedded gene *egfptxt*. Constitutive expression of the fluorescent protein effectively unifies both genetic functions- that which programs the protein structure and that which stores the digital information. Importantly, the optical output of green fluorescence reveals the genetic locale of the embedded data, yet the gene sequence cannot be accessed without the extrication of its defining vector. That is, while genes are routinely sequenced, especially when their position within a DNA framework is known and targeted, the competing presence of a decoy gene would generate noise in the sequence output and therefore occlude the critical sequence information. Without the sequence information, the embedded gene cannot be uniquely amplified, easily identified as containing data, or decrypted. Thus, successful decryption of an embedded DNA sequence requires elimination of its decoy.

The pairing of pEGFPtxt and pEGFPdec is marked by several features. Overall, both plasmids function the same way, with *trc*-based gene expression and pMB1-based replication. However, subtle architectural changes influence each functionality to some extent and also change the structural stability of this framework in the presence of Sau3AI. In addition to the genetic variation between *egfpdec* and *egfptxt*, the other genetic difference is in the antibiotic resistance: pEGFPdec carries *nptII* while pEGFPtxt carries *bla*Δ*7*. To examine the capacity to selectively isolate pEGFPtxt, we tracked the *nptII* and *bla*Δ*7*sequence availability before and after Sau3AI treatment. The plasmid combination was isolated from Top10 and each gene was PCRed before and after treatment with Sau3AI. In Fig. 3.8c, both *nptII* and *bla*Δ*7* were amplified when Sau3AI was not applied, but only *bla*Δ*7* was amplified after Sau3AI treatment. The band intensity of each gene was plotted as a mean gray value measurement in Fig. 3.8d.

Prior to Sau3AI treatment, *nptII* and *bla*Δ*7*were amplified to an equal level, indicating that their relative concentration upon harvesting from the host strain was approximately equal. It is not clear whether the plasmids are maintained at exactly equal ratios intracellularly or whether their relative levels reflect the observations of Fig. 3.2c. Regardless, after Sau3AI treatment, their relative levels were dramatically disrupted such that *nptII* was nearly undetectable by PCR, yet *bla*Δ*7* amplification was minimally affected. By this demonstration, Sau3AI-based extrication of data embedded on pGATCfree was deemed feasible.

The decryption treatment was repeated, this time probing instead for *egfp*. Fig. 3.8e shows the vector combination before and after Sau3AI treatment, where approximately half of the DNA in the Sau3AI$^-$ lane was fragmented in the Sau3AI$^+$ lane, while the remainder was uncut. It follows that the former constitutes pEGFPdec while the latter is pEGFPtxt. The Sau3AI$^+$ treated plasmid mixture was PCRed using Txt-NcoI-F and Txt-HindIII-R for universal amplification of the multi-cloning site, annealing upstream and downstream, respectively. The final lane in Fig. 3.8e identifies *egfptxt* for extraction and sequencing. Finally, the output text embedded in *egfptxt* is revealed by inputting its sequence file as a string in DNAdecrypt(string) via MatLab; the converted text is WEBLAB@UMD #infobio#GATCfree48 (Fig. 3.8f). Computerized decryption, then, harnesses a direct link between a sequence file obtained through standard DNA sequencing and a pre-programmed decryption function containing the customized code definitions. Additionally, the capacity for compact information storage was demonstrated by using common internet-inspired features for hyperlinking online text via handles (@) and hashtags (#). These promote minimal text size through these single characters having designated functions; text is channeled directly to a relevant addressee via @ and categorized by subject matter via #, eliminating the necessity for excessive spacing and punctuation. The same

principles are applied to the *egfptxt* embedded text. WEBLAB@UMD tags the vector's founding location while #infobio denotes purpose (abbreviated for infobiology) and #GATCfree48 identifies the vector's enabling functionality; the text is conveyed using 30 characters to accommodate storage within *egfp*.

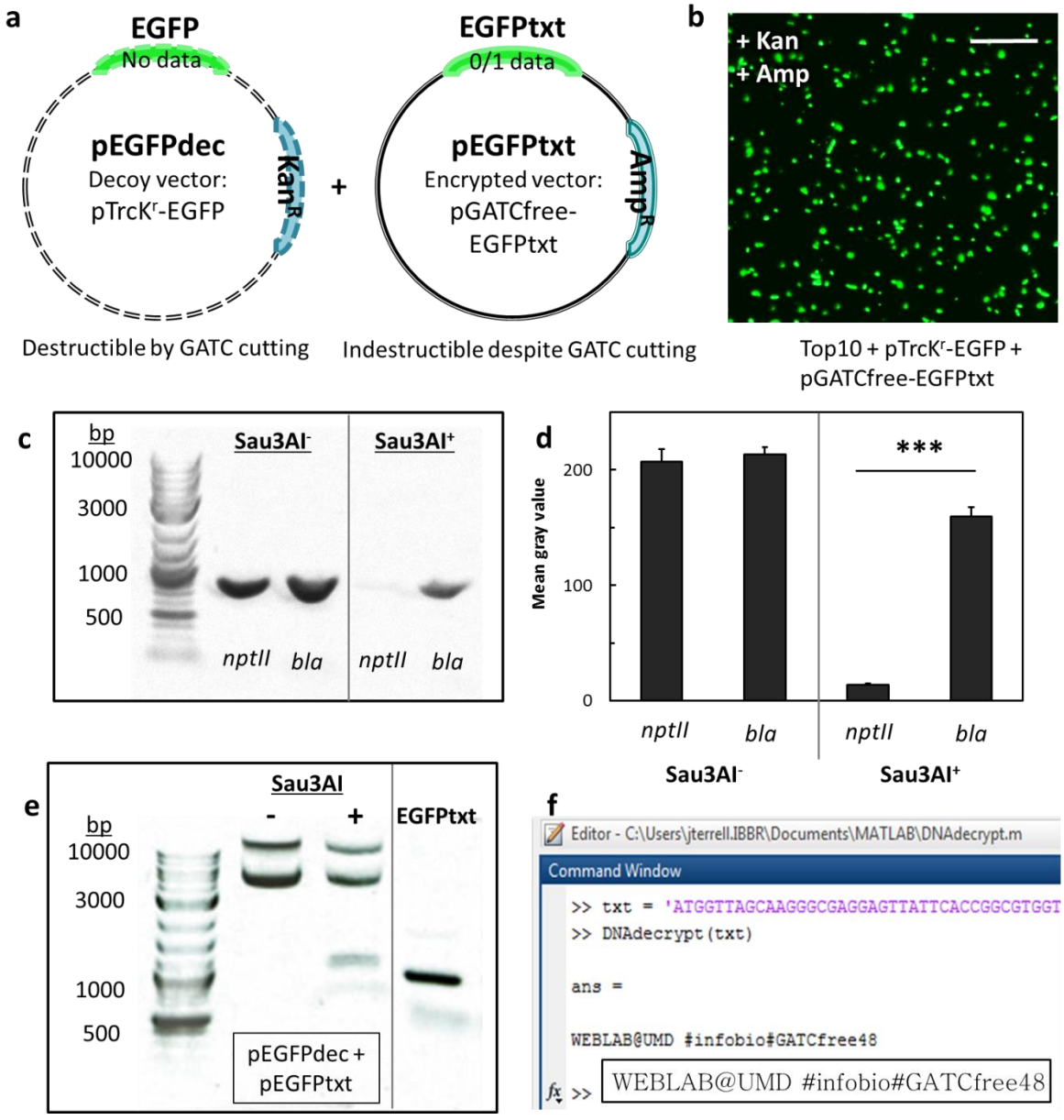


Figure 3.8 Distinction and decryption of a data-embedded *egfp* gene among obscuring sequences. (a) Scheme of paired plasmids for "interlocking" embedded *egfp* data with an empty *egfp* sequence as a decoy. pEGFPdec is represented by pTrcK^r-EGFP and pEGFPtxt is represented by pGATCfree-EGFPtxt.

Selective GATC cutting degrades the decoy vector to eliminate sequence noise. (b) Simultaneous fluorescence expression of eGFP and eGFPtxt in Top10 cells transformed with the plasmid pair. (c) Bands of *nptII* (from pTrcK^{r}) and *bla* (from pGATCfree) PCR products before and after Sau3AI treatment. (d) Mean gray value quantification of DNA bands in panel c. (e) plasmid pair isolated from carrier cells before (-) and after (+) Sau3AI treatment. EGFPtxt PCR product using the digest-treated plasmid pair as a template. (f) Decryptiong of EGFPtxt DNA sequence to a text message.

3.4 Concluding Remarks

From this work, we found that plasmid DNA can be systematically modified for prevalent sequence alterations of GATC sites within genes, regulatory regions, and intergenic segments, yet suitable functionality was maintained. The resulting GATC-free nature of the plasmid constituted a unique feature could be exploited for its selective retrieval from an otherwise inseparable mixture with other plasmids. Here, restriction enzyme action specifically targeted GATC sites for cutting; not only was pGATCfree protected, but the fact that GATC is natively abundant in unmodified DNA resulted in severe degradation of other plasmids, which were reduced to much smaller fragments. We recognized that this feature is advantageous as an encryption key for locking and unlocking potential data.

Since data storage in DNA has been previously demonstrated through gene watermarking, this technique was applied for data storage in *egfp* on pGATCfree. Despite its limited storage capacity (with at least 21 base pairs designated per character), the gene was thoroughly enriched with information, especially since shorthand format was used. Typically, a watermark merely serves as an intrinsic identification feature for differentiation from counterfeits. However, our purpose was to embed digital data while also using the architecture of pGATCfree to add a layer of encryption; the plasmid-incorporated data was locked once paired with a decoy plasmid holding a data-empty version of the same gene to cloud sequence information. The cell hosting the plasmid pair played an active role in relaying relevant information about the data through

protein expression of the embedded gene that resulted in cellular fluorescence, as opposed to being a passive carrier. This aspect could be harnessed as an optically-triggered command for automated data processing in the future. To this end, we demonstrated that the stored sequence information of the gene could be automatically decoded using a computer program with the appropriate code definitions. This work lends itself toward biosecurity applications using DNA as a digital data storage medium.

Chapter 4: Digitally-programmed information relay from DNA to a chemical signal.

4.1 Introduction

Quorum sensing (QS), a phenomenon of microbial intercellular signaling, has encouraged the use of these well-defined cell networks for genetically engineered applications, such as for population-wide artificial cell responses and even synthetically mimicking the signal. The biological underpinnings of this work are based on manipulating QS communication through the programmed generation of a QS signal whose identity and context can be interpreted by cells. Recently, our collaborating work with M Servinsky *et al* (2015) showed that fixed quantities of the signal molecule, autoinducer-2, direct response from discrete subsets of engineered cells within a population, or quantized quorums[85]. In other words, coordination of response across a population is determined by threshold levels of the autoinducer signal, such that bacteria communicate in a concentration-based molecular 'language'. The focus of this work is to investigate the capacity to program transduction of autoinducer-2 (AI-2), which is the dominant form of QS in *E. coli* and *S. Typhimurium [86]*.

AI-2 is a byproduct of a biochemical pathway within the activated methyl cycle [87]. The event of methylation of a variety of substrates by methyltransferases requires the 1-carbon transfer from S-adenosyl methionine (SAM), a universal methyl donor. The demethylated donor, S-adenosyl homocysteine (SAH), is shunted through a Pfs-LuxS pathway; the enzymes Pfs and LuxS are AI-2 synthases that convert SAH to AI-2 in a two-step process. Pfs and LuxS, have been expressed and purified from *E. coli* and subsequently used *in vitro* for AI-2 synthesis from SAH [88]. In additional *in vitro* applications, the enzymes have been assembled into devices for synthetic AI-2 signaling [89], particularly in response to a concept paper published

by LeDuc et al (2007)[90], where a biological "nanofactory" was proposed as a hypothetical bioconstruct. As conceived, the nanofactory would consist of a targeting moiety such as an antibody, a sensing module that could recognize inputs (raw materials), a synthesis module that could convert the input into an output (product), and, potentially, a self-destruct module. Fernandes et al (2010) built a quorum sensing-inspired nanofactory that fused Pfs and LuxS as an assembly line [91]. Thus, Pfs and LuxS comprised the nanofactory's sensing and synthetic components for AI-2 generation. The fusion protein was denoted HLPT, with denoted "H" and "T" his- and tyr- amino acid tags on the N and C termini. Unlike the original concept, the Fernandes construct had an additional, optional domain (protein G) so that the entire construct could self-assemble to include a targeting antibody (whose Fc portion binds generically to protein G) [49,60]. In their example, the nanofactory was targeted to *E. coli* cells within a mixed population using an assembled anti-*E. coli* antibody and elicited an AI-2 triggered response by surface-targeted AI-2 synthesis. Thus resulted a proof-of-concept that a quorum sensing response may be elicited synthetically without the presence of a quorum.

Here, this fusion protein HLPT has achieved further versatility: coupling the event of methylation, specifically DNA methylation, as a precursor step creates an opportunity for sequence-encoded programmability in the reaction pathway. Thus, our inclusion of an additional DNA methyltransferase module coupled to the pre-existing nanofactory is a logical extension. It is novel in that the use of methyltransferases in synthetic applications is virtually non-existent; efforts largely focus instead on the effect of methylation on the particular substrate but generally disregard its role in generating SAH [92]. *E. coli* DNA adenine methyltransferase (EcoDam) was identified as an ideal methyltransferase for testing DNA methylation patterns because 1) its recognition sequence (GATC) doubles as a restriction site for a methyl-dependent family of

restriction endonucleases, and 2) an EcoDam knockout strain (ER2925) is available as a source of demethylated DNA for a substrate. EcoDam should use SAM as a methyl donor for methylating adenine of 5'-GATC-3' sites within DNA. As a nanofactory-coupled module, EcoDam serves the dual purpose of delineating catalysis according to GATC substrate availability as well as converting SAM to SAH for further AI-2 synthesis.

EcoDam has been structurally characterized through crystollagraphic studies by JR Horton et al (2006)[93] and its functional features can be summarized as follows. The structure of EcoDam consists of a site-recognition domain and a catalytic domain. The first domain is a five-helix bundle that guides the enzyme along a DNA tract, recognizing GATC by the affinity of residue K9 for Gua and a beta-hairpin that interacts with ThyCys to position the enzyme centrally across the methylation site. Finally, the enzyme is anchored by intercalation of the aromatic sidechain of Y119 into the Thy base pair region, which also loosens the local helical structure. Once bound, a seven-fingered catalytic domain flips out Ade through stacking and hydrogen bonding interactions with H228 and V261, while R137 and N120 stabilize the orphaned Thy. The flipped Ade is shifted to the active site pocket, which is secured by the introduction of the methyl donor, S-adenosyl methionine (SAM) for methyl transfer to the N6 position of Ade.

Other work by S Urig et al (2002) studied the methylation reaction behavior to conclude the EcoDam acts with processivity, or methylates repeatedly along a single DNA strand before dissociating [94]. EcoDam surveys DNA by a random walk and in this way, is able to methylate on average 55 sites prior to dissociating, as determined by a combination of experimental data and kinetic models. The methylation rate was measured to be on the order of 20 turnovers per minute, particularly for chromosomal DNA. Further, it was discovered that despite GATC being

a palindrome- its reverse complement also GATC in the 5'→3' direction, EcoDam only methylates along a single strand.

This fundamental understanding of the mechanism by which EcoDam methylates DNA led to the hypothesis that *de novo* GATC-templating of a DNA sequence should yield an equivalent number of methylation events, regardless of the custom spacing and frequency within the sequence design. Moreover, since same-strand GATC sites are methylated at a faster rate (by 100-fold) compared to those on different strands, the methyl transfer over a time period could be favored by both GATC number and proximity. In this work, methylation rates were measured through the conversion to AI-2 output. AI-2 production was demonstrated as a function of GATC content on DNA substrates of increasing length and complexity: dsDNA oligomers, in vitro assembled linear DNA up to 1500 bp, and on unmethylated GATC-modified plasmids sourced from dam- *E. coli* cells. Finally, rolling circle amplification is explored as a strategy to generate significant GATC repeats while overcoming the tendency of GATC-rich DNA toward amplification-inhibiting secondary structure. EcoDam in combination with an AI-2 synthesizing nanofactory is compelling because it enhances function, transforming the commanding force for small molecule manufacture from simple enzyme kinetics to being DNA-instructed.

4.2 Materials and Methods

Materials. The following enzymes were purchased through New England Biolabs: EcoDam (*Dam* methyltransferase), S-adenosyl methionine (SAM), DpnI and DpnII, NcoI-HF, HindIII-HF, Q5 and Phusion polymerases for routine PCR, and phi29 polymerase for rolling circle amplification. Isopropyl β-D-1-thiogalactopyranoside (IPTG), S-adenosyl homocysteine (SAH), Ellman's reagent, and agarose powder were purchased from Sigma. Exo-resistant random

hexamers were purchased from Thermo-Scientific. All single-stranded DNA oligomers and GATCblocks (GBlocks) were purchased through Integrated DNA Technologies. Equipment included a thermocycler (BioRad), gel electrophoresis apparatus (Thermo Scientific), a gel imager (HP Alpha), and luminometer (Promega).

Purification of HLPT. HLPT production followed the general procedure outlined by R Fernandes et al (2010) [91]. Briefly, BL21 + pTrc-HLPT cells were grown in LB + 50 μg/mL ampicillin and induced with 1 mM IPTG upon reaching mid-exponential phase. Cultures were induced for 6 h at 37 °C, then pelleted. Metal ion affinity chromatography was used to purify the HLPT protein: the pellets were sonicated and supernatants were pumped through a Ni^{2+} chelated sepharose column (GE Healthcare) and then eluted with immidazole. Finally, the eluted protein was dialyzed in 10 mM PBS at 4 °C for 24 h.

Methyl-initiated AI-2 synthesis reactions. Unless otherwise specified, methylation reactions were prepared with 0.08 U/μL EcoDam and 2 μM HLPT in 10 mM PBS. For reactions with oligomeric substrates, 320 μM SAM; for other reactions, SAM was used at 64 μM. Positive controls used specified concentrations of SAH instead of SAM. DNA was absent in negative controls. For reactions using oligomer DNA substrate, primers were mixed at equal concentrations and in MilliQ water and incubated on a thermocycler for touchdown annealing. Touchdown annealing was performed by an incremental decline in temperature from 95 to 4 °C in 10 s intervals. Annealed oligomers were then added to reactions at either 10 or 50 μM. GATCblocks were used as methylation substrates at 0.03 μM, equating to 1.5, 3, and 4.5 μM AI-2 potential for A, AB, and ABC, respectively. Plasmid substrates were prepared at 0.02 μM. Methyl-initiated reactions were incubated at 37 °C for 4 h unless otherwise specified. For post-reaction DpnI and DpnII digests, reaction aliquots comprising approximately 500-600 ng

plasmid were recovered. Digestion reactions were supplemented with the stock buffer recommended in their product information.

Ellman's Assay. Homocysteine, the byproduct of AI-2 production, was quantified with Ellman's reagent, 5,5'-dithiobis-(2-nitrobenzoic acid) or DTNB (Sigma). AI-2 was twice-extracted from reactions with an equal volume of chloroform and recovered from the aqueous phase. The Ellman's assay was prepared as described previously [95] [96]. Briefly, AI-2 extractions were mixed with 100 μM DTNB (diluted from 2 mM stock in 50 mM sodium acetate) in 100 μM Tris (pH 8). Absorbance at 412 nm was measured and the concentration of homocysteine was determined using 13600 $M^{-1}cm^{-1}$ as the extinction coefficient.

Bioluminescence Assay. *Vibrio harveyi* BB170 was cultured overnight at 30 °C in AB medium with 10 μg/mL kanamycin [86]. The 16 h BB170 culture was diluted 1:5000 into fresh AB medium. Diluted BB170 (180 μL) was mixed with each reaction sample (20 μL) and incubated in 5 mL culture tubes for up to 4.5 h at 30 °C, 250 rpm shaking, then recording the bioluminescence measurement from each sample via luminometer. Bioluminescence is reported as a fold change compared to the negative control measurement.

Rolling circle amplification. Touchdown annealing of the DNA template, either 2 ng or 1 μL of a Gibson Assembly reaction, and 20 μM Circ-F as the padlock primer was performed in 1x phi29 reaction buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 10 mM $(NH_4)_2SO_4$, and 4 mM DTT). Then on ice, the reaction volume was doubled to include 5 U phi29 polymerase, 0.025 U pyrophosphatase, 2.5 μg BSA, and 50 μM exo-resistant random hexamers in 1x reaction buffer. The reactions were incubated for 16 h at 30 °C and then run on a 0.7 % agarose gel with SyberSafe DNA stain (Invitrogen).

4.3 Results and Discussion

4.3.1 Dpn enzymes for tracking the methylation state of DNA

The Dpn restriction endonuclease family, discovered in *Diplococcus pneumoniae* G41, specifically cleave GATC sites, naturally serving as a defensive mechanism against bacteriophage infection [97]. Furthermore, DpnI and DpnII have unique selectivity according to the methylation state of GATC[79] (Table 4.1). DpnI is methyl-dependent, requiring both adenines on complementary strands to be methylated in order to generate a blunt cut. DpnII, on the other hand, is methyl-sensitive, generating a 4 bp overhang only when the GATC adenines on both strands are unmethylated. Therefore, cut band patterns of DNA treated with DpnI and DpnII should reveal the methylation state of the DNA, whereby band patterns should develop in an opposite fashion with I compared to II.

Table 4.1 Methylation dependency for recognition of GATC cut site by Dpn restriction endonucleases.

DpnI	DpnII
CH_3 | GA/TC CT/AG | CH_3	/GATC CTAG/
CH_3-depenendent	CH_3-sensitive

We first characterized the cleavage capability of DpnI and DpnII on unmethylated and methylated plasmid DNA; simultaneously, we tested a Δ*dam* strain of *E. coli* (ER2925) as a source of plasmid DNA that is unmethylated at GATC sites. While ER2925(Δ*dam*, Δ*dcm*) is commonly used to transfer shuttle vectors from an *E. coli* host to organisms that restrict methylated DNA, such as eubacteria or archaea[98], it is less commonly used for plasmid maintenance due to its higher mutation rate[99]. Because EcoDam is critical to mismatch repair

of chromosomal DNA, its absence leads to an 8 - 250fold greater rate of spontaneous mutations than a *dam*$^{+}$ strain [100]. Considering that plasmids for our experimental purposes were only intended as a substrate for EcoDam and not as a genetic vector, we did not check for mutation prevalence of ER2925-isolated plasmids, but we did seek to ensure that ER2925 could reproducibly provide high yields of unmethylated plasmid DNA.

The plasmid pTrc(A^{r}) was isolated from either a *dam*$^{+}$ or *dam*$^{-}$ *E. coli* host, either Top10 or ER2925, respectively. ER2925-sourced pTrc was treated with EcoDam and S-adenosyl methionine (SAM) as a methyl donor. Negative controls for the methylation reaction excluded either EcoDam or SAM. Then, Top10-sourced DNA (used as a positive control of methylated DNA) and EcoDam reaction products were treated either with DpnI or DpnII. The restriction digestion reactions were loaded into an agarose gel to size-separate DNA fragments. Fig. 4.1 shows the resulting band patterns of enzyme-treated DNA, with DpnI-cut fragments on the left and those cut by DpnII on the right. Lane A, loaded with plasmid from a *dam*$^{+}$ host, was cut into several fragments by DpnI but uncut (with no migration) when treated with DpnII. This result confirms that methylation prevents DpnII digestion but creates susceptibility to DpnI. The band pattern of Lane D, comprising DNA that was methylated *in vitro* by EcoDam, matched that of A, which indicates that the DNA was fully methylated. To confirm that methylation occurred *in vitro*, and not by the cell host, negative controls were also evaluated, lacking either the methyltransferase (B) or the methyl donor, SAM (C). Both lanes show a reverse effect compared to the positive control in Lane A: the DNA was left undigested when treated with DpnI but cut by DpnII. The mirrored pattern effect between lanes A and B/C comparing DpnI and II results indicates the DNA introduced to lanes B and C was unmethylated. From this result, it can be concluded that ER2925 is a suitable source of unmethylated plasmid for *in vitro*

reactions and that both EcoDam and SAM are essential for methylation to occur (showing no residual methylation in the absence of either). By contrasting to cell-methylated and *in vitro*-methylated DNA, it is clear that DpnI and DpnII are effective in distinguishing methylation state. Finally, the *in vitro* reaction performed for lane D shows methylation efficacy that is comparable to intracellular levels.

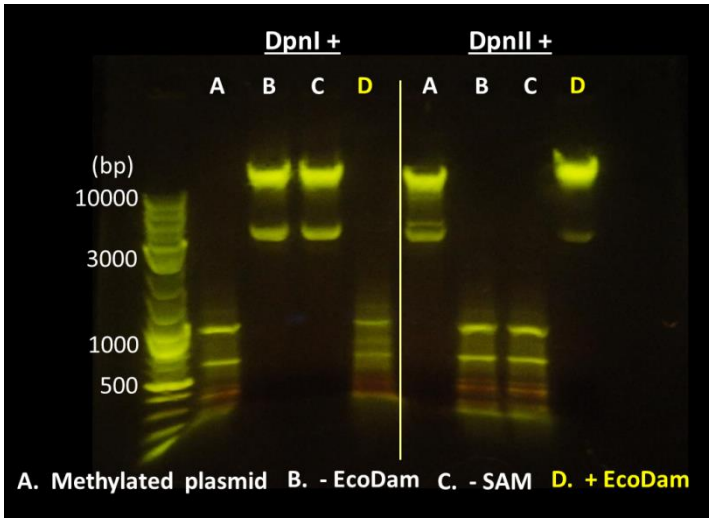


Figure 4.1 Restriction endonuclease-based identification of the methylation state of plasmid substrates from a *dam*-deficient host.

4.3.2 Methylation activity of EcoDam

Next, DpnI and DpnII were used to visually determine the time scale of DNA methylation. EcoDam was introduced to 4 μg pTrc at either 4, 8, or 12 U with equal SAM levels. At time intervals, aliquots were isolated and incubated with either DpnI or DpnII. Digested samples were run on an agarose gel and imaged. Fig. 4.2 shows DpnI and DpnII digestion results across the timepoints and with increasing EcoDam activity. Initially, the plasmid is completely unmethylated- uncut by DpnI and fragmented by DpnII. At every timepoint thereafter, and under every condition, the DNA is cleaved by DpnI treatment (Fig. 4.2, upper). However, a subtle shift in band patterns is observed by DpnII treatment. In Fig. 4.2 (lower), while the first

timepoint (0.5 h) shows a grouping of bands that are dramatically larger than t_0, the band number decreases along the subsequent timepoints, with the smaller bands of the group gradually disappearing. The point of DpnII preclusion is at 3 h when treated with 4 U EcoDam, 2 h for 8 U EcoDam, and between 1-2 h for 12 U EcoDam. At these timepoints, the band pattern (which likely comprises bands of non-specifically nicked, relaxed, and supercoiled plasmids) matches that of DpnI at t_0, where the former indicates complete preclusion of DpnII due to methylation at t_f, and the latter indicates incapacity of DpnI cleavage due to a lack of methylation at t_0.

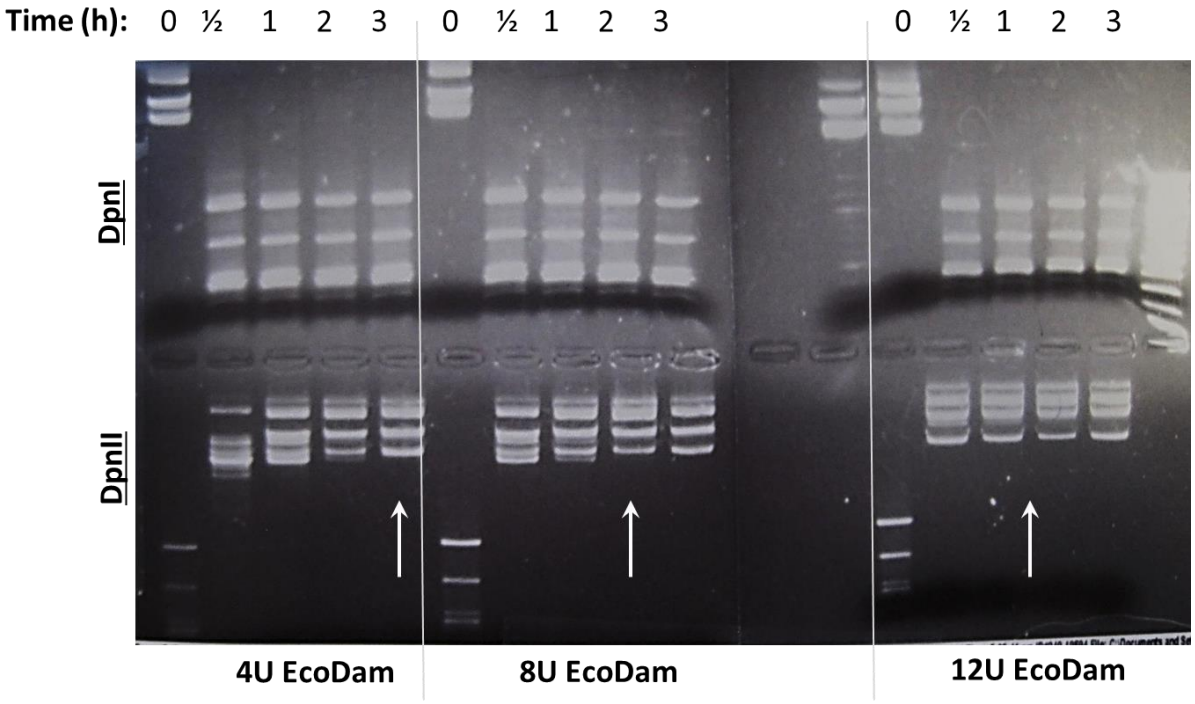


Figure 4.2 Optimization of EcoDam methylation rate for large DNA loads using digestion analysis.

The results show, firstly, that DpnII enables improved resolution of methylation events, where cleavage changes due to methylation events over time were readily observed at $t > t_0$ for DpnII but not DpnI. Secondly, the time scale required for complete methylation was characterized for a range of EcoDam activities. According to the product information, 4 μg of DNA should be fully methylated by 4 U EcoDam in 1 h at 37 °C. However, this guideline is true for λ

bacteriophage DNA, which differs from plasmid substrate in that it is significantly longer and therefore promotes processivity along long DNA tracts as opposed to frequent dissociation from smaller substrates upon their rapid methylation. In the case of plasmid DNA, we found complete methylation to take 3 h, but that by doubling EcoDam activity, 2 U per μg reduce the completion time to 2 h. Tripling the activity to 3 U per μg appear to further reduce the time to between 1 and 2 h. In this case, the final band pattern is nearly set by 0.5 h, but a residual band that is second-smallest in size is apparent (albeit faint) past 1 h. This indicates that the associated GATC sites are the least favorable to an otherwise completely methylated plasmid, and consequently, rate limiting. It appears, then, that the presence of excessive EcoDam does not accelerate methylation of this site. Thus, the middle condition, 2 U EcoDam per μg DNA, was applied to subsequent experiments, noting that 2 h should be sufficient for complete methylation.

Lastly, the DpnII trends, in particular, show based on the positioning of band patterns that upon methylation over time, the size of DNA fragments appear to grow. This indicates that methylation across plasmids occurs in a coordinated fashion rather than stochastically. In the latter scenario, one would expect the same band pattern for smaller fragments to persist over time while growing fainter and for large bands to gradually appear in a similarly equal distribution. Instead, small bands disappear in the order of their size, while the larger bands shift to their final position. This effect is due to EcoDam processivity, where the GATC sites in closest proximity (which define the smallest DpnII-cut fragments) are the most likely to be methylated first and within the same binding event; this quickly grows the methyl-protected portion of the DNA, while the most distal positions are eventually methylated. Additionally, GATC binding preferentiality- due to flanking sequences- has also been shown to contribute to methylation

order on a similar plasmid [79]. For both reasons, methylation order across plasmid substrates appears to proceed in a generally synchronized manner.

4.3.3 Characterization and rate kinetics of methylation-coupled AI-2 synthesis

Next, we sought to expand the reaction pathway to the methyltransferase-initiated synthesis of AI-2 by focusing on optimization of the Pfs-LuxS fusion (HLPT) and working upstream toward the methylation step. To optimize the reaction conditions for the 2 h timeframe demonstrated with EcoDam, maximum AI-2 yields produced by the fusion were targeted. Two concentrations of HLPT were tested across a range of SAH substrate concentrations. After 2 h incubation at 37 °C, the percent conversion of initial SAH to AI-2 was measured via Ellman's assay. The Ellman's assay enables quantification by reacting with the byproduct of AI-2 synthesis, homocysteine, which is produced with 1:1 stoichiometry. Fig. 4.3a plots the AI-2 yields obtained by HLPT catalysis over 2 h when incubated with SAH ranging from 50 μM to 1 mM. We found 50-100 μM enabled the most rapid conversion, particularly when treated with a double-dose of HLPT.

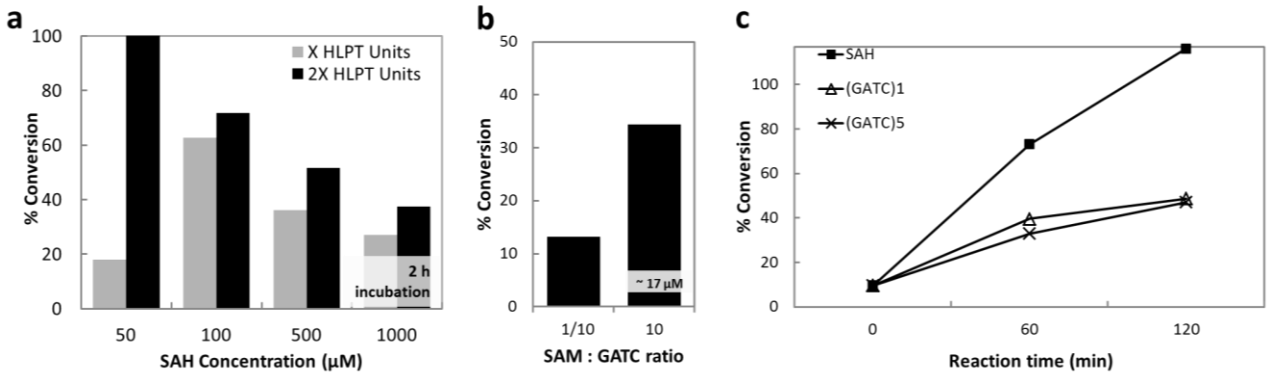


Figure 4.3 Optimization for autoinducer-2 generation *in vitro*. (a) SAH-dependent AI-2 synthesis over 2 h. (b) Methylation-dependent AI-2 synthesis over 1 h with varied ratios of required substrates. (c) AI-2 synthesis kinetics based on different pathway starting points.

Next, AI-2 synthesis was coupled to EcoDam-catalyzed methylation of DNA. Not shown, we found incompatibilities between the preferred buffers for HLPT and EcoDam. EcoDam lost activity when dialyzed in PBS, which is typically used to dilute HLPT. HLPT, by contrast, lost activity in a standard EcoDam buffer (50 mM Tris-HCl, 5 mM β-mercaptoethanol, and 10 mM EDTA, pH 7.5). The traces of Tris-HCl, DTT, and EDTA from the EcoDam storage solution were found enable sufficient EcoDam activity in PBS, such that the final universal buffer comprised 500 μM Tris-HCl, 500 μM KCl, 10 μM DTT, 100 μM EDTA, and 0.5 % glycerol in PBS.

Using the universal buffer, EcoDam and HLPT were both added along with SAM as a methyl donor and GATC-containing DNA as a methyl acceptor. In this way, methylation must precede AI-2 synthesis because the byproduct of a methylation event, SAH, is required as the substrate for HLPT. However, since EcoDam requires two substrates, we evaluated the efficacy of the complete reaction pathway when each substrate, in turn, was provided in excess. First, GATC sites on DNA were included in 10-fold excess compared to the SAM concentration. Then the condition was altered such that SAM was included in 10-fold excess compared to GATC number. Conversion percentage was calculated based on the limiting reagent in each condition and is shown in Fig. 4.3b. For SAM-limiting conditions, we found a relatively low yield of less than 15 % conversion. However, when we switched the ratio to favor an excess of SAM and limiting GATC, conversion increased to 35% within 1 h. The faster rate of conversion when SAM is in excess is probably explained by several factors. First, EcoDam binds to DNA first, but then SAM is required for stabilization at a GATC site. Secondly, since SAH is a competitive inhibitor of EcoDam, the greater availability of SAM competes favorably to promote continued methylation while SAH can be diverted for HLPT processing. Regardless, GATC-limiting conditions are favorable for the goal of a pre-programmed DNA substrate in order to delineate AI-2 output.

Finally, AI-2 synthesis kinetics were evaluated over a 2 h time period at one of two reaction starting points: either the event of methylation or SAH conversion. Methyl-initiated synthesis was prepared using a double-stranded oligomer as a DNA substrate containing either 1 or 5 GATC pairs (Table 4.2). Methyl-initiated and SAH-initiated reactions were prepared with equivalent concentrations by considering the number of GATC sites per oligomer. Fig. 4.3c shows the conversion of each reaction over 2 h. Within 2 h, the SAH-initiated reaction reached 100 % conversion (surpassing 100 % due to error in measurement). Methyl-initiated reactions proceeded at approximately the same rate regardless of whether $(GATC)_1$ or $(GATC)_5$ was used, reaching, however, only 40 % conversion at 2 h. Therefore, methyl-initiated reactions resulted in slower conversion to AI-2 than by SAH initiation. While EcoDam was previously tested by Fig. 4.2 to fully methylate DNA in 2 h, the nature of the substrate likely contributed to a slower rate than reported in the product information (where plasmid DNA was 10-fold shorter in length than λ DNA); again the oligomeric DNA in Table 4.2 is 100-200-fold smaller than plasmid DNA. Previous literature has stated that methylation requiring frequent obligatory dissociation is significantly slower than along longer DNA substrates [93]. Thus, the length of oligomeric substrates in this experiment could be responsible for the reduced conversion. The dissociation mechanism, however, did not observably influence the rate between $(GATC)_1$, having a single GATC pair, compared to $(GATC)_5$ with adjacent GATC repeats. Thus, at the oligomeric length scale, processivity of EcoDam does not have the prevailing influence on AI-2 conversion rate; however, the impact of the methyl-initiation step as a whole reduces the AI-2 conversion rate, meaning that SAH is processed faster than methyl transfer.

Table 4.2 Double-stranded oligomers with fixed GATC content as methylation substrates for fixed autoinducer-2 output.

Oligo	$(GATC)_1$ 20mer	$(GATC)_5$ 40mer
Sequence	5′- CAGTTTAGGATCCATTTCAC -3′ 3′- GTCAAATCCTAGGTAAAGTC -5′	5'- CATGGACTAGTGATC GATCGGATCCGATC GATCTCTAGAA -3' 3′- GTACCTGATCACTAG CTAGCCTAGGCTAG CTAGAGATCTT -5′
GATC content	2	10
AI-2 potential at 10 µM	20 µM	100 µM

4.3.4 Bioluminescence as a quantifiable quorum sensing response by *Vibrio harveyi* to autoinducer-2

The BB170 engineered strain of *Vibrio harveyi* was tested for bioluminescence-based discrimination of fixed concentrations of autoinducer-2 (AI-2) between 0 and 40 μM. AI-2 was synthesized *in vitro* by the nanofactory; the product concentration was determined via Ellman's assay, after which serial dilutions between 0 and 40 μM were prepared in PBS. AI-2 samples were mixed 1:10 with BB170, prepared as a 5000-fold dilution from overnight culture in supplemented AB medium, and cultured for 4.5 h at 30°C. Bioluminescence was measured per sample in relative light units and normalized to a negative control comprising BB170 mixed with PBS. The bioluminescence output of BB170 as a function of AI-2 exposure is plotted in Fig. 4.1.

Output from BB170 showed incrementally brighter bioluminescence in response to increasing concentrations of AI-2. The bioluminescence level did not saturate in the range of 0 to 20 μM AI-2 and each bioluminescence output was statistically significant within this range. The upper limit tested, 40 μM, did not include replicates for statistical analysis. The lowest concentrations, below 5 μM, showed the best resolution, where a 50-fold increase in luminescence was observed between 0 and 5 μM, but luminescence increased by only 14 % more at 10 μM. While the absolute bioluminescence level for an AI-2 input is not a reproducible measurement, BB170 assays are commonly used for comparative analysis of AI-2 activity between samples. Here, we show that this analysis can be extended to fixed AI-2 concentrations (which is not derived from a quorum sensing cell culture) and also demonstrate, roughly, a dynamic range at which differing AI-2 concentrations can be reconciled by BB170. Importantly, this experiment demonstrates that the quorum sensing 'language' is based on the *amplitude* of a molecular input. In other words, the AI-2 concentration introduced to BB170 directly influences the extent to which the population undergoes phenotypic change, in this case, by upregulating a bioluminescence pathway.

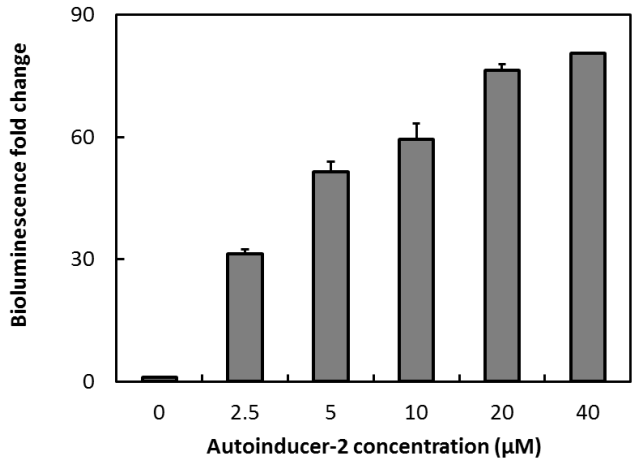


Figure 4.4 Dose response of bioluminescent BB170 reporter to the quorum sensing signal autoinducer -2.

4.3.5 GATC-dependency of AI-2 production during methyl-initiated synthesis

Since BB170 is responsive to AI-2 level and its bioluminescence is an accurate indicator thereof, we expected that the strain is capable of elucidating a 'scripted' DNA message by which AI-2 amplitude is programmed. As a prototype, we worked with the short double-stranded oligomers (Table 4.2) to act as substrates for a precursor methylation event followed by conversion of the SAH byproduct to AI-2. Fig. 4.5 shows the bioluminescence levels produced by BB170 induced by the reaction solutions post-initiation. Experimental samples were prepared with either 10 or 50 μM GATC content as methyl substrates. That is, the exact oligomeric concentration was adjusted to achieve 10 or 50 μM AI-2 potential based on its GATC content. Shown in Table 4.2, due to the existence of GATC as pairs on dsDNA, the AI-2 potential of (GATC)1 is twice the 20mer's concentration and ten-fold higher than the concentration of the (GATC)5 40mer. Oligomers inscribed with GTTC rather than GATC were used as DNA substrates for negative controls. For positive controls, SAH-initiated reactions were used with equivalent AI-2 potentials (10 or 50 μM SAH). Bioluminescence fold changes in Fig. 4.2 are baselined to a blank containing only SAM and SAH as substrates.

The 10 μM samples show a bioluminescence response to both methyl-initiated and SAH-initiated AI-2 synthesis reactions that is significantly higher than the control with a GATC-lacking oligomer. Because

the bioluminescence response to the methyl-initiated reaction is approximately 40 % that of the SAH-initiated reaction, it appears that AI-2 production of the former did not proceed to the same extent as the latter. A lower yield is consistent with results from Fig. 4.3c, which showed that the rate kinetics of AI-2 production upon methyl-initiation is slower than with SAH-initiation. However, the 50 μM results in Fig. 4.5 match between oligomers and with the positive control. Here, equivalent output was generated despite differing starting concentrations- 25 μM of the (GATC)1 20mer and 5 μM of the (GATC)5 40mer. This implies that their matching AI-2 potential of 50 μM was effectively programmed and implemented based on the GATC content. That yields from methyl-initiation were closer to SAH-initiation for the 50 μM level compared to the 10 μM level could be due either to saturation of the bioluminescence signal (where 50 μM can be extrapolated as an upper limit in Fig. 4.4) or alternatively, saturation of the enzymatic velocity (where Fig. 4.3a implies that SAH concentrations exceeding 50 μM cannot be completely processed over 2 h- such as 60 % conversion of 100 μM SAH). Regardless, the range of 50 μM AI-2 potential would have exposed sub-optimal AI-2 conversion at below-saturation bioluminescence levels. Overall, the AI-2 responsive bioluminescence from BB170 reveals that AI-2 production can be controlled by GATC-content for methyl-initiated synthesis, where GATC content defines the AI-2 output to which cell phenotype is influenced.

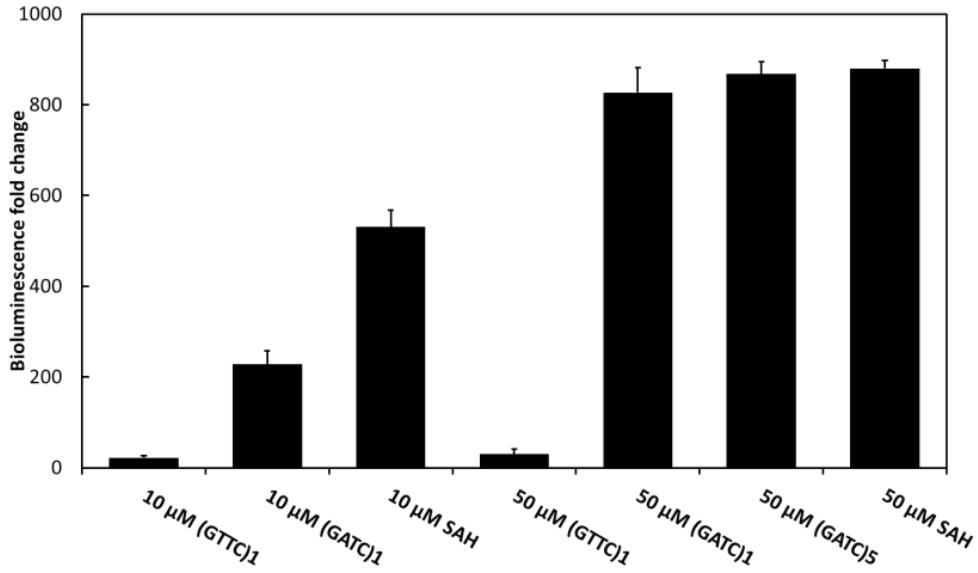


Figure 4.5 Quorum sensing response to autoinducer-2 conversion using methylatable oligomeric substrates as the limiting reagent. Concentration labels indicate AI-2 potential, or the concentration of the limiting substrate reagent. For methyl-initiated reactions, the concentration reflects the accounted-for GATC content and the label indicates the oligomer's GATC composition.

4.3.6 Assembly of designer GATC-rich DNA blocks as gradated inscriptions

Next, we sought to further enrich the GATC content of DNA substrates such that its variation could influence greater deviations in BB170 bioluminescence. Here, GATC density is critical for increasing GATC content, such that AI-2 potential can be significantly enhanced while maintaining reasonable DNA concentrations. Since DNA with high GATC density regions is not naturally present as a substrate or template, we chose to purchase DNA with a synthetic, non-genetic custom sequence. *De novo* DNA synthesis is constrained by repeat frequency, length, and the potential for secondary structure; as such, the inclusion of GATC-rich segments presented a design challenge due to extensive repetition and the strong tendency toward secondary structure due to the palindrome nature of GATC. Thus, the purchase of a single full-length DNA piece with extensive GATC-rich regions was not possible. Since sequences are more stable while double-stranded, our approach assembled multiple GATC-rich 500mer dsDNA segments *in vitro*. Three 500mers were designed with 25 GATC repeats and minimal inter-GATC spacing (0 - 12 bp); these were denoted GATCblock A, B, and C (Fig. 4.6a(i)). The GATC density (GATC bp / total bp) comprised ~ 0.45 in the enriched sections and 0.2 per 500mer block; by contrast, the GATC density of the plasmid pTrcHis is ten-fold less at 0.02. GATCblocks A, B, and C were also designed with 20 bp overlaps for Gibson Assembly, such that the blocks can be covalently connected in ABC order to form a single 1500 bp DNA piece with 75 GATC pairs (Fig. 4.6a(ii)).

Shown in Fig. 4.6b(i), GATCblocks were assembled and amplified to yield the original 500mers (A, B, and C), a double-length 1000mer (AB), and a 1500mer (ABC). Thus, Gibson Assembly and PCR were both compatible techniques for these GATC-rich DNA pieces. Methylation potential was also tested on GATCblock A and analyzed via Dpn digest. While GATC sites were previously demonstrated as methyl-dependent fragmentation sites for DpnI and II, a slightly different observation is expected for GATC-rich DNA. That is, in a methyl-dependent fashion, substantial cutting occurs within the GATC-dense portion, where the short (< 12 bp) fragments are undetectable by standard gel electrophorsesis. The leftover GATC-free portions of the DNA, however, are left intact. For GATCblock A, Fig. 4.6a(i) shows that a 175 bp frame adjacent to the GATC-rich portion should be left intact after Dpn erosion. In Fig. 4.6b(ii), block A was resolved at 175 bp after Dpn I digest and 500 bp with DpnII. Because methylation enables cutting by DpnI and prevents DpnII, this result indicates that block A was successfully methylated.

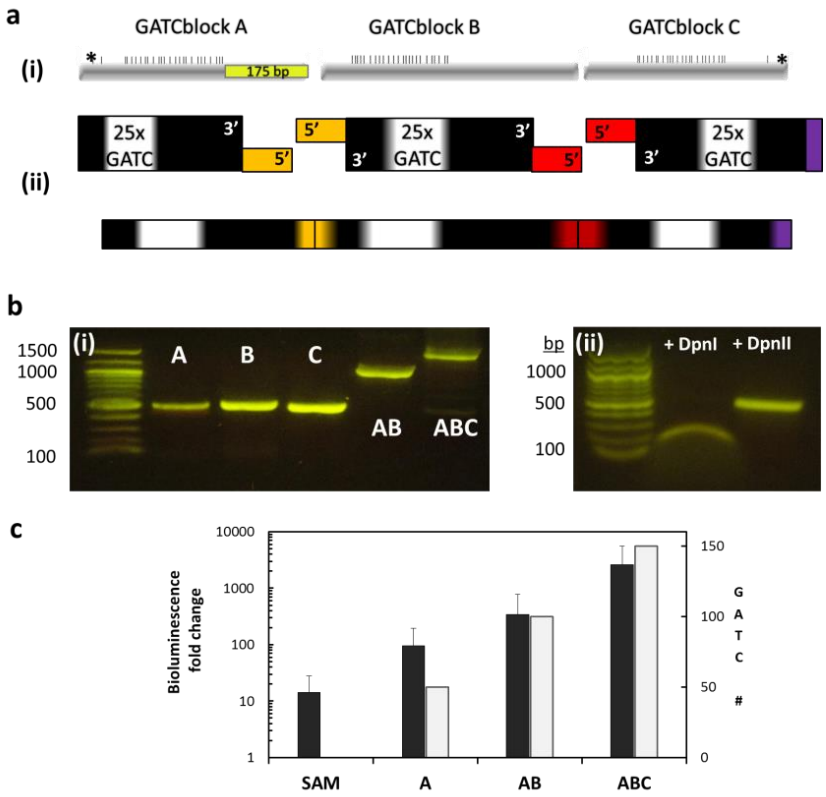


Figure 4.6 Use of linear GATC-rich DNA as substrates to delineate autoinducer-2 output. (ai) Unique maps of GATC sites along a 500 bp DNA blocks A, B, and C. (ii) Design of custom overhangs for ordered ABC assembly. (bi) Gel electrophoresis of separate and assembled GATCblocks. (ii) Dpn digest analysis of GATCblock A post-methylation. (c) Bioluminescence response to autoinducer-2 conversion delineated by GATCblock length and content.

Next, the DNA was tested for its ability to influence BB170 bioluminescence. Due to an incremental increase in GATC content with the DNA length, it was expected that by applying the same amount of DNA, the AI-2 output should increase for ABC, compared to AB and A, respectively. In Fig. 4.6c, the bioluminescence response to methyl-initiated AI-2 synthesis increase by approximately half an order-of-magnitude with incremental lengthening of the DNA substrate. Here, because equal DNA amounts were applied, it is the influence of GATC-rich repeats on the DNA that delineate the extent of quorum sensing response from BB170.

4.3.7 Inscription of plasmid DNA with GATC-enriched patterns

The GATCblocks were inserted onto a pTrc plasmid in order to maintain the GATC-rich sequences in a cell host for routine scaling-up. Fig. 4.7a maps pTrc for its native GATC sites and the available insertion sites for GATC-rich DNA. The vector pTrc has 42 GATC sites, considering GATC exists as a palindromic pair, and was renamed pGATC42 to track the methylation site number and AI-2 potential per plasmid. Three GATC pairs, however, are located in between the cut sites NcoI and HindIII; GATCblock insertion, A and AB, replaced this section with either 50 or 100 sites, for a net gain of 44 and 94 sites, and resulting in pGATC86 and pGATC136, respectively (Fig. 4.7a). In Fig. 4.7b, plasmid isolated from prospective clones was double-digested by NcoI and HindIII to screen for block A in pGATC86 and AB inserted in pGATC136. ABC insertion to obtain pGATC186 was unsuccessful. A 500 bp band for the former construct and a 1000 bp band for the latter were visible on an agarose gel, and distinctive from the pGATC42 empty vector. Additionally, because the clones had a *dam*$^{+}$ genotype, GATC sites were cut by DpnI digest in order to characterize the GATC-cut band pattern. Fig. 4.7c shows DpnI digest patterns compared to pGATC42, including GATC-enriched derivatives and also a series of plasmids where GATC content was systematically reduced to 6 per plasmid by site-directed mutagenesis (Chapter 3). Distinctive band patterns, according to the unique spacing and abundance of GATC sites on each plasmid are apparent. Each band pattern, then, could be considered a methyl-dependent signature or barcode that represents differing GATC content, or in other words, the particular AI-2 inscription.

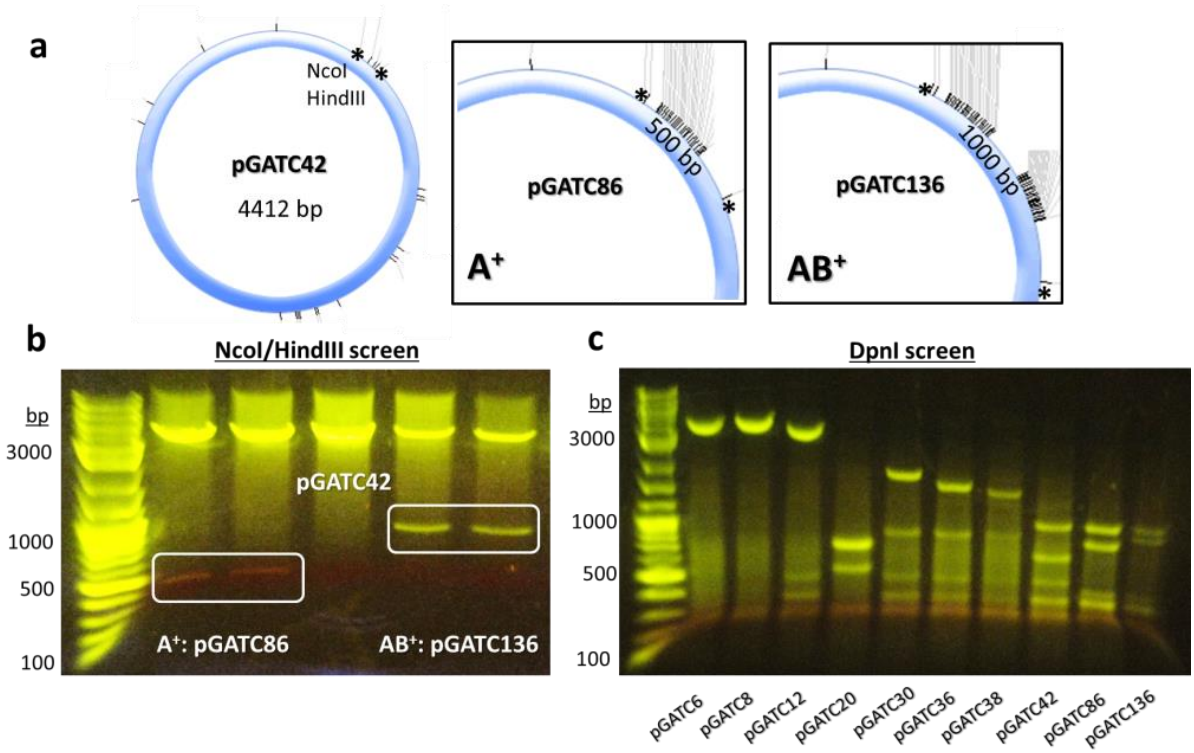


Figure 4.7 Construction of GATC-rich blocks within plasmid DNA to modularly augment the AI-2 inscription size and incorporate a veritable band pattern. (a) Insertion of single (A) and double (AB) GATCblocks into the multi-cloning site (at NcoI/HindIII) of a plasmid with 42 exisitng GATC sites. (b) NcoI/HindIII digestion to screen for GATCblock insertion. (c) Dpn-cutting of GATC sites to reveal band patterns of a plasmid series with varying GATC content.

4.3.8 Inscribed plasmid DNA for AI-2 delineation with post-read documentation

A subset of the plasmid series was chosen to delineate varied AI-2 signal amplitudes and subsequently transformed into the Δ*dam* strain ER2925, from which demethylated plasmids were isolated. Using pGATC6, 42, 86, and 136 as substrates for methyl-initiated AI-2 production, we noted that the AI-2 potential increases by approximately 1, 2, and 3-fold, respectively, compared to pGATC6. Bioluminescence response to the AI-2 produced by each reaction is shown in Fig. 4.8a. Again, the addition of each GATC-enriched block influenced the bioluminescence to increase incrementally, with near-absent bioluminescence delineated by pGATC6. Next, the inscription barcode was probed by GATC-targeted digest. Using DpnI and II for methylation-

specific action, it was possible to simultaneously scan the plasmid identity and whether the inscription had yet been "read." Fig. 4.8b shows the "scanned" pattern obtained through gel electrophoresis under DpnI and DpnII conditions. For lane II, all plasmids are mostly undigested, having migrated less than the reference ladder. Some faint smearing, particularly for pGATC136 could indicate residual digestion. DpnI digestion, by contrast, is readily apparent in all lanes, where the DNA has unanimously been fragmented to smaller pieces than lane II. According to the methyl effect on DpnI and II, the results confirm near-complete methylation, or in other words, that the inscription has been read and enzymatically relayed. Finally, the barcodes generated by DpnI match their characteristic patterns obtained in Fig. 4.7c, confirming the identity of each plasmid construct.

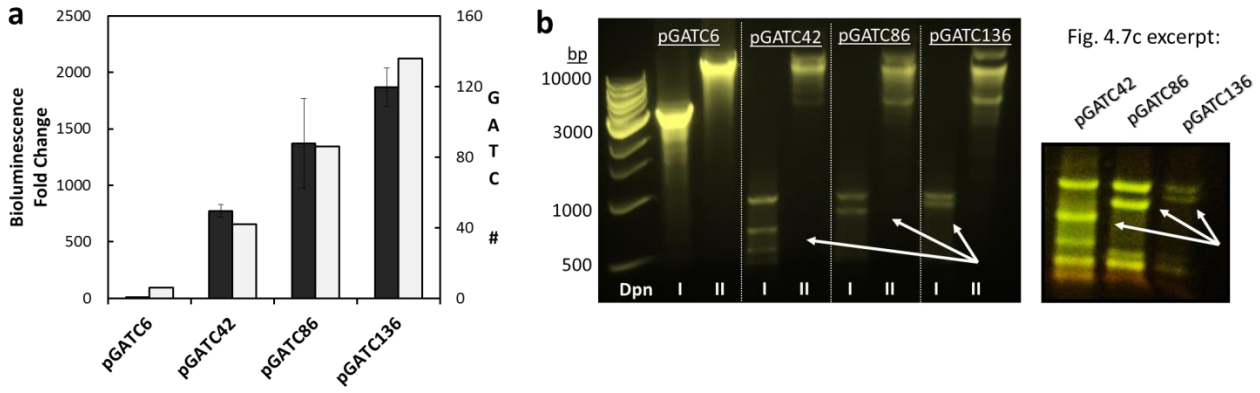


Figure 4.8 Delineation of autoinduced bioluminescent response from GATC inscriptions on plasmids and confirmation of AI-2 conversion based on plasmid band pattern. (a) Bioluminescence response to AI-2 amplitude programmed by GATC inscribed plasmids. (b) Post-reaction Dpn digest to scan plasmids for "barcode-like" identification and documentation of their methylation state.

4.3.9 Repetitive extension and pattern creation of GATC-rich templates

The results established that GATC inscriptions can be written into DNA in rich formats to delineate a message of AI-2 amplitude. AI-2 amplitude reflects the natural 'language' of quorum sensing communication. Here, defined AI-2 messages were relayed from DNA at

detectable levels by cells to elicit commensurate quorum sensing responses. However, we recognized that the aforementioned GATC-inscribed constructs were limited in their ability to be further enhanced with GATC content by standard cloning methods (noted by unsuccessful attempts to incorporate ABC into a plasmid). Therefore, we sought a different tactic to generate highly enriched, yet demethylated DNA with easily differentiated barcodes.

First, we demonstrated that the design of A, B, and C GATCblocks allowed for their further exploitation towards enhanced DNA length and GATC complexity. New primers were used to PCR-extend each GATCblock with altered overlaps for assembly (Fig. 4.9a(i)); specifically, the overlaps matched such that the complements of each block could be assembled in ABC order (Fig. 4.9a(ii)). Thus, the bodies of Arev, Brev, and Crev match in sequence structure to A, B, and C, yet their assemblies differ due to 5' and 3' end alterations. For example, by linking A + B, a 175 bp segment of A is linked to a 50 bp segment of B while Arev and Brev are structured such that the other 50 bp end of A becomes linked with the other B section of 250 bp. In this way, the GATCfree frames were linked in a new manner compared to ABC, resulting in a different band pattern upon erosion of the GATC sites. Furthermore, ABCrev, once compiled, was intended for assembly to the 3' end of ABC; this was designed such that cleavage by NruI at a 'pro-overlap' accesses a matching 3' overlap on ABC and 5' overlap on ABCrev to assemble the two together into a single 3000 bp length (Fig. 4.9a(iii)). As noted above, since ABC and ABCrev each hold a unique band pattern, their compilation results in a barcode that is specific to $(ABC)^2$ by encompassing that of each (Fig. 4.9a(iii)). GATC-erosion reveals the barcode and was illustrated by virtual DpnI digest for each incrementally assembled GATCblock length in Fig. 4.9a(iv), revealing unique barcodes dependent on GATCblock content.

Since the longest construct, $(ABC)^2$, did not PCR reproducibly, we strategized an alternative method for its procurement- rolling circle amplification. Rolling circle amplification (RCA) amplifies circularized DNA via isothermal polymerase action. That is, thermocycling is not necessary to regenerate primed single-stranded templates, as occurs for traditional PCR. Instead, because the template is not linear, the polymerase circles the template repeatedly, meanwhile continuing to extend the single product strand by DNA synthesis. Typically, phi29 DNA polymerase is used due to its superior polymerization rate, propensity for strand displacement, and relatively high fidelity [101]. However, one disadvantage of phi29 is its 3' endonuclease activity. To circumvent this counterproductivity, exo-resistant random hexamers- having two thiophosphate linkages in the 3' nucleotides (5'-NpNpNpNp^{s}Np^{s}N-3') - can be used for priming while protecting free ends from digestion [102]. Random hexamers also improve product yield compared to target-specific primers by hybridizing prevalently along the template; not only are the amplification of long templates (> 100 bp) enabled, the use of random hexamers has been shown to contribute to a 10,000-fold amplification of the template [102].

To expand RCA for diverse amplification purposes, the technique of padlock probing came about, intended for linear templates. Padlock probes circularize linear DNA by bringing together blunt ends to create a loop [103]. Here, a padlock primer was developed to anneal the GATCblock assemblies for initiation of RCA (Fig. 4.9b). Hyrbidization specificity to the template is accomplished by touchdown annealing, followed by the addition of exo-resistant hexamers to stabilize the ring with interspaced annealing. Then RCA conditions are applied, using phi29 polymerase, an excess of nucleotides, and pyrophosphatase to combat an RCA-inhibiting buildup of pyrophosphate byproduct. Polymerization products were analyzed by gel

electrophoresis in Fig. 4.9c. For all linear templates, the use of a padlock probe enabled amplification of severely long DNA chains, which presumably are GATC-dense.

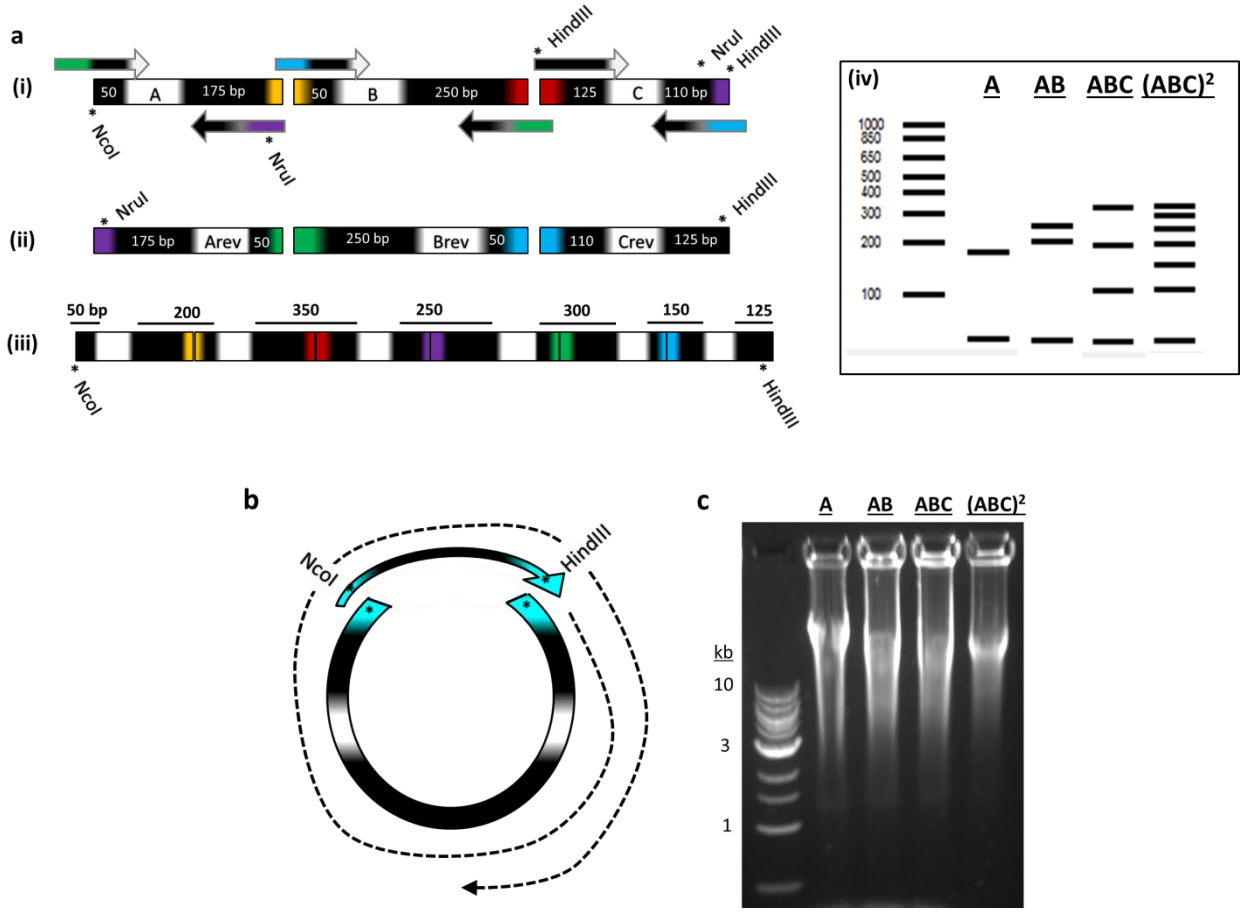


Figure 4.9 One-step amplification of unmethylated GATC-rich substrates towards continuous AI-2 flux. (ai) Alteration of GATCblock 5' and 3' ends through PCR primers with partial complementarity. (ii) Altered GATCblocks with "reversed" GATCfree and GATCrich ordering. (iii) Linear assembly of all GATCblocks into a ~ 3000 bp product with distinct GATCfree and GATCrich bands. (iv) Resolution of unique "barcodes" on a virtual gel, resulting from Dpn erosion GATCrich portions. (b) Use of a padlock primer to circularize a DNA template for rolling circle amplification (RCA). (c) Implementation of RCA to amplify long (> 10 kb), tandem-repeats of GATCblock templates of increasing template length.

4.4 Concluding Remarks

The concept of GATC inscriptions, paired with an enzymatic pathway for conducting methylation-delineated signaling, can serve as an innovative way to tune cell behavior, especially

since DNA methylation is an unexplored frontier for bioengineering. On the other hand, because it is well-understood for its essential role in natural cellular processes, the vast supporting biological evidence strongly correlates to our observations in the contrived *in vitro* system. A realism, being that traditional methylation research has sought only to understand that which is naturally relevant, is that EcoDam mechanistics have already been keenly studied but not yet applied. Our work introduces the opportunity to tune the DNA parameters via GATC inscriptions for deliberate, engineered methylation control.

This work integrates diverse biological components for sophisticated performance, components including plasmid DNA, coupled enzymes (some of which are covalently assembled as a chimeric protein), and engineered responsive cells. All varieties have merits for synthetic biology, yet are rarely used in combination. DNA is easily manipulated using commercial biotechnology kits for gene insertion, and here, for designing custom-methylation substrates. *In vitro* proteins, especially fusion proteins, are advantageous in their utility for clean and defined catalysis. Finally, engineered bacterial cells are highly customizable and robust biological machines. In our development of these diverse biological tools, we have highlighted piece-by-piece assembly for creating simple building techniques and interchangeable parts to facilitate an ease of customization.

This described work results in a smart, information-relaying system, where information is represented as the small-molecule signaling molecule AI-2, "understood" by bacterial cells, and the spatiotemporal nature of the AI-2 message can be pre-programmed into DNA to serve as a "script." Our take on this type of system is consistent with many current efforts toward infobiology, or the ability to store and transmit information using DNA and cells, respectively, as

biological strategies [13,25,30,34]. Here, this vision is further advanced by demonstrating that a signal can be autonomously derived from a DNA inscription and directly interpreted by a cell.

Chapter 5: Nano-guided cell networks as conveyors of molecular communication.

This chapter was primarily reproduced with permission from a paper recently accepted for publication: JL Terrell, HC Wu, CY Tsao, NB Barber, MD Servinsky, GF Payne, and WE Bentley, "Nano-guided cell networks as conveyors of molecular communication*," Nat Commun*, **2015,** *just accepted*.

5.1 Introduction

It has become increasingly apparent that a wealth of molecular information exists, which, when appropriately accessed, can provide feedback on biological systems, their componentry and their function. Thus, there is a developing niche that transcends length scales to concurrently recognize molecular detail and at the same time provide understanding of the overall system[104,105]. 'Omics' efforts have enabled tremendous access to molecular information through systems-level data mining. Meanwhile, an emerging, alternative scheme is to develop nano- to micro-scaled tools that intimately engage with biological systems through monitoring and interacting at the molecular level, with synthetic biology being one such tool[4,33,106-108].

While synthetic biology is often viewed as an innovative means for "green" product synthesis through the genetic rearrangement of cells, their biosynthetic capabilities and their regulatory networks can instead be tuned for executive function[2,10,16]. That is, cells can be rewired to survey molecular space[33,109,110] as they have sophisticated capabilities to recognize, amplify, and transduce chemical information[111]. Further, they provide a means to connect biological systems with traditional microelectronic devices and in doing so present a potential interface between chemically-based biomolecular processing and conventional vectors of

information flow, such as electrons and photons[112-114]. Specifically, through engineered design, cell-based molecular processing can be further coupled to enable external abiotic responses. Cells, then, represent a versatile means for mediating the molecular "signatures" common in complex environments, or in other words, they are conveyors of molecular communication [115-117].

Further, beyond clonal cell-based sensors, there is an emerging concept of population engineering to establish microorganisms in deliberate networks that enable enriched system identification through a combination of distinctive yet coexistent behaviors, including, perhaps, competitive or cooperative features[10,118-123]. We posit the use of cell populations assembled *in parallel*¸ where multiple microbes with distinct molecular recognition capabilities work congruently. An advantage is that populations, as opposed to few cells, can facilitate thorough sampling since the presence of many cells increases their spatial breadth and per-cell data contributions (Fig. 5.1a). Each cellular unit undergoes independent decision-making and contributes a datum to its entire constituency. The prevalence of data provided within the population, then, substantiates a collective output by the system based on the molecular landscape. As follows in a multi-population system, molecular input thus influences the outcomes of each population, and elicits plural responses when the molecular input ranges overlap the ranges of the sensing populations[119], which can define classification boundaries (Fig. 5.1a). Cell mediated classification was posited *in silico* by A. Didovyk *et al*, where reporter libraries with randomized sensitivities to a molecular cue elicit concentration-dependent fluorescent patterns and these are elucidated by population screening [119]. In the present construct, multiple populations enable multiplexed analysis, resulting, here, in a response gradation that is designed to index the molecular input "signature". Consequently, the fedback

information becomes transfigured beyond a dose-dependent cell-by-cell analysis. That is, the output is predicated by the comparison between the populations rather than accumulation of response within a total population.

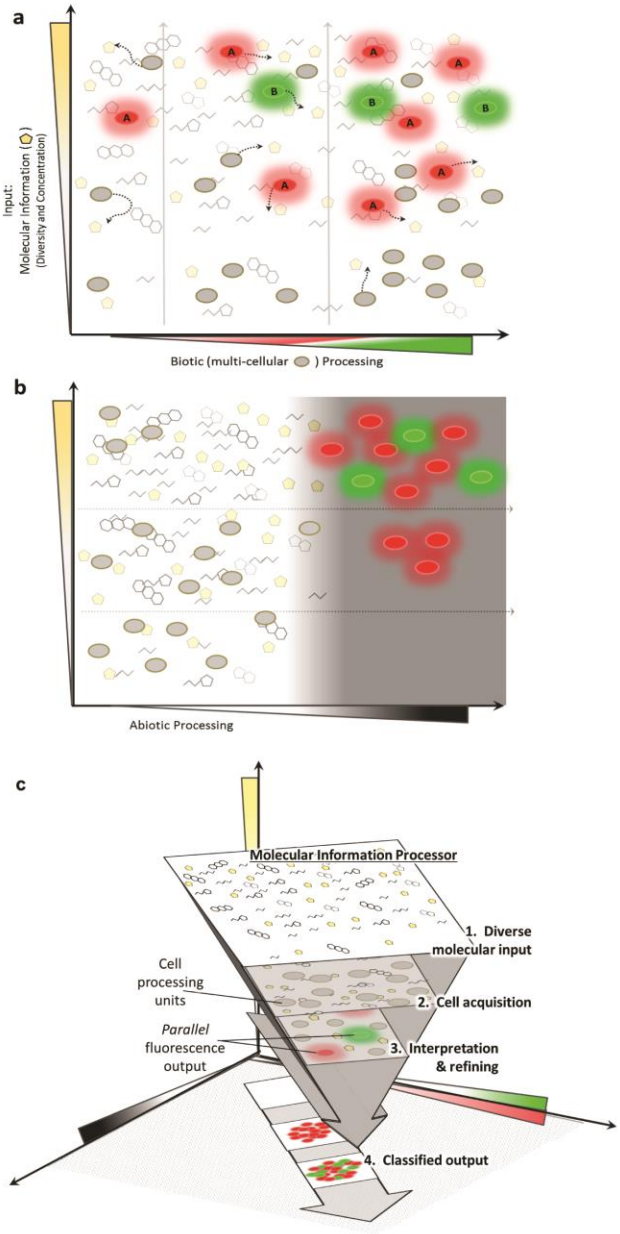


Figure 5.1 Multi-dimensional approach for processing molecular information. (a) Multicellular processing of molecular input is facilitated by molecular capture, cell processing, and response. The degree of response reflects the prevalence of the target molecule. Furthermore, the increase of biotic processing, by both increasing cell number in the population and utilizing multiple populations, enriches the output with multicellular data based on their tuned responses. (b) Abiotic processing, used in

conjunction with biotic processing adds another dimension to cell-based output, by modifying the response through a physical stimulus. (c) A molecular information processor exists as a network that is cell population-based with both biotic and abiotic processing mechanisms. Such a system can interpret molecular information by intercepting diverse molecular input, process it autonomously through independent cell units within the system, and refine the output to only include positive responders. The system's hierarchical structure allows molecular information to be refined into categorized collective output.

With population engineering as a premise for enriched molecular information processing, we engineered cell species, each to achieve an appropriate output through genetic means. There is conceptual basis for incorporation into networks, such as through mobile surveillance and position-based information relay[124,125]. Hence, it is conceivable that, in addition to autonomous molecular recognition and processing afforded by synthetic biology, the use of physical stimuli to enable cell response could confer similar networking properties[18,126]. For example, the complete information-processing "repertoire" can be expanded beyond specific cell responses by the integration of external stimuli that serve to collate cell populations[127]. Specifically, we envision integration of nanomaterials that enable co-responses to molecular inputs, such that cell populations employ traditional reporting functions, i.e. fluorescence marker expression, as well as responses that enable additional processing via the integration of stimuli-responsive abiotic materials (Fig. 5.1b).

In our example, cells are engineered to respond by permitting the attachment of magnetic nanoparticles, such that each fluorescent cell becomes receptive to a magnetic field. Thus, the combination of cell-nanoparticle structures provides further dimensionality for the conveyance of molecular information (via magnetic stimulation). That is, without magnetic collation the fully distributed system would harbor diffuse responses; a magnetically-stimulated system results in acute output due to a *filtering* and *focusing* effect (Fig. 5.1b) [128,129], allowing binned information to be readily, and fluorescently, conveyed.

The detection and interpretation of signaling molecules in our example is based on a microbial communication process known as quorum sensing (QS). The molecules, autoinducers, are secreted and perceived within a microbial community; once accumulated, the autoinducer level indicates that the population size has reached a "quorum" [32,130]. By surpassing a threshold concentration, the autoinducer signaling coordinates population-wide phenotypic changes[131]. We have designed a QS information processor that utilizes two cell populations to interrogate natural microbial communities and generate information about QS activity by accessing autoinducer-2 (AI-2) [132]. Importantly, AI-2 is a chemoattractant for *E. coli*, and hence *E. coli* engineered to sense and respond to AI-2 will naturally move towards its sources to fully sample the prevailing state[16,133]. Each strain evaluates AI-2 with a distinct sensitivity. When "activated" in response to a characteristic level, the cells simultaneously express a fluorescent marker and a streptavidin binding peptide (SBP) on the outer membrane, transported via signal peptide (for inner membrane translocation) and the outer membrane autotransporter pore AIDAc [134-137]. SBP provides a means for cell hybridization through its strong affinity to streptavidin, and here, aids in binding magnetic nanoparticles. This enables the non-genetically coded property of cell translocation within a magnetic field through physically-stimulated focusing and binning. Upon AI-2 recognition, response at the cell and population level includes (1) optical transmission at a unique, designated wavelength (red or green fluorescence, Fig. 5.1a), and (2) translocation within a magnetic field (Fig. 5.1b). Thus, with biotic processing by dual populations, and a post-processing magnetic sweep, the system as a whole interprets a molecular landscape and refines output into color-categorized, or "binned," states (no fluorescence, red, or red and green) through (1) parallel population processing and (2) acute focusing (Fig. 5.1c).

The use of engineered cells as data-acquiring units and selectively equipping each with functional nanomaterials to form a redistributable processing system merges two paradigms: decentralized, active probing at a molecular scale and self-organization of units through structured dependencies on stimuli [138]. The population-based system overall contributes categorized feedback about a biological environment.

5.2 Materials and Methods

Engineered strains. All plasmids and strains used in this study are listed in Table S1of the Supplementary Information. The vectors designed for this study, pSBP-Venus, pSBP-mCherry, and pSBP-eGFP, were derived from pAIDA-I, which was generously donated from Gen Larsson and previously used for covalent surface display of fusions up to 110 kDa in size[135]. The plasmids pSBP-Venus, pSBP-mCherry, and pSBP-eGFP were constructed as described in the Supplementary Information using primers listed in Table S2 and with plasmid maps illustrated in Scheme S1. Plasmids were transformed into chemically competent BL21(DE3) *E. coli* (Life Technologies) for testing T7-regulated expression of the surface display fusion SBP-fluorescent protein-AIDAc. Next, the plasmids were introduced by electroporation into the electrically competent strains CT104 (+ pCT6) and MDAI2 (+ pCT6). Strains were made competent by standard procedures.

Protein expression. Chemically competent BL21(DE3) cells (Life Technologies) were transformed with pSBP-Venus, pSBP-eGFP, or pSBP-mCherry. Cultures were grown to mid-exponential phase, then induced with 500 μM isopropyl B-D-1-thiogalactopyranoside (IPTG), purchased from Sigma (USA). The induced cells were incubated at 37 °C, 250 rpm shaking, for 6 h. Alexafluor488 and Alexafluor594-labeled streptavidin (Life Technologies) were prepared

to a working concentration of 20 μg/mL in 10 mM PBS. Dylight405-labeled streptavidin was prepared from a ThermoScientific labeling kit and diluted to a working concentration of 500 μg/mL. Culture aliquots were washed once in PBS, centrifuged (4000 g, 5 min), resuspended in the fluorescent streptavidin solution, and labeled for 1 hour at room temperature. Finally, cells were again washed in PBS and resuspended for imaging. Imaging parameters for each fluorophore were consistent for each composite.

Cell culture. *E. coli* cultures were prepared by 1 % reinoculation into LB medium from an overnight culture and grown at 37 °C, 250 rpm shaking, until mid-exponential phase, at which point experiments for triggered expression were initiated. Antibiotics were added according to the plasmids contained by the strain. Ampicillin was used at 100 μg/mL for pCT6 transformants and kanamycin at 50 μg/mL for pSBP transformants. Conditioned media were isolated from W3110 *E. coli* by pelleting culture aliquots and filter-sterilizing the supernatant. *L. innocua* cultures were inoculated into a brain-heart infusion (BHI) medium and grown at 30 °C, sampling at cell densities specified in results. For cocultures with *L. innocua*, all cell types were centrifuged at 4,000 g for 5 min and resuspended at the same density into antibiotic-free medium, varied according to the experiment. To initiate cocultures, the reporter cells CT104 (Type 'A') and MDAI2 (Type 'B') were added to *L. innocua* at 1 % each to a 48-well plate. Cocultures were then grown at 30 °C, 220 rpm shaking, overnight for 12-14 h. *L. innocua* was prepared for coculture by isolating the strain at OD intervals during growth and resuspension. For coculturing with CT104, *L. innocua* was grown to OD 0.3 and then resuspended in diluted media: 2.5 % BHI (7.5 % LB, 90 % PBS); in addition to OD 0.3, cells were diluted 10 and 100fold before coculturing. For coculturing with MDAI2, *L. innocua* was grown to OD intervals

between 0.05 and 0.6 and the media was replaced with 25 % BHI in LB. Initial cell number ratios were obtained by colony count calculations from plated cells.

Flow cytometry. Fluorescent cell counts for SBP-Venus reporters were obtained using a FACSCanto II flow cytometer (Becton Dickinson), the 488 nm laser and 530/30 green filter, and BioFACS Diva software. Histogram images **(**Supplementary Information, Fig. S5c) were obtained using FlowJo software (Tree Star). Aliquots from culture were diluted 4-fold in PBS and counts of at least 30,000 were recorded for each sample. Specifically, 50,000 events were counted in the analysis for Fig. 4b and gated for Venus fluorescence at an intensity of 2×10^3. Fig. S14 results were based on subpopulations gated above an intensity of 1×10^3 out of 50,000 events.

Confocal imaging and analysis. Fluorescence imaging utilized a Carl Zeiss (Jena, Germany) LSM700 confocal microscope. Microscopy images were processed and analyzed using ImageJ (National Institutes of Health). Profiles of fluorescence mean gray values were generated for image slices with approximately one cell per vertical space. Alternatively, the brightness threshold was adjusted uniformly across images, after which a particle analysis was used to count cells. To distinguish cell type within coculture images, a combined analysis counted particles based on size threshold and ratio of green to red intensity. This analysis is described in the Supplementary Information and depicted in Figs. S9 and S10.

Scanning electron microscopy. Induced BL21(DE3) + pSBP-venus cells and streptavidin-conjugated magnetic nanoparticles (mNPs) were prepared separately as well as mixed together to isolate cells with surface-bound mNPs. MNP-decorated cells were isolated by co-incubating for 20 min on ice, then magnetically recovering a pellet, which contained observable cells. Then

samples were fixed and dehydrated. After each step, cells were pelleted by centrifugation at 3000 g, 4 °C, 5 min. Magnetic nanoparticles and cell-mNP mixtures were pelleted magnetically. Samples were first washed twice with cold PBS, then incubated at room temperature with 4 % paraformaldehyde (Sigma-Aldrich) in PBS. Subsequently, the samples were washed once with PBS and twice with DI water. Next, gradual replacement with ethanol was performed by sequential 10 min incubations on ice with 30, 50, 70, and 100 % ethanol in water. Finally, ethanol was gradually replaced with hexamethyldisilazane (HMDS, Sigma-Aldrich) by 10 min incubation in 50 % HMDS solution in ethanol followed by two incubations with 100 % HMDS. Samples were ventilated on a coverslip in a chemical fume hood for 30 min, and then further dried at 37 °C for 1 h. Prior to imaging, the coverslips were sputter-coated with carbon using a MED 010 evaporator (Balzers Union, Liechtenstein). Sample imaging and elemental mapping used a SU-70 UHR Schottky field emission SEM equipped with an energy-dispersive x-ray spectrometer (Hitachi, Tokyo, Japan).

Magnetic-activated cell focusing. Magnetic nanoparticles (mNPs, 100 nm, fluidMAG-Streptavidin) (Chemicell, Berlin, Germany) were washed in cold PBS and resupsended at the same density. For magnetic enrichment of multi-species surveillance systems, mNPs were applied directly to overnight cultures at 2 % of the culture volume. After incubating the cells and mNPs in a 96-well plate at 4 °C for 20 min, a 2.2 x 1.6 cm (dia. x ht.) neodymium N42 magnet (K and J Magnetics) was placed under each well for 5 min to collect mNPs on the bottom surface. The collected pellet consisted of magnetically-coupled contents; the supernatant was removed and the sample was resuspended in a minimal volume (~ 2 µL). To image the particle-coupled contents within a magnetic field, a 1.6 x 0.8 mm (dia. x ht.) N52 neodymium magnet was taped behind a coverslip. 2 µL of suspended mNPs and cells bound with mNPs was added

to the opposite side of the coverslip, directly on top of the magnet, set for 2 min, after which another coverslip sealed the sample for imaging at 200x magnification. Characterization of streptavidin-coated magnetic nanoparticles for SBP-surface-expressing cells is provided in the Supplementary Information.

5.3 Results

5.3.1 Coexpression of surface-localized SBP and fluorescent marker protein

First we established expression of a fusion protein consisting of a fluorescent marker (eGFP and variants) and SBP. Importantly, for SBP to function as a coupling agent between cells and magnetic nanoparticles, we used AIDAc (kindly shared by Jen Larssen)[135] to export the chimeric protein to *E. coli's* outer surface. In Fig. 5.2a, we depict expression of three different constructs using Venus, eGFP, and mCherry for optical transmission, and an outer-membrane autotransporter (AIDAc translocator domain) for surface localization. These constructs are mapped in the Supplementary Information (Fig. S1). After induction with isopropyl B-D-1-thiogalactopyranoside (IPTG), cultures were probed for surface expression of the SBP portion of the tagged fluorescent protein. Cells were incubated with fluorescently-labeled streptavidin; the fluorophore of the streptavidin probe was orthogonal to the expressed fluorescent protein. The multiple fluorescence emissions were analyzed by confocal microscopy without spectral overlap. The fraction of cells (f_c) that exhibit colocalized fluorescent protein and the streptavidin-tagged antibody is reported in Fig. 5.2b, showing that SBP-Venus cells bound streptavidin at a slightly lower frequency than SBP-mCherry and SBP-eGFP which exhibited statistically similar fractions ($f_c = 0.7$).

That is, microscopy results related to the colocalization analysis are depicted for pairings of Venus and blue-streptavidin (-SA), eGFP and red-SA, and mCherry and green-SA (Fig. 5.2c). Strong signals were observed in both filter sets (the fluorescent protein (Column I) and the labeled streptavidin (Column II)). Overlaying each image reveals colocalization, as indicated in Column III, where arrows point to examples of strong colocalization. Additionally, Column IV plots fluorescence intensities across horizontal sections of the images, where cells that exhibit colocalized fluorescence are indicated by superimposed peaks. For + pSBP-Venus cells, those with both a blue and yellow signal are observed as pale blue-violet in the overlaid image. Cells with +pSBP-eGFP and +pSBP-mCherry and labeled streptavidin emit both green and red signals; their colocalization appears yellow. Controls in the Supplementary Information (Fig. S2) verify that fluorescent streptavidin (all colors) has specificity for only SBP-expressing cells over negative controls. Colocalization indicates that not only are both components of the fusion, SBP and the fluorescent protein, expressed, but that SBP is accessible to bind streptavidin on the cell's surface. Also, the cell surface was immunostained for the fluorescent protein; eGFP was tested in particular and was positively tagged, indicating surface exposure (Supplementary Information, Fig. S3). Also in Fig. S3, confocal microscopy confirms precise localization of the eGFP and Alexafluor594-labeled antibodies within the confines of individual cells. This is the first use of AIDAc for cell surface anchoring of fluorescent proteins, each having been functionalized with an affinity peptide.

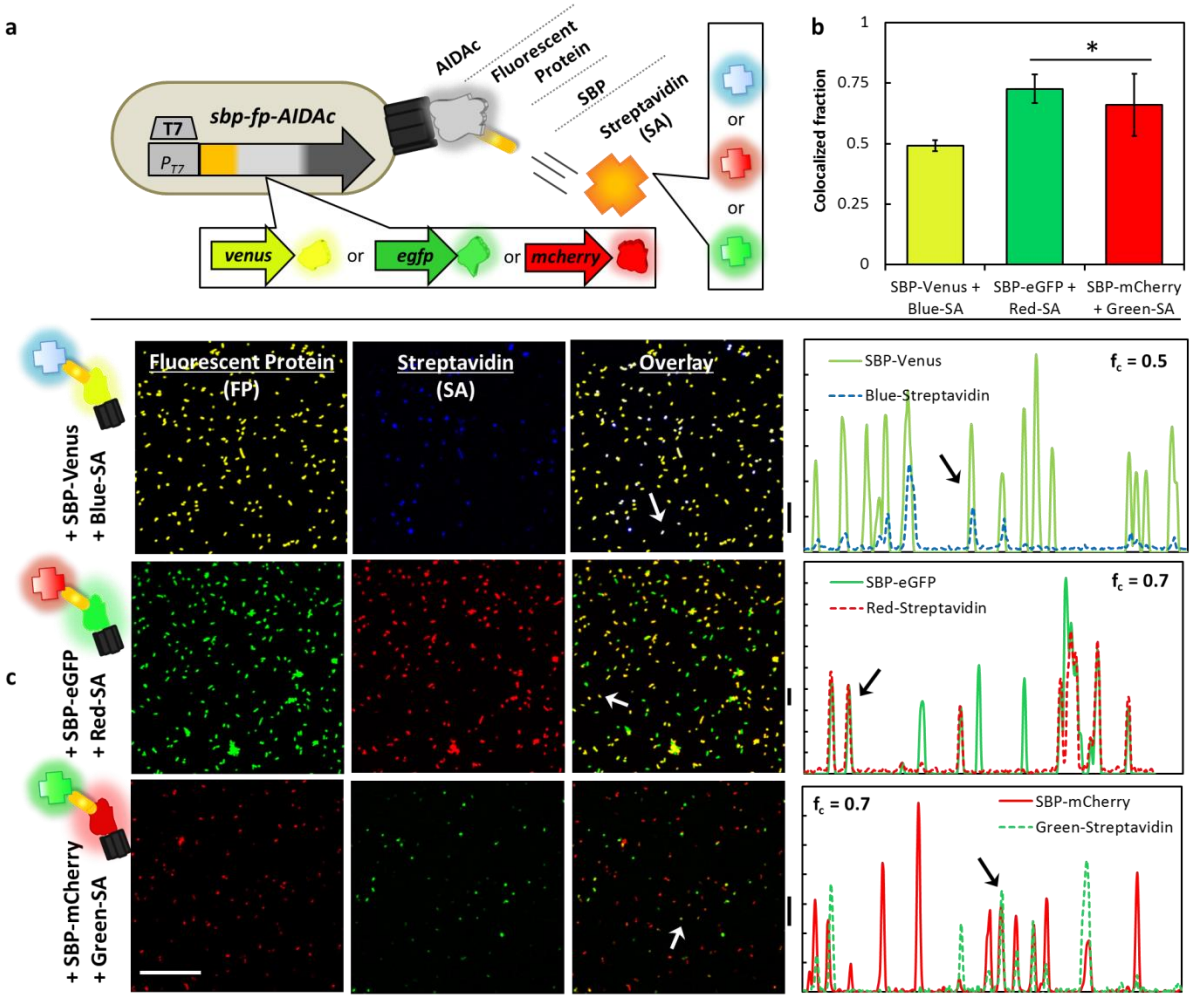


Figure 5.2 Cells express functional, interchangeable protein components for combined fluorescence and streptavidin-linked surface coupling. (a) A T7 cassette was used to express a chimeric protein consisting of a membrane autotransporter domain (AIDAc), one of several fluorescent protein options, and a streptavidin binding peptide (SBP). Fluorophore-tagged streptavidin (SA) was used to bind SBP and probe for surface expression. (b) Colocalized fractions(f_c) of cell population for each fluorescent protein (FP) construct (streptavidin-stained fluorescent cells as a fraction of total cell number expressing FP). * indicates that f_c for +SBP-eGFP and + SBP-mCherry is statistically greater than +SBP-Venus cells, but similar to each other. (c) Composites show cell fluorescence in Column I from the fused fluorescent protein (FP); in Column II, from labeled streptavidin; and an overlay of both in Column III, where arrows point to representative cells with strong colocalization. Plotted in Column IV are the fluorescence mean gray values (y-axis) across a representative horizontal slice of the composite image (x-axis). Vertical bars between Columns III and IV identify the position of each analyzed slice. Arrows indicate peaks that match the highlighted cells in Column III. Inset are the f_c values. Fluorophores with non-overlapping spectra were paired. Row 1, Venus expression (yellow-green) was paired with Dylight405-labeled SA (blue). Row 2, eGFP expression (green) was paired with Alexafluor594-labeled SA (red). Row 3, mCherry expression (red) was paired with Alexafluor488-labeled SA (green). Scale bar in lower left indicates 50 μm.

5.3.2 Cell hybridization with magnetic nanoparticles through affinity interaction

Given that expression of a fluorescent protein tagged with SBP enabled external binding of streptavidin, we employed this interaction for fastening streptavidin-functionalized materials directly to the cell surface. We chose streptavidin-conjugated magnetic nanoparticles (mNPs), 100 nm in diameter (an order of magnitude smaller than a cell), for binding to a cell surface (Fig. 5.3a) in order to impart the abiotic magnetic properties. Scanning electron microscopy (SEM) was used to observe surface interaction between cells surface-expressing SBP and streptavidin-functionalized mNPs. In the Supplementary Information, Fig. S4a-b shows electron micrographs of *E. coli* cells (dimensions 1.5 – 2 μm in length) and the mNPs (approximately 100 nm in diameter). The SEM image in Fig. 5.3b, shows a magnetically isolated SBP-expressing cell with streptavidin-mNPs. The sample was prepared by mixing SBP-expressing cells with streptavidin-mNPs, then collecting or "focusing" into a magnetized pellet via magnetic field, then separating from unbound cells in the supernatant. The cells were then washed and resuspended. In Fig. 5.3b, clusters of surface-bound mNPs are observed. Additionally, the elemental composition was analyzed with energy dispersive X-ray spectroscopy (EDS), shown in Fig. 5.3c by an element map superimposed with carbon (red) and iron (green). While the cell appears to be of a uniform carbon composition, the particles localized at the cell surface (highlighted with arrows) were found having a strong iron composition; thus, elemental analysis confirmed particle identity as iron oxide mNPs. Additional characterization of magnetic functionality, including detailed SEM and fluorescent microscopic analysis prior to and after application of magnetic fields, is described in the Supplementary Information (Fig. S4).

In sum, the well-known affinity interaction between streptavidin and the peptide SBP is harnessed to endow cells with non-natural abiotic properties. Here, coupling a functionalized nanomaterial to the surface-displayed peptide physically extends the fusion protein and also adds physical (magnetic) functionality to the cell.

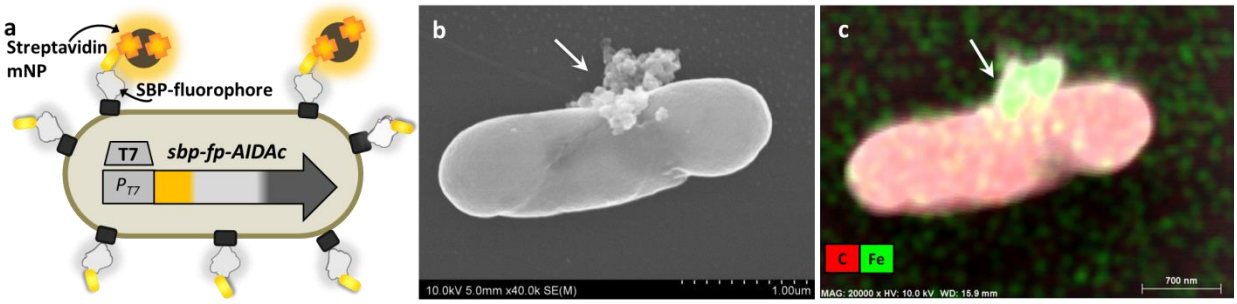


Figure 5.3 Cells equipped with magnetic nanoparticles (mNPs) through selective interaction of streptavidin with surface proteins. (a) Cell surface binding of streptavidin-conjugated magnetic nanoparticles occurs by binding to cell surface-anchored streptavidin binding peptide (SBP). The fusion of T7-expressed SBP-fluorescent protein (FP)-AIDAc enables the protein's surface accessibility. (b) Scanning electron micrograph of an *E. coli* cell with surface-bound particles. (c) Element map of carbon (red) and iron (green) through energy dispersive spectroscopy.

5.3.3 Linking SBP-fluorophore expression to autoinducer-2 recognition

The expression system for pSBP-Venus was then put under autoinducer-2 control so that the protein is expressed in the presence of AI-2 instead of IPTG. That is, we coupled the native QS signal transduction circuitry to the reporter cassette. In order to ensure ample expression (as the native operon is fairly weak), we placed expression of T7 RNA polymerase under control of the natural quorum sensing circuitry.[139] Phosphorylated AI-2 activates the system through derepression of the regulator LsrR, naturally upregulating AI-2 import and phosphorylation,[140] and, by design, the T7 RNA polymerase on a sensor plasmid.[139] When *sbp-venus* is included downstream of a T7 promoter region on a second plasmid, expression is then triggered by AI-2 uptake (Supplementary Fig. S5a). Then, we used two host sensor strains engineered to provide

varied AI-2 sensitivity (denoted Responders "A" and "B"). In "A", *lsrFG*, genes required for internally phosphorylated AI-2 degradation[141,142] are deleted. Also, both strains lack the terminal AI-2 synthase, *luxS*, so they cannot produce AI-2 and, instead, must "receive" AI-2 from an external source (Fig. S5a). The phenotypic difference between A and B is the threshold level of AI-2 that activates the genetic response[85,143]. Fully constructed, these cells are designed to take up and process AI-2 in order to generate fluorescence output (that co-functions with streptavidin-binding). Efficient transport of this functionality to the membrane under AI-2 induction was demonstrated in each host (Supplementary Fig. S6). Also, this process did not alter growth kinetics for either cell type (Fig. S5b).

5.3.4 Establishing parallel ranges of cell interrogation for quorum sensing information

Importantly, the engineered cells each provide a characteristic response to the level of autoinducer-2. Recently, we showed that AI-2 level influences the quorum size of responding engineered populations but does not alter the expression level within each quorum[85]. Here, we evaluated our engineered AI-2 responders, again for quorum size (or in other words, percentage of AI-2-responsive cells in the population), this time varying the compositions of molecular input and the configuration of responders (Fig. 5.4a). First, we added AI-2, synthesized *in vitro*, to each of the two responder populations (Fig. 5.4b). We also added conditioned medium (CM), the spent medium from an AI-2 producer culture containing metabolic byproducts, as well as AI-2[132,144] (Fig. 5.4c). We also mixed the responder populations and added AI-2 to gauge responses in complex cultures (Fig. 5.4d).

Specifically, in Fig. 4b, A and B populations were incubated at mid-exponential phase with *in vitro*-synthesized AI-2[88,91] at concentrations: 0, 2, 10, 28, and 75 μM. After 12 h, samples

were observed for fluorescence by confocal microscopy and then quantified by FACS (Fig. S5c in the Supplementary Information). We found SBP-Venus expression for responder A cells occurred at the lowest tested level (2 μM AI-2), where 56 % of the population expressed SBP-Venus and this fraction increased rapidly with AI-2 reaching a maximum of 90 % at 28 μM. For type B, a more gradual trend was found; only ~1% were fluorescent from 0-2 μM, and this increased from 9 to 46 % as AI-2 was increased to 28 μM. Finally, the highest fraction of fluorescing cells was found at the highest concentration tested, 75 μM.

We next isolated conditioned media (CM), which contains a dynamic composition of unfiltered metabolites and media components, from W3110 *E. coli* cultures at intervals during their exponential growth, throughout which AI-2 accumulates (AI-2 levels for the samples are indicated in the Supplementary Information, Fig. S7). CM aliquots were mixed with either A or B cells and cultured in triplicate for 12 h. Through FACS analysis it was found, again, that a larger subpopulation of A expressed Venus compared to population B at any concentration (Fig. 5.4c). There was no statistically relevant expression from B until the CM samples from 0.23 OD cultures. In all cases, population A recognized AI-2 presence, including from media isolated at a W3110 OD of 0.05, the minimum cell density tested in this study.

The sensitivities of both strains to AI-2-mediated induction corroborate previous literature[16,85]. These trends demonstrate that strains engineered for altered sensitivity to molecular cues provide discrimination of concentration level. That is, the identical plasmid expression system was transformed into different hosts, providing robust and distinct levels of expression.

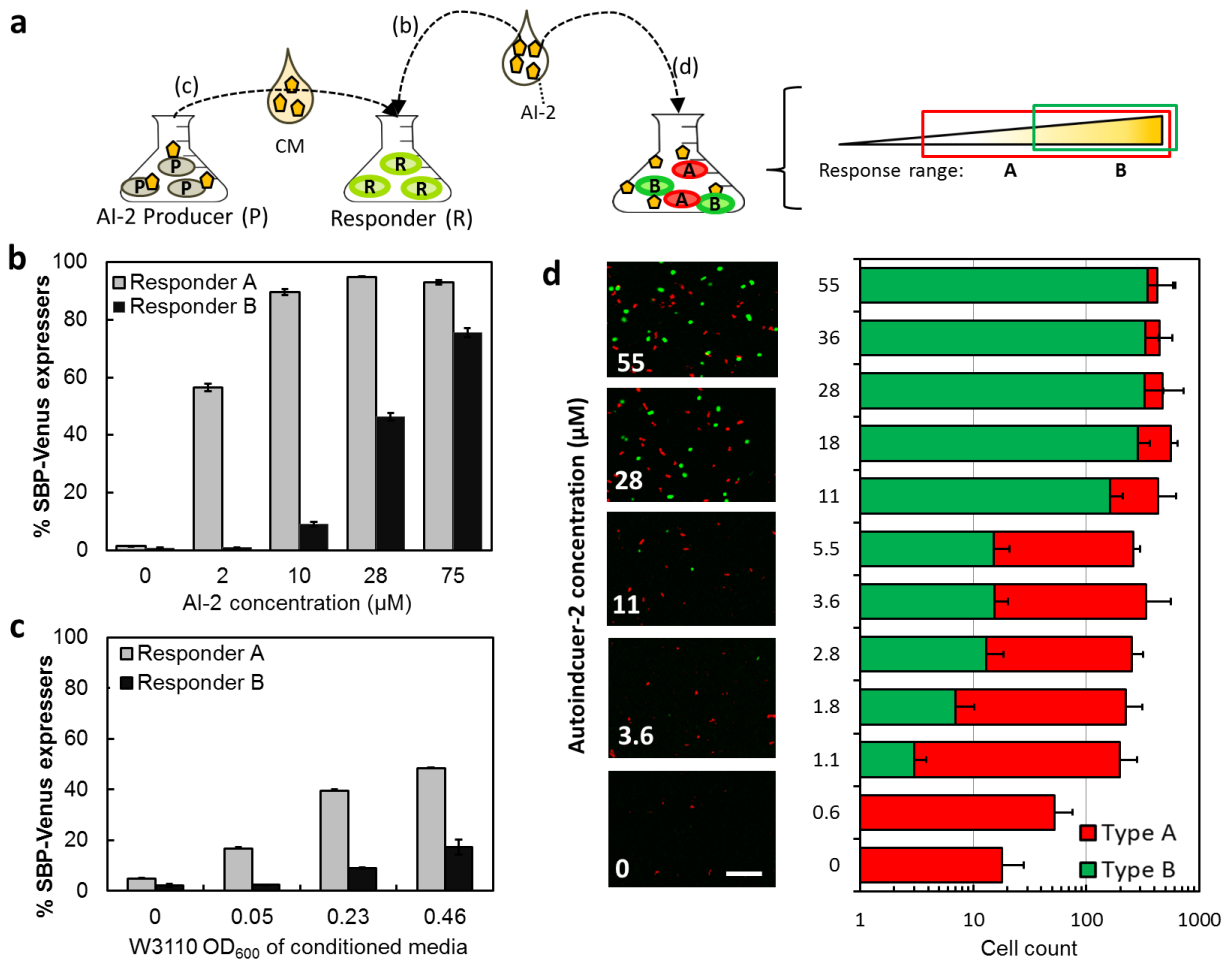


Figure 5.4 Population response to autoinducer-2 (a) from purified and raw materials sources, where fluorescence output is linked to small molecule input. Fluorescence from type A and B responders, with altered genetic responses, was analyzed after exposure to autoinducer-2 (AI-2) in mono and mixed culture environments. (b) Venus expression from *in vitro* synthesized AI-2. (c) Venus expression from conditioned media (CM) isolated from growing cultures of W3110 *E. coli*, sampled at indicated OD. Data are averages from triplicate cultures with standard deviations indicated. (d) Red and green fluorescence responses to AI-2 during co-incubation of type A (pSBP-mCherry+, red) and type B (pSBP-eGFP+, green) cells. Representative fluorescence images show colocalization of red and green cells. Scale bar indicates 10 μm. The average cell count per responder cell is plotted against AI-2 concentration with standard deviations, as determined by image analysis in quadruplicate.

Having developed cell types A and B with differential ability to detect AI-2, we next altered the reporters so that each cell type expressed a unique SBP-fluorescence fusion for color-coded designation. Cell type A was engineered with pSBP-mCherry and type B with pSBP-eGFP resulting in red and green fluorescence, respectively. These populations were mixed together in equal proportion at mid-exponential phase, introduced to a range of AI-2 concentrations, and incubated overnight. Populations A and B exhibited equal growth rates when cultured alone and together (Supplementary Fig. S8c); it followed that the cocultures should comprise a 1:1 ratio of each constituent. Fluorescence output is shown by representative images in Fig. 5.4d. Also in Fig. 5.4d, the green and red cell count is plotted from a quadruplicate analysis for each input concentration.

Coculturing enables parallel processing as the molecule-rich environment is perceived by each cell, and is processed uniquely per cell type. Yet, since each sensing mechanism is a living and proliferating population, we tested whether the potentially altered dynamics of coculturing would permit the same sensitivities as isolated culturing. We evaluated the Monod-type saturation constant for each population independently and in co-cultures. We found, in Fig. 5.4d, the general trends in response to an increasing AI-2 level were as predicted by modeled response curves (Table S3), which were also well-correlated to Fig. 5.4b data (Supplementary Fig. S8a-b). That is, the saturation constants that describe dependence on AI-2 were unchanged when measured in co-cultures. Phenomenologically, as expected, an initial accumulation of red type A responders was found. Then, at higher AI-2 levels, we found an emergence of a green subpopulation (type B). Above 28 μM, there was no longer an apparent differential response that would otherwise enable discrimination of AI-2 concentration; based on the consistency with modeled behavior, coculturing contributed to this dampening effect. However, the overall

fluorescence output is enriched by the combination of multiple populations since the ranges of sensitivity overlap and effectively expand that of the master population (Supplementary Fig. S8d). Specifically, because the fluorescence of B is described by a larger saturation constant, its fluorescence continually increases at higher AI-2 concentrations, while the fluorescence of A remains unchanged. Thus, coculturing between A and B enables resolvable output that is lower than the detection limit of B (due to A) yet also surpasses the upper limit at which A saturates by the inclusion of B. The choice to fluorescently differentiate A and B was important because the output would otherwise be biased by extracellular components including the existence of non-sensing cells. Due to color designation of A and B, a color "pattern" emerges as a feature of the parallel response which we recognize is independent of the absolute fluorescence of the population.

5.3.5 Consensus feedback by combining biotic and abiotic processing

We hypothesized that the value of cell based sensing would be enhanced if the cell output could be collated in an unbiased manner that in turn, were easily "read" using optical means. We engaged magnetic processing, which represents an abiotic processing step that enhances the signal by focusing the collective response. Hence, cells were equipped with streptavidin-conjugated mNPs (Fig. 5.3). The ability of a magnetic field to refine fluorescence output through filtering and focusing is described in the Supplementary Information (Fig. S11). Thus, in our combinatorial approach, fluorescence feedback about molecular information within a microbial community entails biotic processing through constituencies of two independent cell types in conjunction with magnetic post-processing that is enabled by guidance at the nanoscale (Fig. 5.5c). Moreover, since the fluorescence feedback data is provided through two constituencies, consensus from each independently provides an aggregate output; in our

example, the output becomes relayed as a distinctive "binned" category due to finite color-combinations generated from constituencies A and B (Fig. 5.5c).

Again, type A transmits red output (SBP-mCherry $^{+}$) and type B transmits green (SBP-eGFP^{+}). These were first co-incubated with titered concentrations of AI-2, to obtain results similar to those of Fig. 5.4d. By coupling magnetic nanoparticles to the responsive parallel populations, we tested for aggregate two-color output to provide informative feedback within a set of outcomes ranging from no color, red only, to red + green. After overnight coincubation and a magnetic sweep with streptavidin-mNPs, fluorescence results are shown in Fig. 5.5a where the recovered cells are displayed above a magnet's center in order from highest to lowest AI-2 level (top left to bottom right). The processing output generated by the range of conditions was quantitatively assessed for contributions from A and B responders. The spatial density of each fluorophore, or the area occupied by fluorescent responders as a percentage of total visible area, was quantified and plotted in Fig. 5.5b. Here, the trend of increasing fluorescence with AI-2 is followed by both A and B cell types; however, red A cells accumulate at a higher rate than green B cells. This relationship between A and B processing is not only consistent with their previous characterizations (Fig. 5.4) but indicates that the aggregate output is unbiased regardless of assembly with magnetic nanoparticles and magnetic-stimulated redistribution (Supplementary Information, Fig. S12a).

Next, A and B cells were added together to probe the quorum sensing environment of *Listeria innocua*, an AI-2-producing cell type that is genetically and ecologically similar to the pathogenic strain *Listeria monocytogenes*[145]. The environment was biased toward low and high cell density conditions by altering nutrient levels to develop contrasting scenarios of AI-2 level. Preliminary characterization in the Supplementary Information indicated that *L. innocua*

proliferation is unperturbed by the presence of *E. coli* responders (Fig. S13) and that type A cells detect AI-2 at low cell *Listeria* densities limited by sparse nutrients; then with rich nutrient availability, cell proliferation permits a higher AI-2 level that can be detected by type B (Fig. S14). Replicating these conditions, we expected red fluorescence to be observed at low culture density and for green fluorescence to be reported when high (Fig. 5.5c). Two conditions were tested: *L. innocua* was proportioned to responder cells at 20:1 in dilute media to establish a low culture density condition or alternatively, a ratio of 200:1 in rich media for a high culture density condition. After overnight coincubation and a magnetic sweep (applied directly to the triple strain cultures) with streptavidin-mNPs, the recovered cells are displayed above a magnet's edge (shown in Fig. 5.5d). Acute focusing of the fluorescence signals, contributed by each subset population of the processor (A and B), is visually apparent. The magnetic field had a physical effect of repositioning the "on" subsets to be tightly confined within the magnetic field.

The processing output generated by the contrasting culture conditions was again assessed for the respective contributions of A and B, and for changes in spatial signal density due to the magnetic sweep (Fig. 5.5e). The analysis was based on images provided in Supplementary Information (Fig. S12b). Data in Fig. 5e indicates that red type A cells are prevalent regardless of culture condition (except negative controls). However, compared to the low AI-2 condition, the abundance of green cells is 100-fold higher in the high AI-2 condition. Additionally, the ratio of green to red was consistent prior to and after magnetic concentration, substantiating observations in the distributed system. Further, data shows that magnetic refining increased per-area fluorescence 100-fold or 10-fold in low and high cell culture studies, respectively.

Based on the thresholds established for responder populations A and B, we found color-coded binning corresponded to AI-2 level, where "red-only" represented less AI-2 than "red + green"

(Fig. 5.4d). Thus, we found a binned output was established via this multidimensional molecular information processing system and that this matched the expectations. Red feedback (from Responder A) indicated dilute AI-2 accumulation occurred in the low density culture. In the dense cultures, high AI-2 accumulation turned on both A and B for combined red and green feedback.

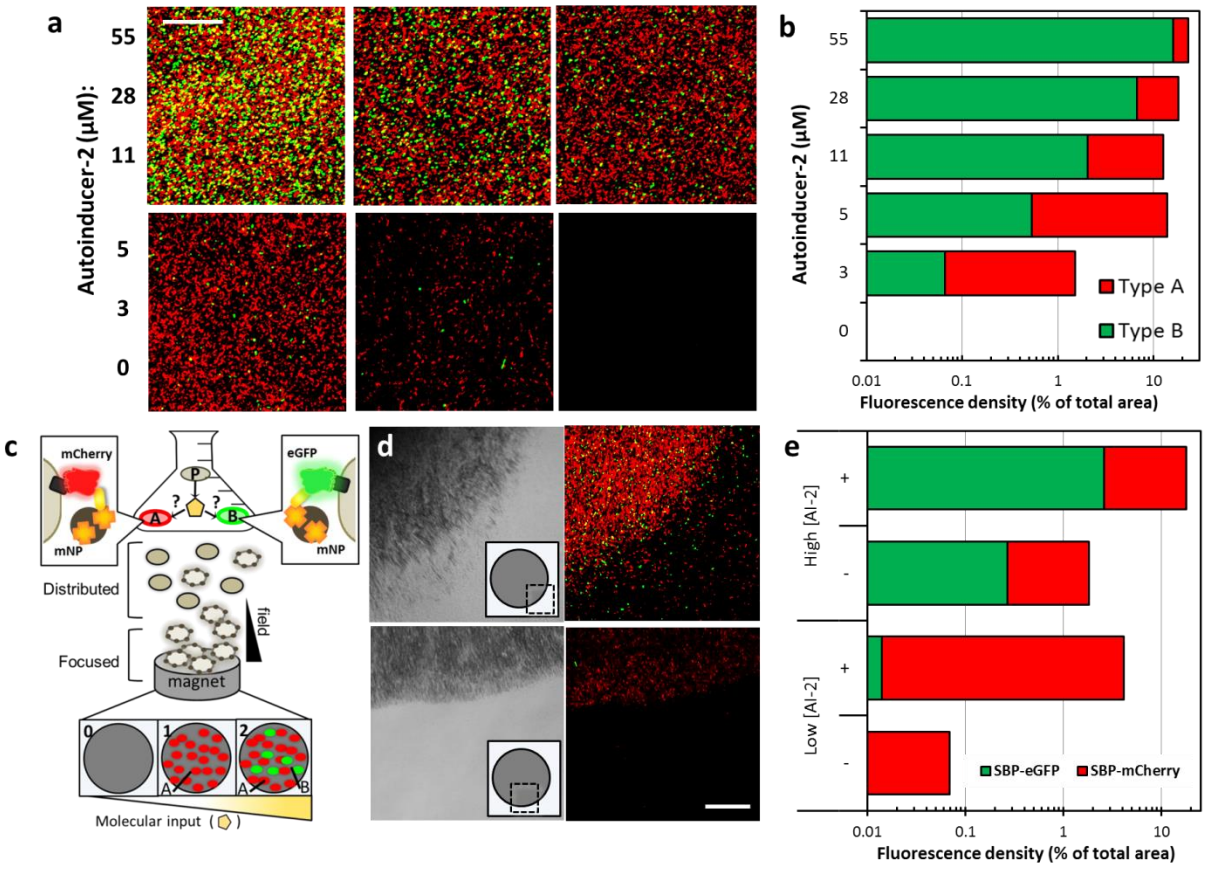


Figure 5.5 Binning molecular information through cell-based parallel processing and focusing the fluorescence into collective consensus output. (a) A and B cell types were coincubated with AI-2 levels ranging from 0 – 55 μM AI-2, then imaged after magnetic nanoparticle coupling and magnetic collation. Fluorescence results (centered directly over the magnet) are shown from high to low input (top left to bottom right). (b) Quantification of red and green fluorescence densities per AI-2 level. (c) The process of accessing molecular information begins by distributing type A and B cell processing populations within the environment of an AI-2 producer. A and B independently express fluorophore fusions, linked with available nanoparticles, upon processing autoinducer-2. Magnetic focusing translocates fluorescing responders only. Imaging the magnetically collated cell aggregate reveals classified fluorescence output representing the AI-2 composition of the sampled environment. (d) Brightfield (left) and fluorescence

(right, red and green filters) images positioned over the edge of a magnet, as defined by the inset. The sample in the bottom image pair was isolated from an environment of low AI-2 accumulation. The sample in the top image pair was isolated from a high AI-2 accumulation environment. (e) Quantification of visual space occupied by collated cells (eGFP and mCherry expressers) while distributed and magnetically focused. Scale bars indicate 50 μm.

5.3.6 System response patterns delineated through the relationship between parallel populations

Our example demonstrates the concept of an amorphous processing system that utilizes several biotic and abiotic components for multi-dimensional information processing. Interestingly, a binning effect was enabled: our system yields an index of color-categorized feedback that characterizes the sampled environment. In Fig. 5.6, we present a means to extend our approach to multidimensional systems, those with more than one molecule of interest and at different concentrations. That is, by appropriate design of the cell responders we can further enrich the methodology, its depth and breadth of applicability. We depict 10 hypothetical pairs of responses – those that can be driven by appropriately engineering cells to portend altered genetic responses. For example, rows 1 and 3 provide genetic outcomes as a function of analyte (AI-2) concentration. The hypothetical depictions are feasible as "designer" signal transduction and marker expression processes enabled by synthetic biology[12,34,119]. Rows 2 and 4 demonstrate the corresponding visual planes, where red cell numbers (x-axis) are plotted against green (y-axis), illustrated by the first example (i). If one divides the two dimensional space into quadrants (no color, majority red, majority green, and equivalent ratios of red and green), it becomes apparent that the relationship between cell types influences the "visual" or optical output. Thus, the 10 arbitrary response sets yield a variety of pairings that can provide unique visual patterns for categorizing molecular information. We have simplified the analysis by placing dot marker symbols at the various co-incident datapoints, revealing visual patterns. In this way, the ability to

incorporate unique responses to a multitude of molecular cues, all within a single pair of cells, or through further multiplexation with additional cell populations becomes apparent.

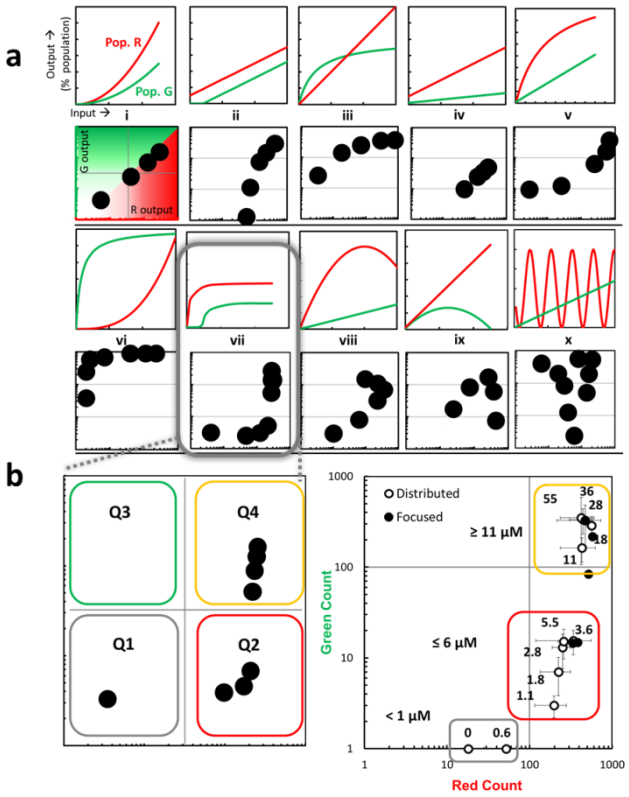


Figure 5.6 Two-color fluorescence output delineated by hypothetical population pairs. (a) Ten hypothetical response trends to molecular input for pairs of fluorescence-reporting cell populations are suggested. Included on a second plot, the red and green output are plotted against each other across the response trend (i-x). This establishes a visual field, showing the extent of bias toward each color (illustrated in example i). (b) The two-population pairing shown in trend vii defines visual output that inherently bins into three quadrants: Q1, negligible color, Q2, red bias due to majority red cell output, and Q4, combined red and green output due to equal ratios of each. (c) Fig. 5.5d and 5.6b data is plotted in the red vs. green plane, where each data point represents an input concentration of autoinducer-2 (labeled, µM) and is plotted as the output red and green cell number. The output is binned into Q1, Q2, and Q4 between the concentration thresholds defined on the plot.

Our AI-2-conveying cell network is roughly similar to example vii in Fig. 5.6a and the AI-2 response curves in Fig. 5.4 (characterized by Table S3 equations). Example vii establishes output into three basic quadrants, including Q1 (negligible color), Q2 (majority red), and Q4 (roughly equal red and green) (Fig. 5.6b). We recast the data from Fig. 5.4d and Fig. 5.5b as a

phase-plane portrait in Fig. 5.6c. This reveals the mechanisms by which the output is binned and how the originating cell response curves lead to this pattern consistently before and after magnetic refinement. In the Supplementary Information (Fig. S15), we demonstrate a parameterization of the red and green response curves that suggest the methodology is robust, that when cells are appropriately engineered, one could "tune" system characteristics to enhance or diminish a binning effect. We suggest that the utility of subcellular genetic tuning extends well beyond per-cell performance. Rather, we suggest such a strategy may be used to guide the dynamics of population architecture for actuation of by-design response patterns at a systems level.

5.4 Discussion

While cell based sensors work well in well-defined assay conditions, extension to complex environments remains a challenge. They grow, they move, they perturb their environs, they report in a time and concentration-dependent manner, small numbers of sensor cells may require signal amplification, etc. Also, increasingly, bacterial cells are engineered for user specified "executive" functions in complex environments[15,146,147]. Their performance depends on their ability to filter out extraneous noise while surveying the molecular landscape, and providing informed actuation.

By making use of a diversity of biotic and abiotic features, our multidimensional system of "responder" populations exemplifies several key metrics that promote executive performance in such environments: active molecule capture, post-capture refining of the detection output, and finally the utilization of multiple feedback thresholds[148-150]. Here, cells facilitate AI-2 recognition *autonomously* and *actively* because, as a distributed network, they reside

planktonically, chemotaxing to and continually processing signals over time. When AI-2 is detected, a processor cell's cognate machinery responds by upregulation of the native QS operon, leading to rapid signal uptake and thereby creating an active-capture signal processing mechanism. To maximize information acquisition and account for a potentially heterogeneous molecular landscape, cells serve as molecular sampling units among a distributed population, which leads to data fed back as a consensus of fluorescent "datapoints". Then, distributed data collection can be selectively reversed via the incorporated abiotic feature: magnetic nanoparticles, fastened externally on the cell through affinity-guided self-assembly. As such, responding cells obtained this extendable feature, thereby becomes sensitized to repositioning within a magnetic field.

The layered nature of the processor here, from the subcellular to multicellular scale, permits a series of selective steps: it commences with the AI-2-triggered expression cascade which releases a tight repressor, surface localization of both the fluorescent protein and SBP tag, and finally nanoparticle binding for recovery. Additionally, multiple layers of amplification result in orthogonal fluorescence feedback. The AI-2 detection event leads to whole-cell fluorescence through expression of many protein copies[85]. Then their physical collection further amplifies the signal, yielding a macroscopic composite of many individual cell units. When utilized as a network of multiple constituencies, responder cell types A and B contribute individual recognition results (off, red, or green) to a single consensus output. Finally, due to their overlapping thresholds for recognition of the same molecule, in this case, AI-2, parallel processing by A and B responders can contribute to visual interpretation of information about the molecule. Outcomes are classified into a finite number of states: here, output to no fluorescence, red, or red and green, with each addition of color as a metric of a higher interval of AI-2. In

many respects the elucidation of layered information networks as demonstrated here is analogous to computer information processing via information theory[151-153].

Here, however, interrogation of biological systems requires a reliable means for accessing molecular information- that which is communicated between biological species *and* that which can be relayed to the end user. The responder cells need not be present in high concentration, nor must they all be collected in the present format. We suggest engineered biological mechanisms are well-poised to serve at this critical interface between information acquisition and user interaction. Thus, the functional design of components for autonomous self-assembly, decision making, and networking is requisite in the field of micro- and nano- scaled machines. Our combinatorial approach allows for cells to independently assess, yet collectively report, on molecular information. Its processing is enabled through appropriate integration of synthetic biology and nanomaterials design. We suggest this approach provides a rich opportunity to direct many formats of multi-population response through genetic tuning and systems-level engineering. Further development of cellular networks and incorporation of alternate abiotic attributes can expand the depth and breadth of molecular communication for user specified actuation.

Chapter 6: Summary

6.1 Conclusions

From this dissertation work, we have extended the capabilities of biological and bioengineered entities to communicate information through the development of bio-enabled conduits of information flow. The transmitted information varied in nature from biological to non-biological: a functional metabolic product (Chapter 2), digital data in DNA (Chapter 3), molecular information converted from DNA-stored data (Chapter 4), and the conversion of molecular information to binned optical output (Chapter 5). The information was also transmitted along various trajectories: between spatially arranged addresses (Chapter 2), between a physical (genetic sequence) and digital format (readout message, Chapter 3), from a DNA template to a cell population through enzymatic transduction (Chapter 4), and finally, across length scales by amplification of a small molecule input to an aggregated cell population for provision of multi-pronged optical output (Chapter 5). Finally, in many cases, molecular information was processed through digitization: conversion to computer-compatible binary code (Chapter 3), representation as discrete signal amplitude (Chapter 4), and the transduction from a continuous-state signal to finite-state optical output (Chapter 5).

All of the information flow networks were supported by the integration of diverse biological and biocompatible tools. Additionally, a convergence of protein and genetic engineering strategies yielded a robust combination of cells, nucleic acids, and chimeric proteins that were individually tuned for precise assembly and controlled performance. Ultimately, the communication relay in each case was guided by synchrony between intracellular and

extracellular components. Furthermore, their bottom-up inclusion within a network allowed for modularity and reversibility. In this way, the systems present customizable network platforms.

Our multi-dimensional demonstrations of information flow have expanded the potential complexity that can be incorporated into bio-networks. Each chapter highlighted robust attributes that can be further merged: their spatial organization, on-site information storage, and a relay from signal-to-cell, as well as cell-to-user. Thus, the incorporation of molecular communication through interactive bio-componentry is a powerful tool for developing robust, electronics-free operational continuity. Because cells and molecules are supreme molecular recognition agents, their integration and interaction in synthetic formats is instrumental for advancing biotechnological tools toward 'smart' diagnosis, therapy, and general stasis *in vivo* and in the environment.

6.2 Future Directions

First, due to growing opportunity for the synthetic biologist to contribute to bioMEMs performance, further integration with devices and materials ought to be more widely explored. Our work in Chapter 2 demonstrates the feasibility of arranging several biological tools into an integrated system with defined processingstations. Our following work develops additional biological tools, all of which related to the field of infobiology, and all of which would prove highly applicable to a bioMEMs device format [154] for on-chip data processing of information that is biological/chemical in nature.

Expanding on the content established by this dissertation, we may recognize additional opportunities for integration of the componentry itself. For instance, there are architectural similarities in the DNA sequence modifications of Chapter 3 and Chapter 4, where GATC

quantity was modified in both cases. We envision the combination of DNA constructs for a merged infobiology application of digital DNA storage that can be directly translated to a cell-readable format. By combining GATCblocks (such as ABC or even, perhaps, $(ABC)^2$) with pGATCfree, up to 50 % of the circularized sequence would contrast its other half- one being GATC-rich (with 300 GATC sites via $(ABC)^2$) and the other GATC-free. Then through rolling circle amplification of this construct, each GATC-rich and GATC-free segment would be alternately repeated approximately four times (assuming an order of 20 kb amplified products). Resultantly, the GATC-banded DNA tract could be tested for its ability to influence the *frequency* rather than amplitude of AI-2 conversion. Based on the known processivity of EcoDam described in Chapter 4, its rate-limiting step for AI-2 synthesis (Fig. 4.3), and our data that suggests a high degree of synchronization in methylation order across substrates (Fig. 4.2), we anticipate that the programmable generation of AI-2 pulses based on GATC inscriptions is a feasible endeavor. Further, work by T. Gordonov et al (2013) suggests that electrochemical quantification of the AI-2 synthesis byproduct homocysteine may provide an appropriate strategy for real-time detection of AI-2 pulse generation [155]. AI-2 pulses, then, could constitute real-time data flow, whereby pulse size could represent coded information, analogous to Morse Code.

Secondly, as Chapter 5 concludes with an engineered bacterial consortium, whose subpopulation ratio is controlled by AI-2 level, there is a direct link between DNA-programmed AI-2 amplitude for, in turn, programmable cell response. The operations employed by each component of this system could be readily organized on-chip, as was demonstrated in Chapter 2. Further, by incorporating magnetic elements at electroaddresses, as was described by Y Li et al (2015), the magnet destination for mNP-equipped cells could be spatiotemporally defined [156]. Reversible mobilization and immobilization within a magnetic field could be a powerful added

capability, especially in guiding the controlled assembly of cell populations (or signal-selected subpopulations- shown by us using an AI-2 signal).

Such capabilities could be useful toward enhancing the biological functionality of otherwise inert systems, or in other words, rendering them "bioactive." Along these lines, we have developed bioactive entities of increasing complexity by encapsulating the necessary functional machinery. First, A Gupta et al (2012) demonstrated capsule formation using alginate and chitosan electrostatic interactions (Fig. 6.1a) [157]. By encapsulating the nanofactory of R Fernandes et al (2010)[91] within, the bioactive capsules engaged with cells to influence population phenotype through AI-2 communication (Fig. 6.1b) [157]. Next, similarly biofabricated capsules were instead localized with AI-2-producing cells to constitute a multi-cellular recruiting entity (Fig. 6.2a) to which engineered cells (trackable via red fluorescence) migrate through AI-2 chemotaxis (Fig. 6.2b) [85]. In this way, similar polysaccharide scaffolding could be used to divergent purposes- to create fixed stations by on-chip electroaddressing, to build a mobilized vessel for confinement, or for biomimicry by loading a defined biomolecular system to endow phenotypic behavior.

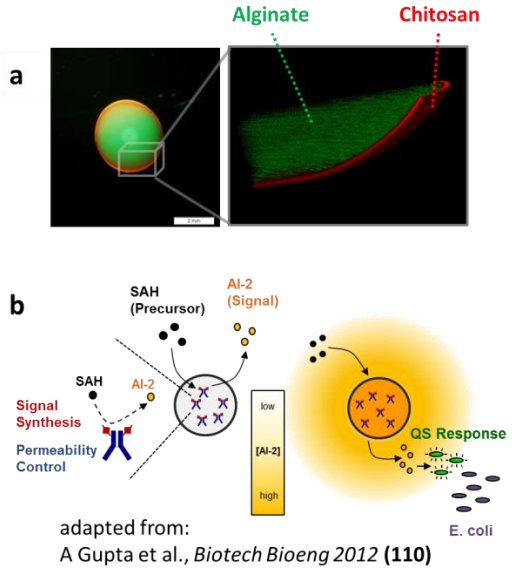


Figure 6.1 Bioactive capsules that confer response from *E. coli*. Through encapsulation of the protein machinery for AI-2 synthesis with chitosan and alginate, diffusion of the signaling molecule from the capsule influences the quorum sensing behavior of surrounding cells. Work published by A Gupta et al in *Biotechnology and Bioengineering* (2012) [157].

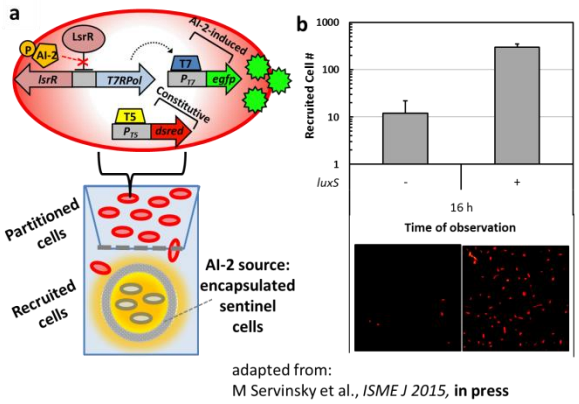


Figure 6.2 Population recruiting by an encapsulated multicellular aggregation. (a) Encapsulation of sentinel cells allows for signal-based recruiting to a unit source. The cell recruits, meanwhile, were altered for constitutive red expression, and therefore, continuous tracking, with a second signal-induced fluorescence response (green- not shown here). (b) Capsules harboring sentinel cells with the gene for signal synthesis (luxS) summon significantly more red-recruits across the partitioned fraction. Work published by M Servinsky et al in the *ISME Journal* (2015) [158].

Beyond encapsulation, other collaboratory work has resulted in genetically regulated artificial cells. L Martini et al (2011) first demonstrated inducible riboswitch activity within a lipid vesicle (Fig. 6.3b) loaded with transcription and translation machinery for ribo-regulated protein expression (Fig. 6.3c) [159]. Then, R Lentini et al (2014) used the ribo-regulated artificial cell to translate its inducer signal (not recognized by natural cells) into a cell-recognized signal (Fig. 6.3a) [106]. In this way, the artificial cell communicated with *E. coli* to elicit cellular response (Fig. 6.3d) [106]. By incorporation of the genetic equivalents of the proteins used in our AI-2 relay, an artificial cell could be DNA-programmed for AI-2-based quorum sensing communication, having defined signal amplitudes, frequencies, or even signal generation that is non-depleted. The latter could be accomplished by incorporating rolling circle amplification for continuous elongation of a GATC-rich DNA template within the lipid vesicle.

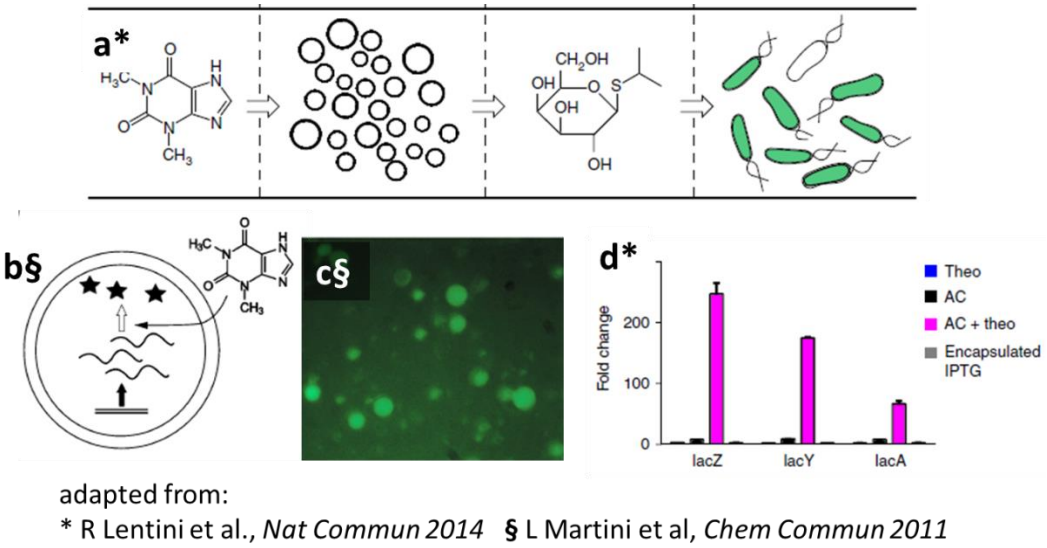


Figure 6.3 Artificial cells for translating chemical cues for interpretation by *E. coli* cells. (a) The translation of a chemical compound mediated by artificial cells (ACs) to influence *E. coli* recognition and response. (b) The AC is a lipid vesicle equipped with a genetic template and the transcription/translation machinery for inducible protein production. (c) Production of YPet fluorescence protein by artificial cells. (d) Gene upregulation in *E. coli* in response to the AC-translated signal.

In conclusion, our work has vast potential for alternative assembly schemes and further integration between components, especially considering the potential inclusion of our collaboratory work. We envision the application of intricate, interactive bio-networks, especially for the purpose of relaying DNA-stored information through small-molecule conversion, cell-programmed response, and futuristically, automated device actuation.

Appendix A: Sequence information for Chapters 3 and 4

A-1: Character-to-ASCII binary conversion

Character	ASCII (7bit)
W	1010111
E	1000101
B	1000010
L	1001100
A	1000001
B	1000010
@	1000000
U	1010101
M	1001101
D	1000100
	0000000
#	0100011
i	1101001
n	1101110
f	1100110
o	1101111
b	1100010
i	1101001
o	1101111
#	0100011
G	1000111
A	1000001
T	1010100
C	1000011
f	1100110
r	1110010
e	1100101
e	1100101
4	0110100
8	0111000
ETX	0000011

A-2: List of Sequences

EGFPtxt

CGATTAAATAAGGAGGAATAAACC**ATG**GTTAGCAAGGGCGAGGAGTTATTCACCGGCGTGGTTCCGATTT
TAGTTGAACTGGATGGCGACGTGAATGGGCATAAGTTCAGCGTGAGTGGCGAAGGCGAAGGCGACGCCAC
CTATGGCAAATTAACCTTAAAATTTATTTGTACCACGGGGAAGCTGCCCGTGCCTTGGCCGACATTAGTG
ACCACATTAACCTACGGTGTGCAGTGTTTCAGCCGTTATCCGGACCATATGAAACAGCATGATTTTTTCA
AAAGCGCGATGCCTGAGGGTTACGTGCAAGAACGTACAATTTTCTTCAAAGACGACGGTAACTACAAGAC
CCGTGCCGAGGTGAAGTTCGAAGGTGACACTTTAGTTAATCGTATTGAACTCAAGGGCATTGACTTCAAA
GAGGATGGCAATATTTTAGGTCATAAGTTAGAGTACAACTACAACAGCCATAATGTTTACATTATGGCGG
ATAAACAAAAGAATGGTATTAAAGTGAACTTTAAAATCCGCCACAACATTGAGGGCTCTGTGCAGTTAGC
GGACCATTATCAGCAAAATACACCCATTGGCGATGGTCCTGTGTTATTACCCGATAACCACTATTTAAGT
ACCCAGAGTGCGTTATCTAAGGACCCGAACGAGAAACGCGACCATATGGTTTTACTGGAGTTTGTGACCG
CCGCTGGTATTACCCTGGGCATGGATGAACTGTATAAATGATGA**TAA**AAGCTT

GATCBlock A

TTAAATAAGGAGGAATAAACCATGGACTAGTGCATGCTAGATCCTAACGTGCTGACACAATCCAGAGCAT
TTGTAAGCCGCATATTGAGTCTGGATCGATCCCTGAAGATCGGGCATGATCGTTCAACGATCAAGTCACG
ATCTGATCTTCCAGCATGATCCCCATGTTGATCCTGATCCTGATCACGTGTCTTGATCCGGCTGACCGAT
CACCTTGATCCGAAGGATCCGATCGAGCATTGTCGATCCAAGTGGACTGATCGTTCGATCCATAGATCGA
GTCACGCGATCATGCACGAAGATCTTCGATCGGCGATCTCCTTCGCCTTCTCGTCGGCTGCGGCGAGCCG
CCACACTTTATGGACGCAATAGGGGACTCTTTTGCTCACATGTGTGAGCTGATACCGCTCGCATCAGAGA
TGGCGTAGTCGCTCAGTCAAAGCGGAAGAACGGCGTATCAGCGACCGTGCTCGGTCGTTCGCAAACCGTA
TCC

GATCBlock B

GCTCGGTCGTTCGCAAACCGTATCCGTTTGAGTACCTTTGAGTGATGCCTTGATCCGGATCGATCAATGA
TCGCAGATCACTCGTCAGATCGAAGTCTGATCCCCTGATCCTCGGATCATGTACGATCAAGTCGAGATCT
AAACGATCAGCCGGGATCTTGATCACGATCGCAGATCACGCGATCATATGATCCGGTTGAAGATCGGGTG
CTGATCGGATCGAACGAGATCATGTTTCGATCCTGATCAAAGGTCAGCTCACTCAAAGGCACCTGATGCG
TGGATACGTCTAACTCTCCATGCCACACTTTAGAACCAGTACCGTTGACCCTGCAGTCCGCGAGAAGTAC
AAACCAGCACTCGCCAGATGCCATTCGTCTTTTCCGGTCCTTGGATACGCCGTAGTCTCGTCTAACATAC
GCCACACTTTCGTAGTCCGAGTGCCACATCTGCTCAGGTACGATACCCACGACATCCTGAGCAGGGACCT
ACTACCAC

GATCBlock C

CATCCTGAGCAGGGACCTACTACCACGCGTAGGTCTTTCTTGTATCCGGCTCTTCCGCTCTCGCTCACTA
GAAGGCGATACCGCCTTTGAGGAGCTACTAAGCCTTGGAGTGACTATTGCGTCTCTGATCGATCCCTGAA
GATCGGGCATGATCGTTCAACGATCAAGTCACGATCTGATCTTCCAGCATGATCCCCATGTTGATCCTGA
TCCTGATCACGTGTCTTGATCCGGCTGACCGATCACCTTGATCCGAAGGATCCGATCGAGCATTGTCGAT
CCAAGTGGACTGATCGTTCGATCCATAGATCGAGTCACGCGATCATGCACGAAGATCTTCGATCGGCGAT
CCTTTGAGTGATGCGCTCGGTCGTTCGCAAAAGGTCAGCTCACTCAAAGGCACCTGATGCGTGCGTTTCC
TATGAGTACGACCGTATTGAGTTGTTCCTCATCGCGAGATCGCATGCTCTAGAAAGCTTGGCTGTTTTGG
CGGATGAG

A-3: List of Primers

Oligo name	Sequence (5'—3')*
Txt-NcoI-F	ATTAAAAATTAAAGAGGTATATATTAATGTATCGATTAAATAAGGAGGAATAAACCATGG
Txt-HindIII-R	CTTCTCTCATCCGCCAAAACAGCCAAGCTT
ABlock-F	TTAAATAAGGAGGAATAAACCATGGACTAGTGCATGCTAG
ABlock-R	GGA TAC GGT TTG CGA ACG ACC GAG CAC G
BBlock-F	GCTCGGTCGTTCGCAAACCGTATCCGTTTG
BBlock-R	GTG GTA GTA GGT CCC TGC TCA GGA TGT CG
CBlock-F	CAT CCT GAG CAG GGA CCT ACT ACC ACG
CBlock-R	CTC ATC CGC CAA AAC AGC CAA GCT TTC TAG AG
ArevBlock-F	TAT TGA GTT GTT CCT CAT CGC GAC GGT CGCTGA TAC GCC G
ArevBlock-R	GAC CTA CTA CCA CCT AAC GTG CTG ACA CAATCC AG
BrevBlock-F	AGC ACGTTA GGT GGT AGT AGGTCC CTG C
BrevBlock-R	TTG TTC CTC AGCTCG GTC GTT CGC AAA C
CrevBlock-F	ACG ACC GAG CTG AGG AAC AACTCA ATA CG
CrevBlock-R	TTC TAG AGC ATG CGATCT CGC GAG CGT AGGTCT TTC TTG TAT C
ABlock-HindIII-R	CTC ATC CGC CAA AAC AGC CAA GCT TGG ATA CGG TTT GCG AAC GAC CGA GC
BBlock-HindIII-R	CTC ATC CGC CAA AAC AGC CAA GCT TGT GGT AGT AGG TCC CTG CTC AGG ATG
CBlock-HindIIIext-R	CTC ATC CGC CAA AAC AGC CAA GCT TTC TAG AGC ATG CGA TCT CGC GA
Circ-F	AAGCTTGGCTGTTTTGGCGGATGAGTTAAATAAGGAGGAATAAACCATGG

Appendix B: Supplementary information for Chapter 5

This appendix was primarily reproduced with permission from a paper recently accepted for publication: JL Terrell, HC Wu, CY Tsao, NB Barber, MD Servinsky, GF Payne, and WE Bentley, "Nano-guided cell networks as conveyors of molecular communication*," Nat Commun*, **2015,** *just accepted*.

1. Plasmids and strains.

Table S1. List of strains and plasmids.

Strain/plasmid	Relevant genotype and property	Source or reference
Strains		
E. coli		
W3110	K12 strain, wild type, λ-, F-, IN(*rrnD-rrnE*)1, *rph-1*s	Genetic Stock Center Yale University, New Haven, CT
MDAI2	W3110 *luxS*::Tcr W3110-derived *luxS* mutant strain	[160]
CT104	W3110 *luxS*::Tcr, lsrFG$^-$, W3110-derived *luxS, lsrFG* mutant strain	[16]
BL21 (DE3)	B strain, F– *ompT hsdSB*(rB- mB-) *gal dcm rne131*λ(DE3)	Invitrogen
V. harveyi		
BB170	BB120 *luxN*::Tn5 (sensor 1- ,sensor 2+), Km^r	[161]
L. innocua		[145]
Plasmids		
pCT6	pFZY1 derivative, containing *lsrR* and *lsrR* promoter region fused with *T7RPol*, Ap^r	[139]
pFZY1	*galK'-lacZYA* transcriptional fusion vector, Ap^r	[162,163]
pET200/	Cloning vector, containing *T7* promoter, Km^r	Invitrogen
pAIDA-I	pACYC184 derivative, containing AIDAc translocation gene, Cm^r	[164]
pSBP-Venus	pET200 derivative, containing *sbp-venus*, AIDAc translocation gene, Km^r	This study
pSBP-mCherry	pET200 derivative, containing *sbp-mcherry*, AIDAc translocation gene, Km^r	This study
pSBP-eGFP	pET200 derivative, containing *sbp-egfp*, AIDAc translocation gene, Km^r	This study

2. pSBP plasmid construct methods and maps.

Table S2. List of primers.

Oligo name	Sequence (5'—3')*
SBP-mCherry-F	CGGGCGAACTGGAACAGCTGCGCGCGCGCCTGGAACATCATCCGCAGGGCCAGCGCGAACCGggaattaac ATGGTGAGCAAGGGCGAGG
mCherry-XbaI-R	gcgcgcgctctagaCTTGTACAGCTCGTCCATGCCGCC
SalI-SBP-F	gtgtgtgtgtcgacATGGATGAAAAAACCACCGGCTGGCGCGGCGGCCATGTGGTGGAAGGCCTGG CGGGCGAACTGGAACAGCTGCGCG
NdeI-AIDAc- F	cccctctagaaataattttgtttaactttaagaaggagatata catATGAATAAGGCCTACAGTATCATTTGGAGCCACTCCAG
AIDAc-SacI-R	gttagcagccgatcgttgagctcgcccttTCAGAAGCTGTATTTTATCCCCAGTGCTCCGG
SBP-XhoI-Venus-F	GCGAACTGGAACAGCTGCGCGCGCGCCTGGAACATCATCCGCAGGGCCAGCGCGAACCG ctcgagATGAGCAAAGGCGAAGAACTGTTCACGGG
Venus-KpnI-R	cgcgcgcgggtaccTTTATACAGTTCATCCATACCATGGCGTAATGC
Venus-KpnI-Linker-F	GGCCGGCATTACGCATGGTATGGATGAACTGTATAAAggtacc GTGAATAACAATGGAAGCATTGTCATTAATAACAGCATTATAAACGGG
Linker-NotI-R	gcgcgcgcgcggccgcACTTTCTTTTGTAGGATTAAGAATGATATTTTTGTTATTAAC
XhoI-EGFP-F	ctctctctctcgagATGGTGAGCAAGGGCGAGGAGC
EGFP-KpnI-R	agagagagggtaccCTTGTACAGCTCGTCCATGCCGAGAGTGATC

*Capitalized letters indicate complementarity to a gene. Blocked letters indicate a restriction endonuclease site. Underlines indicate complementarity to another primer.

Vector Construction. A series of cloning steps were performed to generate three plasmids that utilize the AIDAc membrane transporter for the surface display of SBP-tagged fluorescent proteins: pSBP-mCherry, pSBP-Venus, and pSBP-eGFP. First, one-pot nested PCR was performed using *mcherry* as a template and 0.2 μM SBP-mCherry-F and 0.5 μM mCherry-XbaI-R as oligomers. After 5 cycles, a minimal volume of 100 μM SalI-SBP-F was added to achieve 0.5 μM in the working PCR reaction to further extend the 5' end for another 30 cycles. The vector pAIDA-I and the PCR-generated sequence were double-digested with SalI and XbaI in order to insert *sbp-mcherry* as the passenger in the AIDAc construct. Next, NdeI-AIDAc-F and AIDAc-SacI-R were used as primers to PCR the gene fusion of *signal peptide-sbp-mcherry-AIDAc* from start to stop codon. The destination vector, pET200, and the PCR product were each digested with NdeI and SacI to insert the gene fusion after a T7 promoter. The resulting plasmid was named pSBP-mCherry.

Additional measures were taken to allow for more facile cloning of a variety of fluorescence genes. The one-pot nested PCR strategy was used to amplify *sbp-venus* using SBP-XhoI-Venus-

F, SalI-SBP-F, and Venus-KpnI-R as primers, introducing new restriction sites directly up and downstream of *venus*. A linker upstream of *AIDAc* was also PCRed using primers Venus-KpnI-Linker-F and Linker-NotI-R, which extended the sequence to include an overlap region with *venus*. Next, the two PCR products, *sbp-venus* and the linker, were annealed together in a PCR reaction to link the two sequences together and amplify their combination; no primers were necessary since the sequences shared complementarity at the 3' end of *venus*. The PCR product was digested along with pSBP-mCherry using SalI and NotI to replace *mcherry* upstream of *AIDAc*. The completed construct was named pSBP-Venus.

Finally, XhoI-eGFP-F and eGFP-KpnI-R were used as primers to amplify *egfp*. Digestion of both *egfp* and pSBP-Venus with XhoI and KpnI were followed by routine insertion to swap *venus* for *egfp*. The resulting plasmid was named pSBP-eGFP. Constructs for each engineered plasmid are shown in Fig. S1.

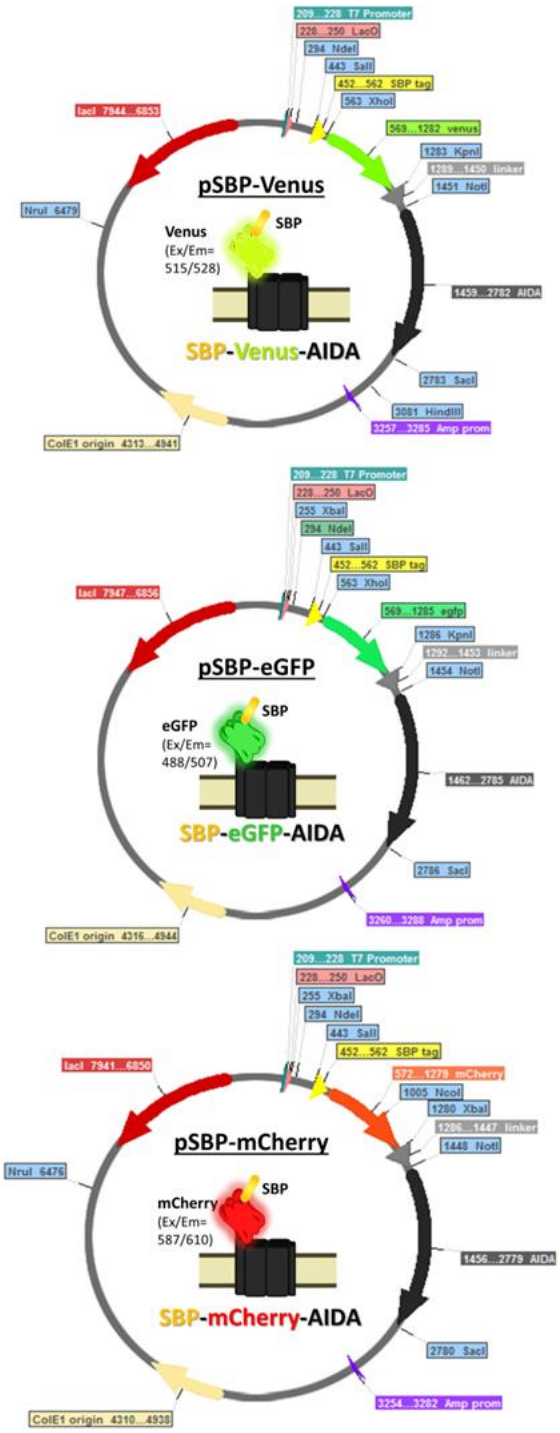


Figure S1 AIDAc-tethered surface displays fused with fluorescent proteins and a streptavidin binding peptide (SBP) tag. Representations are depicted above of the outer membrane containing the AIDAc transmembrane domain with either Venus-SBP, eGFP-SBP, or mCherry-SBP expressed on the outer surface. Below are the genetic constructs of each surface display under T7 regulation.

3. Template gene sequences inserted on pET200 vectors.

Signal peptide

ATGATAAGGCCTACAGTATCATTTGGAGCCACTCCAGACAGGCCTGGATTGTGGCCTCAGAGTTAGCCAGAGGACAT
GGTTTTGTCCTTGCAAAAAATACACTGCTGGTATTGGCGGTTGTTTCCACAATCGGAAATGCATTTGCAG

Streptavidin binding peptide

ATGGATGAAAAAACCACCGGCTGGCGCGGCGGCCATGTGGTGGAAGGCCTGG**CGGGCGAACTGGAACAGCTGCGCG**C
GCGCCTGGAACATCATCCGCAGGGCCAGCGCGAACCG

Venus

ATGAGCAAAGGCGAAGAACTGTTCACGGGTGTGGTTCCGATCCTGGTTGAACTGGATGGCGATGTGAACGGTCATAA
ATTTAGCGTGTCTGGTGAAGGCGAAGGTGATGCGACCTACGGCAAACTGACGCTGAAACTGATTTGCACCACGGGTA
AACTGCCGGTTCCGTGGCCGACCCTGGTGACCACGCTGGGTTATGGTCTGATGTGTTTCGCACGTTACCCGGATCAC
ATGAAACGCCATGATTTCTTTAAATCTGCGATGCCGGAAGGCTATGTGCAGGAACGTACCATCTTTTTCAAAGATGA
TGGTAACTACAAAACCCGCGCGGAAGTTAAATTTGAAGGCGATACGCTGGTGAACCGTATTGAACTGAAAGGTATCG
ATTTCAAAGAAGATGGCAATATTCTGGGTCACAAACTGGAATACAACTACAACAGTCATAACGTGTACATTACCGCC
GATAAACAGAAAAACGGTATCAAAGCAAACTTCAAAATCCGTCACAACATCGAAGATGGCGGTGTTCAGCTGGCCGA
TCATTACCAGCAGAACACCCCGATTGGCGATGGTCCGGTGCTGCTGCCGGATAATCATTATCTGAGTTACCAGAGCA
AACTGTCTAAAGATCCGAATGAAAAACGCGATCACATGGTTCTGCTGGAATTTGTGACCGCGGCCGGCATTACGCAT
GGTATGGATGAACTGTATAAA

mCherry

ATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTC
CGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGA
AGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTAC
GTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAA
CTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGC
TGCGCGGCACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGG
ATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGC
TGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACA
TCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGCATG
GACGAGCTGTACAAG

eGFP

ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCA
CAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCG
GCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGAC
CACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGA
CGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCA
TCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATG
GCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGC
CGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGT
CCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACT
CTCGGCATGGACGAGCTGTACAAG

Linker

GTGAATAACAATGGAAGCATTGTCATTAATAACAGCATTATAAACGGGAATATTACGAATGATGCTGACTTAAGTTT
TGGTACAGCAAAGCTGCTCTCTGCTACAGTGAATGGTAGTCTTGTTAATAACAAAAATATCATTCTTAATCCTACAA
AAGAAAGT

AIDAc

ATGTATAGGTAATACTCTTACCGTGTCAAATTATACTGGGACACCGGGAAGTGTTATTTCTCTTGGTGGTGTGCTTG
AAGGAGATAATTCACTTACGGACCGTCTGGTGGTGAAAGGTAATACCTCTGGTCAAAGTGACATCGTTTATGTCAAT
GAAGATGGCAGTGGTGGTCAGACGAGAGATGGTATTAATATTATTTCTGTAGAGGGAAATTCTGATGCAGAATTCTC
TCTGAAGAACCGCGTAGTTGCCGGAGCTTATGATTACACACTGCAGAAAGGAAACGAGAGTGGGACAGATAATAAGG
GATGGTATTTAACCAGTCATCTTCCCACATCTGATACCCGGCAATACAGACCGGAGAACGGAAGTTATGCTACCAAT
ATGGCACTGGCTAACTCACTGTTCCTCATGGATTTGAATGAGCGTAAGCAATTCAGGGCCATGAGTGATAATACACA
GCCTGAGTCTGCATCCGTGTGGATGAAGATCACTGGAGGAATAAGCTCTGGTAAGCTGAATGACGGGCAAAATAAAA
CAACAACCAATCAGTTTATCAATCAGCTCGGGGGGGATATTTATAAATTCCATGCTGAACAACTGGGTGATTTTACC
TTAGGGATTATGGGAGGATACGCGAATGCAAAAGGTAAAACGATAAATTACACGAGCAACAAAGCTGCCAGAAACAC
ACTGGATGGTTATTCTGTCGGGGTATACGGTACGTGGTATCAGAATGGGGAAAATGCAACAGGGCTCTTTGCTGAAA
CTTGGATGCAATATAACTGGTTTAATGCATCAGTGAAAGGTGACGGACTGGAAGAAGAAAAATATAATCTGAATGGT
TTAACCGCTTCTGCAGGTGGGGGATATAACCTGAATGTGCACACATGGACATCACCTGAAGGAATAACAGGTGAATT
CTGGTTACAGCCTCATTTGCAGGCTGTCTGGATGGGGGTTACACCGGATACACATCAGGAGGATAACGGAACGGTGG
TGCAGGGAGCAGGGAAAAATAATATTCAGACAAAAGCAGGTATTCGTGCATCCTGGAAGGTGAAAAGCACCCTGGAT
AAGGATACCGGGCGGAGGTTCCGTCCGTATATAGAGGCAAACTGGATCCATAACACTCATGAATTTGGTGTTAAAAT
GAGTGATGACAGCCAGTTGTTGTCAGGTAGCCGAAATCAGGGAGAGATAAAGACAGGTATTGAAGGGGTGATTACTC
AAAACTTGTCAGTGAATGGCGGAGTCGCATATCAGGCAGGAGGTCACGGGAGCAATGCCATCTCCGGAGCACTGGGG
ATAAAATACAGCTTC

4. Microscopy of surface-expressed SBP-fluorophore fusions with streptavidin-labeling.

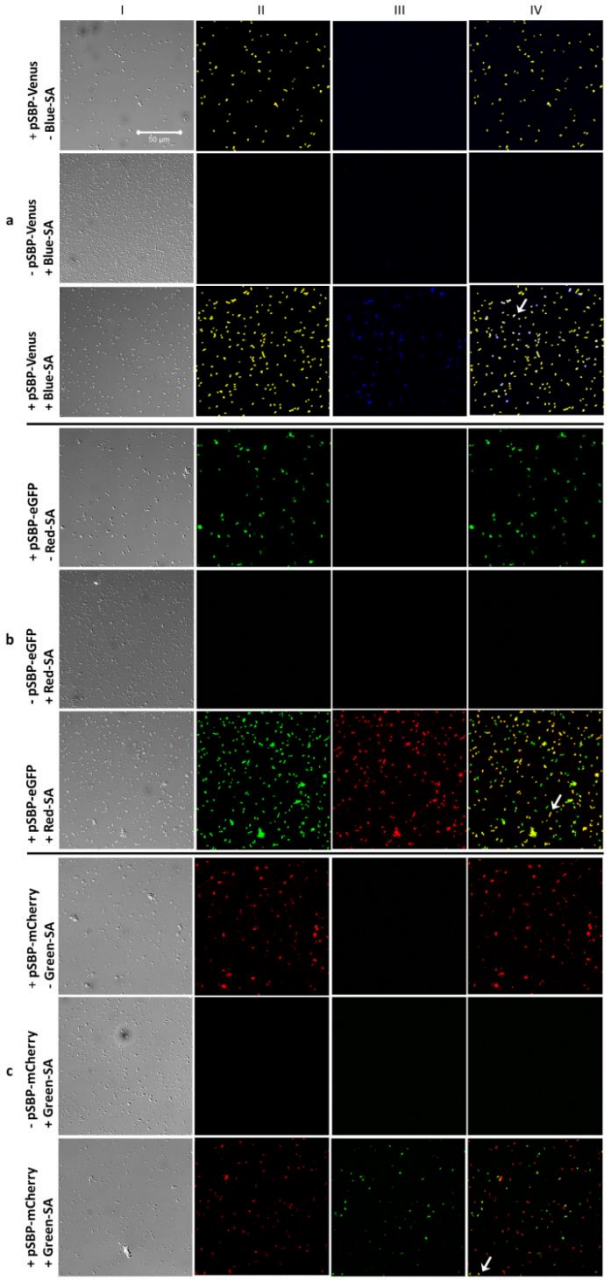


Figure S2 Streptavidin-labeling of SBP to assess its surface accessibility. Fluorophore-tagged streptavidin (SA) was used to probe for SBP expressed on the outer membrane of BL21(DE3). Each of three composites represent a set of controls, BL21(DE3) expressing the SBP-fluorescent protein fusion but unlabeled with streptavidin (row 1) and labeled with streptavidin but without the fusion gene (row 2); additionally, both the conditions of fusion gene expression and streptavidin labeling are positive in row 3. Columns I-IV represent each filter applied to the image: I, brightfield; II, filter for the fluorescent protein; III, filter for the fluorescent streptavidin; and IV, an overlay of II and III. Fluorophores with non-overlapping spectra were paired. **(a)** Venus expression (yellow-green) was paired with Dylight405-labeled SA (blue). **(b)** EGFP expression (green) was paired with Alexafluor594-labeled SA (red). **(c)** mCherry expression (red) was paired with Alexafluor488-labeled SA (green).

5. Superficial immunoprobing of AIDAc-linked green fluorescent protein (eGFP).

Having probed the engineered cell surfaces for the availability of SBP, we also investigated whether the fluorescent protein segment of the fusion protein also translocated across the outer membrane. Specifically, we probed cell surfaces for eGFP using an anti-eGFP antibody. Immunoprobing was performed on cells that had been treated to express SBP-eGFP-AIDAc and on a negative control that did not express the fusion protein. First, cells were washed in 10 mM PBS and incubated with anti-GFP monoclonal mouse IgG (Rockland Immunochemicals), diluted 100-fold in PBS, for 1 h at room temperature. Samples were washed twice in PBS, then incubated with Alexafluor594 (red) anti-mouse polyclonal goat IgG (Molecular Probes) as a secondary antibody, diluted 200-fold in PBS, again for 1 h at room temperature. A sample of cells expressing eGFP was not treated with the secondary antibody to act as a negative control in calibrating the fluorescence signal contributed by the red probe. Fig. S3 shows (i) a scheme of the conditions tested and results from (ii) flow cytometry and (iii-viii) microscopy.

FACS was used to analyze cells for green (eGFP) and red (immunostaining) fluorescence per cell and the intensities of each fluorophore were plotted as coordinates (ii), with the fluorescence map divided into quadrants: Q1, significant red fluorescence, Q2, cells exhibiting both red and green, Q3, non-fluorescing cells, and Q4, green only. In Fig. S3a, immunostained cells expressing eGFP showed strong green fluorescence, mapping in Q4, with a distinct subset mapping to Q2 for cells colocalized with red and green. By contrast, non-immunostained eGFP cells are negligible in Q2, appearing, instead in Q4 (Fig. S3b), and cells that do not express eGFP appear mostly in Q3 due to the absence of both fluorophores (Fig. S3c).

This data was visually corroborated by microscopy. Low magnification shows a large sample size of fluorescent cells (iii), with the boxed areas split into green (iv) and red (v) channels. Additionally, higher magnification was used to visualize the cell surface localization of the fluorescent components (vi-viii). Interestingly, fluorescence appears in clusters, around the perimeter of the cell cross-section. Only cells in Fig. S3a (iii-viii) show both green and red fluorescence. The positions on the cell surface where eGFP is expressed (vii) are identically colocalized (vi) with red (viii); this evidence of positive cell-surface

immunoprobing suggests that when SBP-eGFP-AIDAc is expressed, the entire SBP-eGFP fusion translocates across the outer membrane for many cells, shown by Fig. S3a(vi). While not explicitly tested, we hypothesize that the other tagged fluorescent proteins may translocate as well, especially considering that the SBP portion was at least confirmed to be surface-accessible in all variants (Fig. S2).

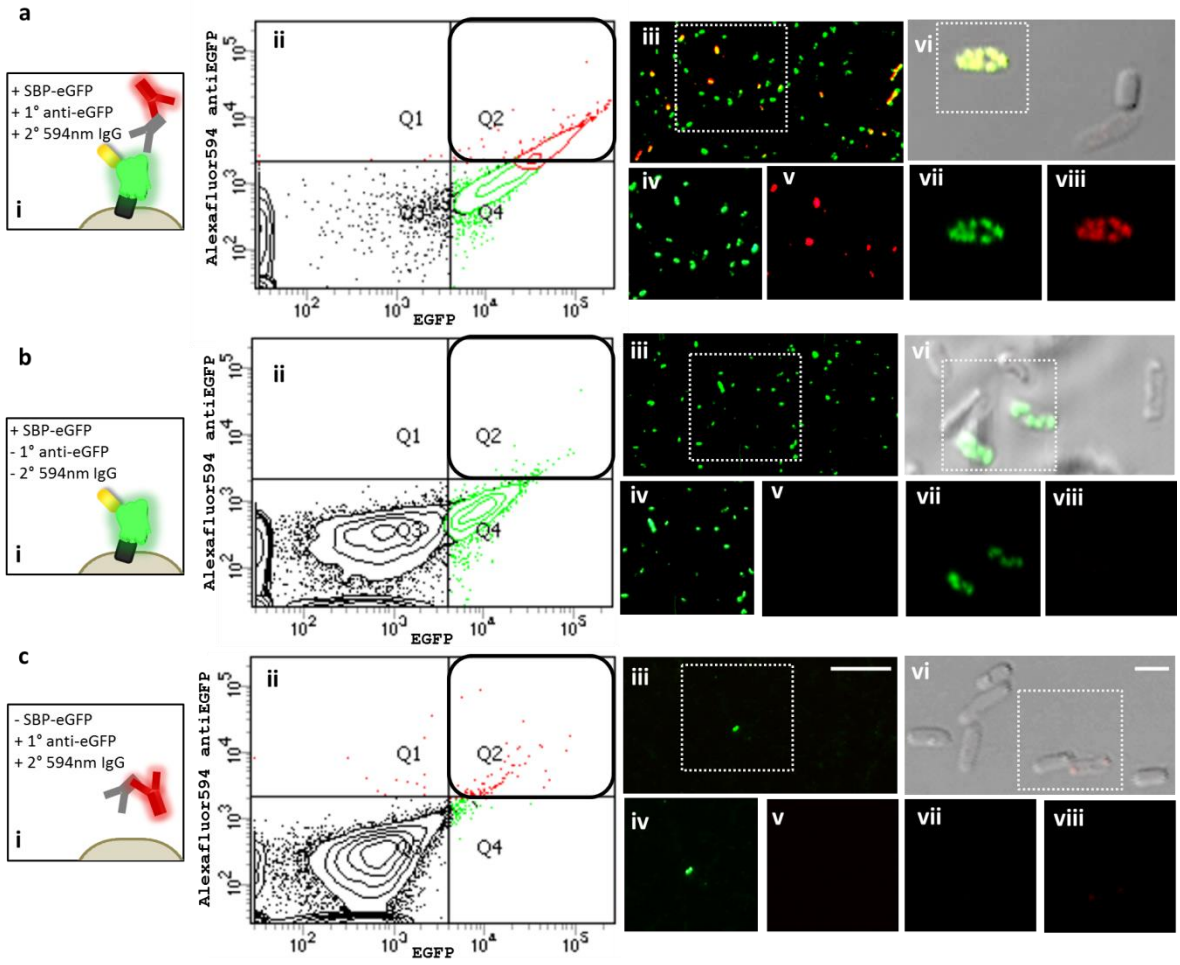


Figure S3 Immunolabeling the fluorescent protein to assess surface accessibility. The external surfaces of cells expressing AIDAc-eGFP-SBP were probed with an anti-eGFP and Alexafluor594-labeled antibody pair. Colocalization of green (eGFP) and red (surface immunostaining) was assessed by flow cytometry and microscopy. Controls included (a) cells expressing the eGFP fusion and antibody-treated, (b) eGFP-expressing cells without antibody treatment, and (c) non-expressing cells treated with the antibody. Data includes: (i) a scheme of the experimental set-up. (ii) FACS distribution of fluorescence intensities, approximated into quadrants – Q1 red cells, Q2, cells exhibiting red and green, Q3, non-fluorescent cells, and Q4 green cells. (iii) 200X magnified fluorescence image showing overlaid red and green filters, with the boxed area split into the (iv) green and (v) red filter. (vi) 1000X magnified transmitted light image with overlaid red and green fluorescence, with boxed area split into the (vii) green and (viii) red filters. Scale bars indicate 25 μm (iii) and 2 μm (vi).

6. Fluorescence and scanning electron microscopy of interactions between *E. coli* and magnetic nanoparticles.

Interactions between cells expressing SBP and streptavidin-conjugated magnetic nanoparticles (mNP, 100 nm dia., Chemicell) were investigated. In Fig. S4a-b, scanning electron micrograph images of each species ((a) cell and (b) mNPs) were used as controls for interpretation of subsequent SEM images showing combinations of the two species. In Fig. S4b, it is observed that the mNPs frequently exist in clusters of various sizes, some of which have dimensions smaller than a cell (~ 1 x 2.5 μm, Fig. S4a) and some of which are greater. It was not clear whether SEM sample preparation contributed to the clustering effect.

To investigate further, fluorescence labeling was applied for FACS and fluorescence microscopy. Fig. S4c shows a scheme of the fluorescence designation and assembly of components. Cells expressing SBP-mCherry (or uninduced for expression as a negative control) were fixed with 2 % paraformaldehyde, washed in 10 mM PBS, and stained with 13 μM propidium iodide (PI, Life Technologies); thus red cellular fluorescence was due to both mCherry and PI. Next, mNPs were diluted 50-fold from stock into the cell samples. Subsequently, FITC-labeled polyclonal anti-streptavidin IgG (Abcam), diluted 200-fold in PBS, was incubated with the cell-mNP hybrids for 30 min on ice. Finally, the samples were either imaged directly (Fig. S4e) or treated with a magnet on ice; the supernatant was separated from the magnetically-recovered pellet, which was resuspended in an equal volume of PBS and remagnetized three times to filter out cells not bound to mNPs.

FACS was used to look at the distribution of red and green fluorescence in the cell population, but to also count the magnetically recovered fractions of controls (+/- SBP-mCherry expression) in fixed volumes. Fig. S4d shows the counted density of cells, collected in triplicate from controls of equal volume. Cells were found in the negative control sample, indicating some extent of non-specific interactions with mNPs. However, an approximately three-fold greater cell density was recovered in the positive control. T-test analysis indicated that the results are statistically significant with $p < 0.05$. Thus, mNPs perform

favorably for selecting SBP-expressing cells compared to non-expressing cells. Furthermore, as shown by the inset FACS plot, a large portion of the magnetically recovered cells exhibited strong fluorescence at both red and green wavelengths, verifying that an abundance of the recovered cells were bound to FITC-labeled mNPs.

Lastly, confocal microscopy was used to visualize cell-mNP complexes. In Fig. S4, samples were compared before (e) and after (f) refining mNP-bound cells within a magnetic field. In this way, the samples were examined for an effect of clustering as a consequence of exposure to a magnetic field. Fig. S4e shows a collection of cells, of which some are unbound and some are bound to green mNPs; of the bound cells, two appear to interact closely with a large mNP cluster (~ 2 μm in diameter) and one cell appears to have at least three single particles bound to the cell surface (highlighted by an arrow). The particle diameters appear to be greater than 100 nm, which could either be due to a larger-than-average size or due to a fluorescence halo-effect.

By contrast, in Fig. S4f, a large multicellular and multiparticular conglomerate is observed. Many cells appear to be irreversibly embedded within the mass and many particle clusters appear to multivalently bind several cells. This massive clustering (characteristically larger than the size of a single cell) was only observed post-magnetic focusing, and is likely due to the strong magnetic force pulling all species to the same destination (magnet's edge) and their consequential confinement there before resuspension. It is reasonable to expect mNPs to bind with multivalency since streptavidin is multivalent (tetrameric with four binding sites) and a single mNP should be conjugated with several streptavidin molecules. Thus, the amalgamation observed in Fig. S4f is likely to occur at a magnet's surface upon redistributing species (from dispersed to focused) within the magnetic field.

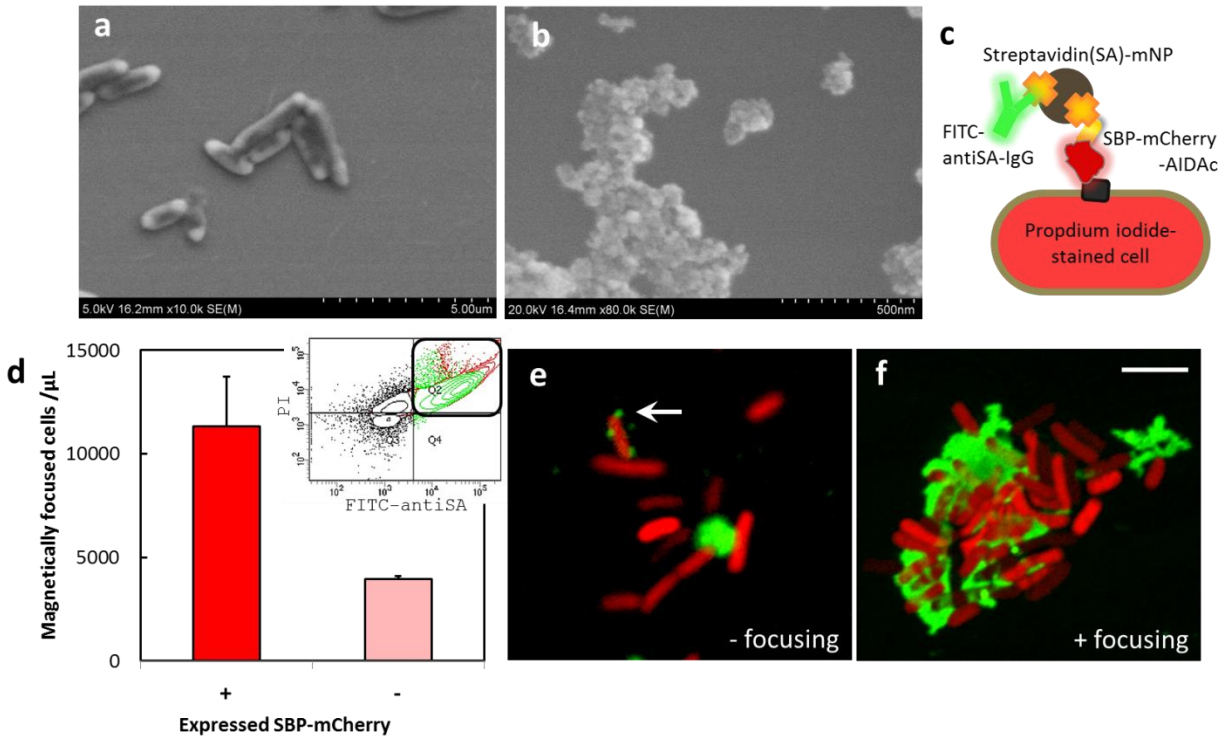


Figure S4 Analysis of cell interaction with streptavidin-coated magnetic nanoparticles. Electron micrograph of (a) *E. coli* K12 cells and (b) streptavidin coated magnetic nanoparticles (mNPs). (c) scheme of immunoprobed cell-mNP complexes. (d) Comparison between recovery via mNP-coupling and magnetic collection for cells expressing surface-displayed SBP-mCherry (+) or non-expressing cells (-). Data is in triplicate and statistically significant across controls with $p < 0.05$. Inset FACS distribution of fluorescence intensities maps the red (propidium iodide-stained) cells that are bound with FITC-anti-streptavidin (SA) labeled mNPs (circled in Q2). (e) X magnification of red cells, some of which have assembled with mNPs (example indicated by arrow). (f) Composition and organization of cell-mNP complexes after the application of an external magnet used to focus the hybrid assemblies. Scale bar indicates 5 µm.

7. Fluorescence imaging and FACS gating for analysis of fluorescence distribution of Venus-expressers.

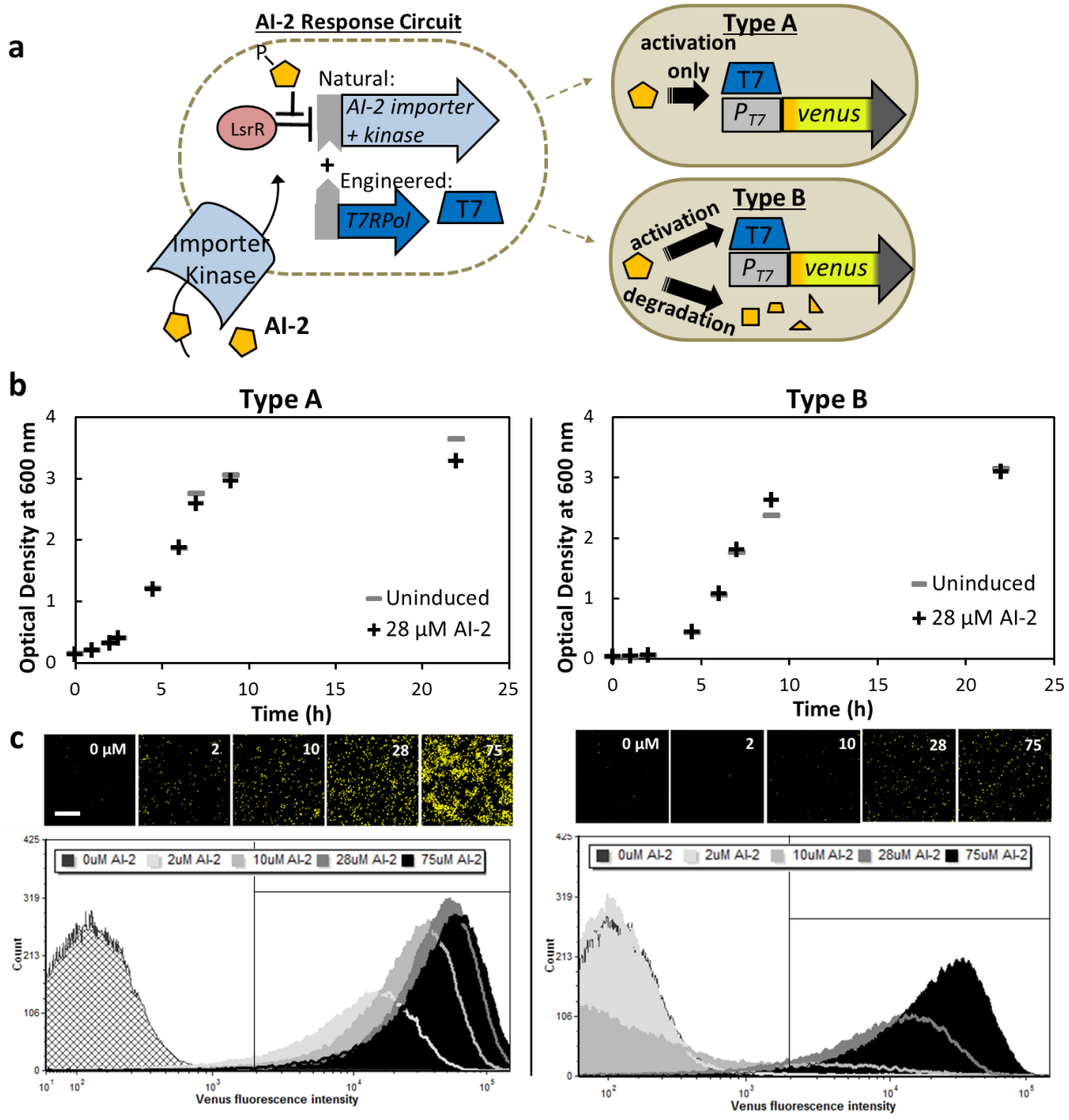


Figure S5 Engineered "Responder" cells for autonomous response to the small molecule autoinducer-2. (a) Two host strains CT104 (Δ*luxS*, Δ*lsrFG*) and MDAI2 (Δ*luxS*) were engineered for Type A and B, respectively, by linking marker chimeric protein expression to the natural response pathway for AI-2. (b) Growth curves of responders Type A and B (+ pSBP-Venus) upon addition of 28 µM AI-2 (at induction point represented by dotted line) compared against an undinduced culture. Optical density at timepoints was measured at 600 nm. (c) Fluorescence images of Responder Type A and B culture samples incubated with a range of *in vitro*-synthesized autoinducer-2 (AI-2) concentrations and corresponding FACS gating for population analysis of SBP-Venus expression. Scale bar indicates 50 µm.

8. Characterization of SBP-fluorophore-AIDAc expression on membranes of AI-2 responders.

Sample preparation. Samples were prepared by incubating cell cultures at 30°C, 250 rpm shaking in LB broth with 100 μg/mL ampicillin and 50 μg/mL kanamycin. Samples were induced with either 28 μM AI-2 and incubated overnight (for comparison across fluorescence protein variants) or 20 μM AI-2 for timecourse analysis. At the designated timepoints, samples were pelleted and washed twice in cold PBS. For cell lysis, the pellets were resuspended in Bugbuster (EMD Millipore) at 5 % of the original volume (per 3 x 10^9 cells/mL) at room temperature for 1 h. The lysis mixture was then centrifuged to separate the cell membrane fraction (in pellet) and cytosolic fraction (supernatant). The pellet was washed twice in PBS, then resuspended in 10 % SDS at an equal volume to the supernatant fraction.

SDS-PAGE. Samples corresponding to 6 x 10^8 cells (either the cytosolic or membrane fraction) were loaded into 10 % acrylamide/bis-acrylamide (37.5:1, BioRad) gels with a protein ladder (Benchmark, Life Technologies) and run at 200 V for 45 min.

Western blot. The gels were transferred to nitrocellulose membranes (BioRad) at 10 V followed by 20 V for 20 min each. The gel was then stained with Coomassie blue and destained to visualize retained protein and the nitrocellulose was blocked overnight at 4°C in 5 % milk-TBS. The membrane was probed for SBP-tagged proteins using alkaline phosphatse-conjugated streptavidin (Life Technologies) at a 1:1500 dilution in 1 % milk-TTBS. Finally, protein bands were determined using NBT/BCIP (Roche) as a colorimetric substrate for alkaline phosphatase.

Analysis of western blots.

Fig. S6 shows western blot images for analysis of the size, location (cytosol or membrane), and relative expression level of SBP-fluorophore fusions. All combinations of the SBP-tagged fusion protein, containing SBP, either Venus, mCherry, or eGFP, a linker, and AIDAc were predicted to be 88 kDa in size. In the literature, AIDAc has been reported at 51 kDa in size; when fused with a passenger, the resolved size of the fusion is often smaller than its prediction[164]. In Fig. S6a, Type A and B cells with

pCT6 and either pSBP-Venus, pSBP-mCherry, or pSBP-eGFP were compared for expression. On the western blot, no SBP-tagged protein was detected in rows to the left of the ladder, which were loaded with the cytosolic protein fractions. Heavy bands are observed to the right of the ladder, containing the membrane fractions. The stark difference in SBP-tagged protein composition between lysate fractions indicates that nearly all of the AIDAc-fused passenger is located on/in membrane rather than cytoplasm. Secondly, in the lanes loaded from the lysate pellet, protein from Type A cells, regardless of the fluorophore variant, show a stronger signal than Type B. This indicates that protein the expression level induced by 28μM AI-2 is higher in Type A cells. Finally, the blots show evidence that the SBP fusion is expressed with different efficiency depending on the fluorescent protein inserted. The band corresponding to each SBP-fluorophore-AIDAc fusion appears near 82 kDa. At this band, fusion proteins containing Venus and eGFP are expressed at higher levels than mCherry, in both A and B cell types. The corresponding SDS-PAGE gel shows nearly the same amount of protein per well, confirming that differences in band intensity on the blot are due to expression rather than a discrepancy in protein loading.

Type A and B cells were compared for SBP-Venus-AIDAc expression over time after induction with 20 μM AI-2. The western blots in Fig. S6b-c resolve the (b) membrane and (c) cytosolic fractions for samples isolated at time intervals after induction (time 0). In Fig. S6b, a band cluster develops over time at 82 kDa for both Type A and B cells, with a stronger signal from A expression. For both Type A and B cells, 3 h is required before the band intensity reaches above-background levels, from which can be inferred as the approximate time requirement for protein targeting to the membrane, maturation, and translocation. Strong expression is maintained from 3 – 7 h, decreasing slightly by 8 h and further diminished after overnight culturing, likely due to slight degradation. The 82 kDa signal is observed as a tight cluster of bands, with the largest two appearing as a pair of distinct bands, and a smear below. While it is not clear whether all constituents represent the full sequence of SBP-Venus-AIDAc, we hypothesize that the paired bands could be identified as the full fusion before (upper) and after (lower) cleavage of the signal peptide, reducing the size by approximately 5 kDa. At low expression levels,

cleavage could occur in all expressed protein, resulting in only the lower band and slight emergence of the upper band in some cases (Type A: 0, 2 h; Type B: all timepoints). Once overexpressed, a percentage of the fusion protein remains uncleaved (other timepoints). The lower smear could be explained by growing protein-in-translation, which is reduced over time (by 24 h) and at lower expression levels (Type B cells).

We note that in Fig. S6c (the cytosolic fractions), no bands at 82 kDA are observed at any timepoint and in either cell type. This absence indicates that at the onset of expression, the protein is immediately targeted to the membrane- evidence of highly efficient surface display.

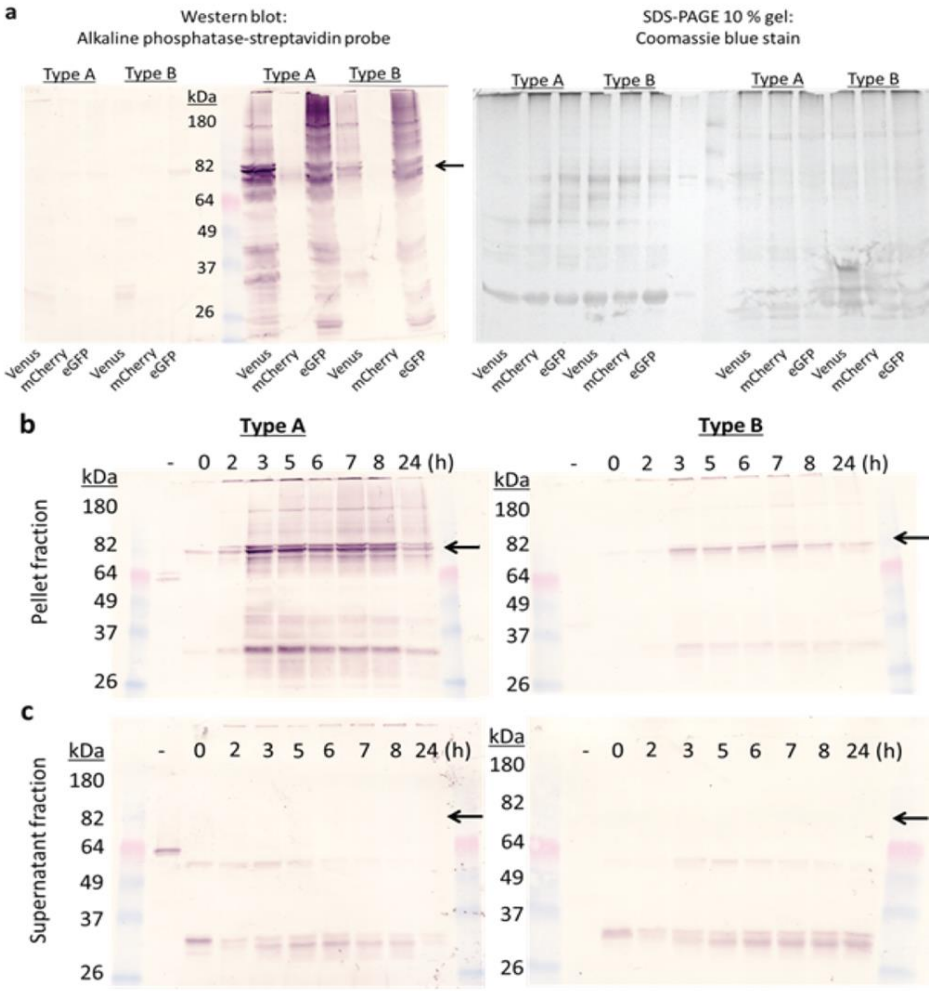


Figure S6 Western blot analysis of SBP-fluorophore-AIDAc membrane expression. Proteins isolated from Type A and B cells were size-separated by SDS-PAGE and blotted, using alkaline-phosphatase-linked streptavidin to probe for SBP-tagged proteins. (a) Western blot and corresponding SDS-PAGE gel for cytosolic (left of ladder) and membrane fractions (right of ladder) of induced cell extracts. Western blots for (b) pellet fractions and of (c) cytosolic fractions for Type A and B samples isolated at time intervals after induction with 28 μM AI-2. Arrows indicate the position of the 82 kDa SBP-fluorophore-AIDAc fusions.

9. AI-2 activity of W3110 *E. coli* during growth.

Autoinducer-2 bioassay. Autoinducer-2 was analyzed by a standard bioassay where bioluminescence intensity corresponds to relative AI-2 concentration.[86] Specifically, *Vibrio harveyi* strain BB170 generates bioluminescence in response to AI-2, and provides a quantifiable measure of extracellular AI-2 level. A standard assay was performed by diluting a BB170 16 h culture 1:5000 into fresh AB medium. Diluted BB170 (180 μL) was mixed with each sample (20 μL) and incubated in 5 mL culture tubes for 4.5 h at 30 °C, 250 rpm shaking, then recording the bioluminescence measurement from each sample. Luminescence measurements were obtained using a luminometer (EG&G Berthold, Gaithersburg, MD) and are reported as a fold change compared to a blank LB sample. To study the dynamics of AI-2 production by W3110, conditioned media were obtained from W3110 during growth in LB medium at 30 °C, 250 rpm shaking. The culture sample was pelleted and supernatant media was filter-sterilized. AI-2 activity from W3110 throughout growth is reported in Fig. S7, normalized to a sample assayed with blank LB medium.

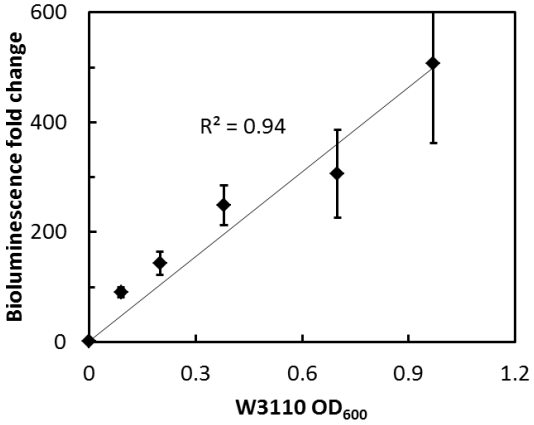


Figure S7 AI-2 activity reported by standard *Vibrio harveyi* bioassay. Bioluminescence units were measured from *V. harveyi* strain BB170 response to AI-2 and fold changes were normalized to a negative control. Samples consisted of conditioned media isolated from a W3110 culture during growth, correlated to the OD_{600} measurement at the time of sampling. Standard error of replicates and linear correlation of data is shown.

10. Evaluation of AI-2 responder coculture inoculated with cell types A and B.

First, we evaluated the robustness of the populations' fluorescence response as accurate AI-2 output. From the data in Fig. 4b, response patterns follow a Michaelis-Menten or Monod-type relationship. Thus, equations were established to mathematically characterize the percent of SBP-fluorescence expressing cells as a function of AI-2 concentration, using an empirically-defined Monod constant and adjusted maximum percentage for each population type (Table S3). The correlation of the FACS data and population models (with the maximum possible % set at 100) are shown in Fig. S8a; the general relationship between AI-2 and the fluorescently responding population is satisfied with R^2 correlations greater than 0.87.

Table S3 Equations representing fluorescent responder percentages as a function of autoinducer-2 concentration

	Responder A	**Responder B**
Equation *f(x) = %([A])* %: % expressers [A]: AI-2 (μM)	$\% = \frac{\%_{\max}([A])}{K_m + [A]}$	$\% = \frac{\%_{\max}([A] - C_{lag})}{K_m + [A] - 2C_{lag}}$
$\%_{max1}$	100	100
$\%_{max2}$	50	50
K_m	1.99	20.86
C_{lag}	--	5

To understand the necessary coculture-adjustments for the models, AI-2 responder strains A and B were evaluated for growth as a consequence of coculturing. First, Type A cells + pSBP-mCherry and Type B + pSBP-eGFP were cultured separately until mid-exponential phase, then induced with 25 μM AI-2. The growth rates, measured as the optical density at 600 nm (OD600), of A and B once induced were compared between single cultures and a coculture comprising a 1:1 volume ratio of each strain. Fig.S8c shows the OD600 trends for the cultures approximately 5 h post-induction. From this data, correlated growth of each culture type is observed throughout exponential (0 to 3.5 h) and stationary phase. Since the growth rates of Type A and B match, equal growth can be surmised in a coculture scenario, where the

total population composition is approximately 50 % each of the two cell types throughout incubation. Accordingly, the coculture growth data matches that of both single cultures, indicating that coculturing causes no change in growth dynamics for either cell type.

The models of A and B populations were used to evaluate our A and B expression data in coculture as a function of AI-2 (obtained from Fig. 4d). Based on Fig. S8c, we hypothesized that to fit the model in a coculture scenario, $\%_{max}$ becomes reduced to 50 (out of 100) to satisfy the strain's composition within the master population, but the response function to AI-2 should remain otherwise independent of coculturing. Fig. S8b plots red A and green B data as raw numbers (data from Fig. 4d diluted ten-fold) and the coculture-adjusted models as cell number (per 100 cells in the total population). General fits between data and models are observed for both cell types. These comparisons suggest that response to AI-2 from A and B populations is independent of coculturing, but that the muted fluorescent output, as a % of the population, is due to a lower $\%_{max}$ during coculture.

Finally, to fully evaluate the established behavior models for Type A and B cells in coculture, we considered the master population response, or in other words, the sum of A and B output, plotted in Fig. S8d. Additionally, the resolution of A, B, and the population as a whole, was considered across the range of AI-2, plotted as the change in response per μM AI-2 (Fig. S8d). From this analysis, the master population takes on properties of A and B at different subsidiary ranges. At low concentrations (< 5 μM), the master population response is primarily due to A turning on. The response in this range has a steep slope, giving both A and A+B a lower detection limit and higher resolution than B alone. However, the slope of the A quickly diminishes as the A-contributed response plateaus at the same range in which B contributes a dynamic response to AI-2 (> 5 μM). At this range, the total responding population is contributed by a steep accumulation of Type B and a more gradual accumulation of A cells, making the overall resolution higher than that of either A or B alone; moreover, the responsive range of the total population is expanded beyond that of either constituent. Thus, coculturing two responders in parallel

enhances the population-based AI-2 sensing while simultaneously generating a color-accumulation index (first, red A and second, green B).

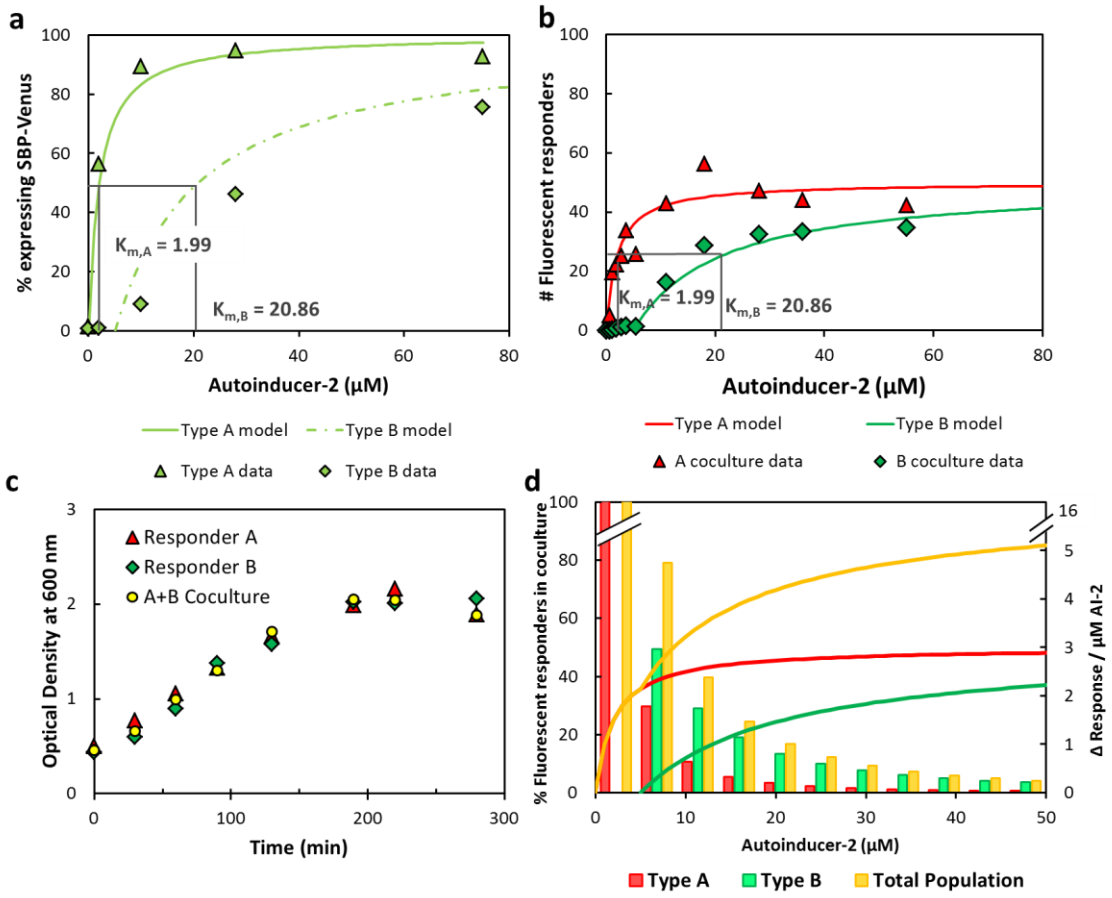


Figure S8 Analysis of AI-2 response by Type A and B populations in coculture. (a) FACS data overlaid with Michaelis-Menten or Monod-type models describing population percentage expressing SBP-Venus as a function of AI-2 concentration. (b) Population models with coculture-adjusted parameters (describing responding A and B number per 100 total cells) overlaid with relative cell numbers of red A and green B in coculture for an AI-2 concentration range. (c) Optical density, measured at 600 nm, of Type A and B cultures, isolated and cocultured (1:1 volumes A to B), as a function of time post-induction with 25 μM AI-2. (d) Modeled coculture response from A and B, described by (b), plotted along with their sum representing the total response from the master population. Additionally, the resolution of each curve (change in response per μM AI-2) is plotted for intervals between 0 and 50 μM AI-2.

11. Image analysis based identification of fluorescent *E. coli* vs. Syto9-stained *L. innocua*.

Image analysis of *L. innocua* and SBP-Venus expresser cocultures. Pixel-based particle analysis performed with ImageJ was used to quantify the unique size distributions of each cell type: *L. innocua* (Syto9-stained AI-2 producer) is definitively smaller than SBP-Venus-expressing BL21(DE3) (Fig. S9). Furthermore, fluorescence imaging was used to observe differential green and red signals through each respective filter. The Syto9-stained producer exhibits stronger green fluorescence than red, while the signal from the Venus-expressing cell is composed of nearly equal green and red intensities (Fig. S10). The size and fluorescence features of each cell type were used to quantify cell numbers for further image analyses.

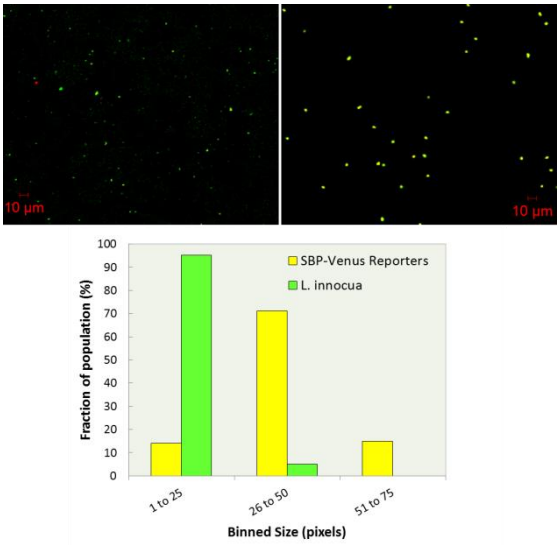


Figure S9 Image analysis of particle size from both Syto-9 stained *L. innocua* and SBP-Venus reporter fluorescence.

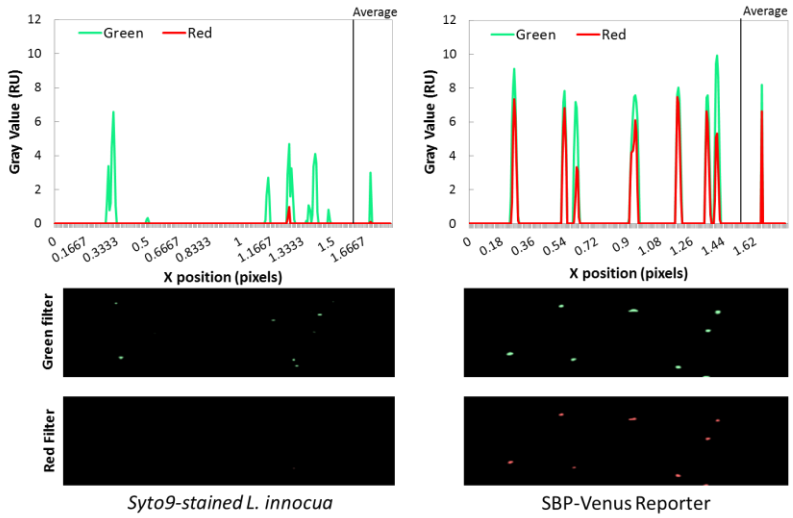


Figure S10 Image analysis of red and green pixels from both Syto-9 stained *L. innocua* and SBP-Venus reporter fluorescence.

12. Magnetic activated cell sorting (MACS) characterization using streptavidin-conjugated magnetic nanoparticles.

Magnetic activated cell sorting. Culture samples were first washed by resuspension in cold PBS. Initially, *L. innocua* was fluorescently stained with Syto9 nucleic acid stain (Life Technologies) according to protocol. Cell densities were readjusted to an OD of 0.2 and then diluted 10-fold in PBS. Mixed cultures were prepared using a 1:1 volumetric ratio of each cell type. Streptavidin-conjugated magnetic nanoparticles (mNPs, 100 nm diameter) were first washed in cold PBS, then diluted from stock concentration approximately 20-fold into cell samples, creating a 50,000-fold particle-to-cell excess, similar to previously published methods.[127]

After incubating the cells and mNPs in a 96-well plate at 4 °C for 20 min, a 2.2 x 1.6 cm (dia. x ht.) neodymium N42 magnet (K and J Magnetics) was placed under each well for 3 min to collect mNPs on the bottom surface. The supernatant was removed and particles were washed twice by resuspending in the same volume of cold PBS and magnetically recollecting; the first supernatant, reproducibly containing the vast majority of unbound cells, is referred to as the eliminated fraction. The recovered sample, consisting of the magnetically collected pellet, was resuspended in the same PBS volume. Fractions were counted by imaging samples at 200x magnification. To image magnetic collection *in situ*, a 1.6 x 0.8 mm (dia. x ht.) N52 neodymium magnet was taped behind a coverslip. A 2 μL mixture of cells and mNPs was added to the opposite side of the coverslip, directly on top of the magnet, set for 2 min, after which another coverslip sealed the sample for imaging.

Focusing magnetically-equipped populations of processor cells. We evaluated the application of magnetic collation as a form of abiotic processing in order to focus magnetically-equipped cell

units into aggregate multi-unit fluorescence data. The ability to selectively reposition "on" cells only was incorporated by design: since SBP is fused to the fluorophore terminus, SBP-surface exposure is conditional to fluorescence output, as is cell hybridization with streptavidin-conjugated mNPs (Fig. S11a). In this way, a magnetic field should filter AI-2 producers and off cells from a diverse distribution and simultaneously focus mNP-attached cells to the magnetic source.

We used induced cells with expressed SBP-Venus as the processing unit and *L. innocua* as the AI-2 producer. The AI-2 responder emitted fluorescence due to Venus and we used a Syto9 nucleic acid stain to make the AI-2 producer also fluoresce. It was verified that the cell types could be tracked by fluorescence microscopy and distinguished from one another. Image processing-based tracking of the two cell types is shown by Figs. S11 and S12. We performed experiments where the cells were mixed with mNPs to briefly allow hybridization, then a magnet was applied. This process formed two fractions: a focused fraction (magnetically recovered cells), contents concentrated at the magnet surface, and a filtered fraction (eliminated cells), which comprised unbound cells remaining in the supernatant and separated from the magnetized portion (Fig. S11a).

First, we directly observed the magnetic behavior of cell-mNP complexes by imaging the samples *in situ* with a magnet. A 1.6 mm diameter neodymium disc was fixed beneath a coverslip; subsequently the cell-mNP mixture was added to the coverslip and a composite image was taken to include directly adjacent fields of view. Fig. S11b shows the resulting image, with merged brightfield and fluorescence views, overlaid with a plot where the number of cells were counted as a function of position across the field of view (x position). The magnet is visible as the dark section on the left-hand side of the image. At the location of the magnet, the *recovered*

cell density is observably higher, by a count of at least 100-fold, than the *eliminated* cell density at a distance from its edge; the sharp drop in cell count occurs within 1 mm of the magnet's edge. Contrasted against a diffuse distribution of processor cells, each transmitting a single optical datum, collective fluorescence output from the focused population is significantly sharper and more obvious. Thus, magnetic enrichment is effective at focusing the individual fluorescence signals due to the physical act of concentration.

To quantitatively characterize the efficacy of magnetic parsing, cells in sample fractions were counted from images taken before (pre-) and after magnetic parsing, at which point fractions were split between eliminated (E) and recovered (R). Magnetic redistribution for processor cells, AI-2 producers, and a mixture of the two were completed in triplicate. Results show consistency magnetic filtering across samples (Fig. S11), where each sample underwent two sequential separations. Venus-expressing processor cells were recovered at greater than 40 % while less than 30 % were eliminated. Conversely, an average of 80 % of the AI-2 producer, *L. innocua*, was eliminated and less than 5 % was recovered. When mixed together, performance was comparable to that of the unmixed samples. Of the 90 % of cells eliminated, 20 % were the processor cells and over 60 % were *L. innocua*. Nearly 40 % of the cells in the original sample were recovered; composed mostly of processors, with less than 5 % being *L. innocua*. Thus, magnetic post-processing significantly refined the mNP-equipped SBP-Venus expressers, filtering out the majority of other constituents. The magnetic field also compacted the position of the Venus-expressing majority, thereby concentrating the fluorescence signal into collective output (Fig. S11). These data verified the specificity of MACS for SBP-expressers over *L. innocua* and provided evidence that self-assembled mNP-cell complexes are capable of selective positioning, prompted by the stimulus of a magnetic field.

Compared to the results described by Fig. S4, where SBP-dependent specificity was evaluated, results in Fig. S11c show improved specificity, although non-fluorescing non-expressers were excluded. Fig. S11c instead compares specificity against *L. innocua*, which in addition to lacking SBP, is also Gram-positive and therefore has a distinctly different cell wall than *E. coli*. It is not clear whether competition for streptavidin binding in mixed cultures, which occurs more favorably for SBP-expressers compared to non-expressers (Fig. S4d) and compared to *L. innocua* (Fig. S11c), results in interactions with enhanced specificity. Regardless, a key advantage is that since SBP-expressers are favorably recovered, their inclusion with non-specifically-bound cells does not impact the reliability of the fluorescent result; fluorescence output is inherent to positive expression and has been separately shown to be specific to the molecular input with low noise (Fig. S5c).

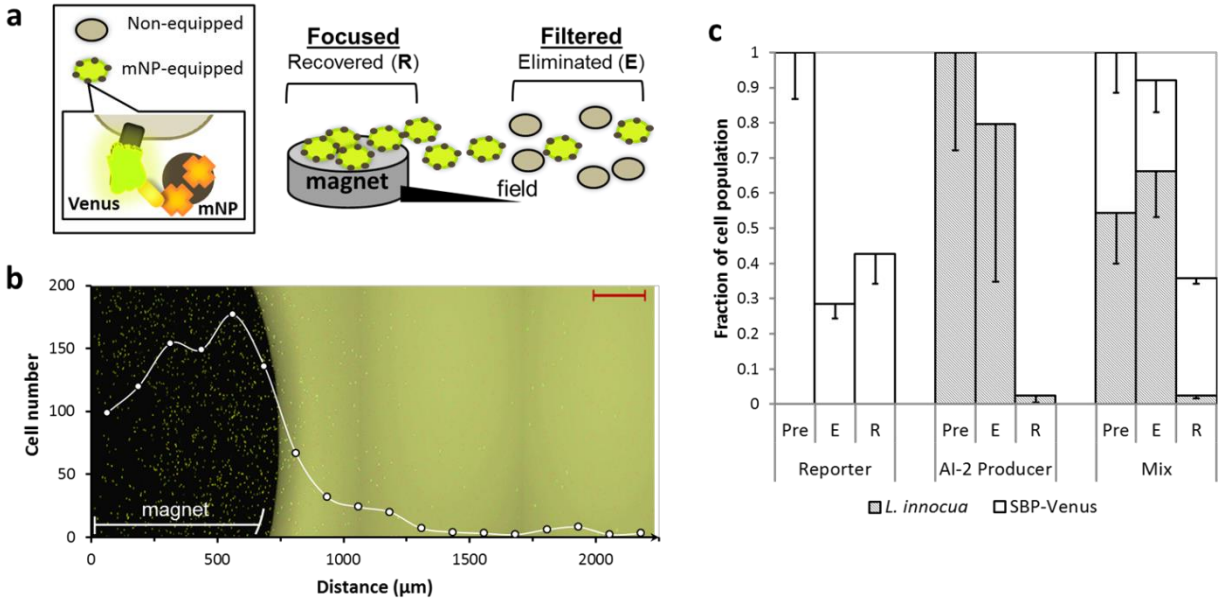


Figure S11 Filtering magnetically-equipped cells from AI-2 producers in order to focus fluorescence within a magnetic field. (a) Magnetic refining was performed applying a magnet to collect SBP-expressing cells equipped with streptavidin-functionalized mNPs into a pellet. Both the focusing and filtering properties of a magnetic field were investigated. (b) Observations of magnetic focusing were visualized by applying a mixture of cells and mNPs directly on top of a coverslip containing a magnet beneath. Tiling was used to compile an image of SBP-Venus expressers over 1 mm from the edge of the magnet. Cells were counted and plotted as a function of distance. Scalebar indicates 250 μm. (c) Magnetic filtering was measured by comparing the recovered counts of magnetically equipped cells against the eliminated count for non-equipped cells. Cells were counted from fluorescence images of the separated supernatant (eliminated fraction (E)) and the pellet (recovered fraction (R)), resuspended in an equal volume. The counts were compared in triplicate to the original number of cells in the mixture and reported as a fraction thereof. MACS was performed on SBP-Venus expressers, Syto-9 stained *L. innocua*, and a mixture of the two.

13. Magnetic refining effect on the feedback categorization and spatial distribution of fluorescent AI-2 responders.

We explored the expression of each cell type while cocultured across a range of AI-2 concentrations. Additionally, streptavidin-magnetic nanoparticles were applied to bind and concentrate all SBP-expressers within a magnetic field. Image analysis was used to calculate the ratio of red Type A to green Type B while distributed in culture and then at the higher density resulting from magnetic focusing; this data is plotted in Fig. S12a for each AI-2 condition. The analysis indicated that similar ratios between responders were maintained before and after magnetic collation. At all concentrations, the ratio is of the same order of magnitude and drops sharply between 5 and 11 μM AI-2, due to the emergence of green responders. Fig. S12b, shows a comparison of red and green output from cocultures with *L. innocua* at either low or high cell densities, both while the AI-2 processing cells are distributed throughout the culture and then after they have been magnetically focused. Visually, a higher fluorescence density is apparent and enables collective interpretation of population feedback.

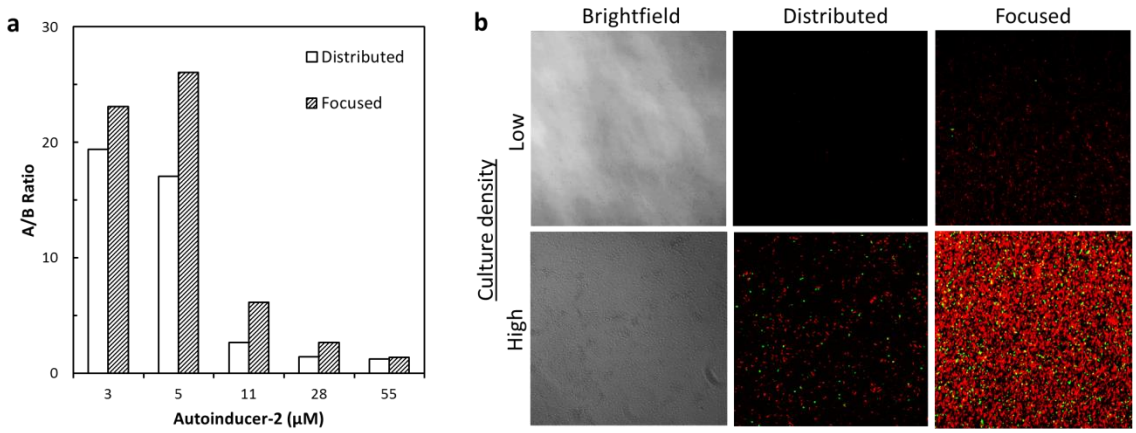


Figure S12 Magnetic refining of two-color responding A and B cells from a master population. (a) Ratio of Type A cells expressing SBP-mCherry to B cells expressing SBP-eGFP while distributed in culture and upon focusing to a magnetic surface. (b) Images of visual space occupied as a result of AI-2 information processing in a *L. innocua cu*lture in parallel by AI-2 responder types A (SBP-mCherry +) and B (SBP-eGFP +). Culture conditions were biased between a low and high cell density. Images depict the multi-dimenstional system of *L. innocua*, responder populations A and B, and magnetic nanoparticles for self-assembly with responders in three settings: brightfield, distributed (culture before magnet application), and focused (during magnet application).

14. Coculturing dynamics between the AI-2 producer *L. innocua* and *E. coli* responders.

The dynamics of coculturing were elucidated through a combination of tracking the optical density (OD) of cultures at 600 nm, plated colony counts, and gram staining. Single- and co-cultures were inoculated identically with relevant strains. Throughout the initial 3 h of incubation, the coculture OD followed that of *L. innocua*, while growth from A and B responders was negligible (Fig. S13a). This indicates that *L. innocua* proliferation was predominant within the coculture, and further, that its growth was unaffected by the presence of *E. coli* cells. After 3h, *L. innocua* reached stationary phase at an approximate OD of 0.9, which is consistent with its reported range[145]. At the same time, the growth of A and B responders became measurable and indicative of their exponential growth (Fig. S13a). The OD of the coculture also continued to increase, presumably due to the exponential accumulation of A and B despite stationary *L. innocua* growth.

Fig. S13b reports the measured final ODs after 10 h growth and the onset of stationary phase in each culture. Observed, A and B responders reach similar ODs (2.61 and 2.73, respectively), while the growth rate for *L. innocua* does not surpass 0.9. The coculture reaches an OD of 2.7, the same as the culture density achieved by A and B. The coculture data suggests that because of an early null net growth from *L. innocua*, the *E. coli* constituents were eventually (after 4 h) able to outgrow *L. innocua*, perhaps by scavenging unused nutrients[145]. However, the fraction of *L. innocua* maintained at stationary phase was not clear from this data.

To investigate the final coculture composition further, culture dilutions were plated onto agar mixed with either brain-heart infusion (BHI) medium, which supports the growth of both *L. innocua* and *E. coli*, or Luria broth (LB) with kanamycin and ampicillin (50 μg/mL each), permitting only colonies from A and B responders. Colony counts revealed a 1.6 coculture ratio between A + B and *L. innocua*, where the responders comprised just under two-thirds of the culture (Fig. S13c). This fraction was found to be consistent with the OD analysis, given the postulation that the *L. innocua* density remains steady. Thus,

evidence suggests that *L. innocua* competes sufficiently for nutrients during *E. coli* accumulation in coculture.

The coculture was determined to transition from a state of *L. innocua* abundance to one predominated by A and B responders. Two visualize these two prevailing conditions, Gram staining was applied to samples and imaged. *L. innocua*, a Gram-positive strain, stains violet and *E. coli*, Gram-negative, stains pink-red. In Fig. S13d, a sample isolated at 3 h reveals a majority of *L. innocua* cells while a 10 h sample (Fig. 13e) shows a prevalence of *E. coli* in addition to *L. innocua*. The images, thus, corroborate the coculture data.

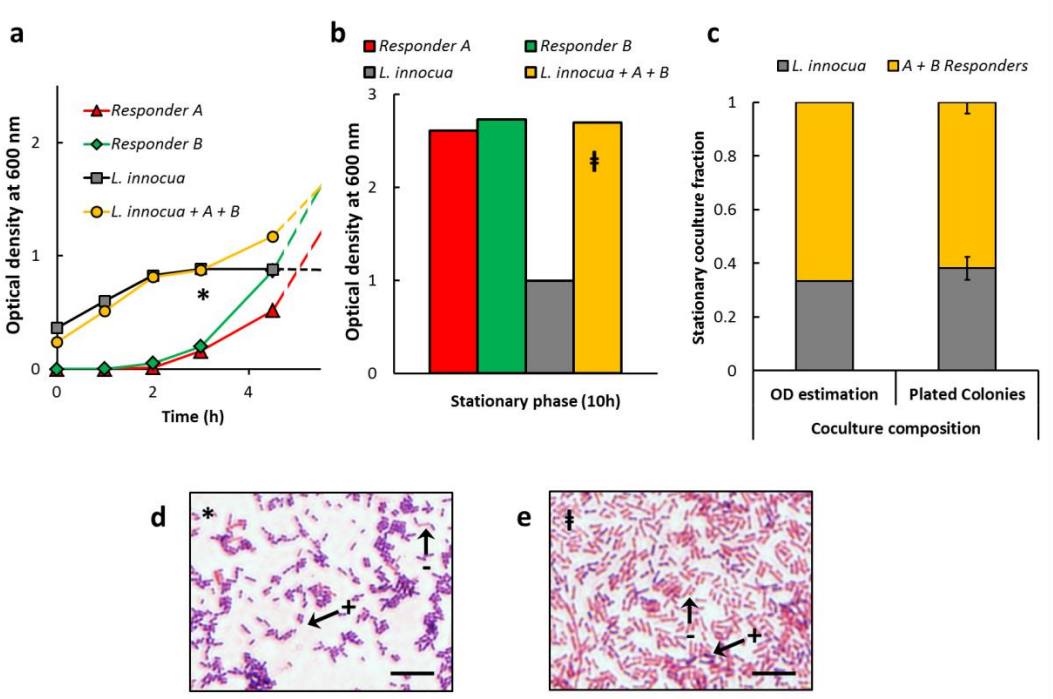


Figure S13 Culture growth dynamics in isolated and mixed-species cultures. The growth of *Listeria innocua* in a mixed medium (3:1 Luria broth to brain-heart infusion) was compared against *E. coli* responders A and B when grown separately and when inoculated identically into a mixed-species culture. (a) Optical density measurements (at 600 nm) of cultures during the first 5 h of growth. (b) Final OD measurements 10 h after all cultures reached stationary phase. (c) The final coculture composition was calculated by counting colonies from species-selective plating and compared against the stationary OD of *L. innocua* as a fraction of that of the coculture. (d) and (e) show Gram staining images of the coculture at representative timepoints: 3 h (*) and 10 h (ł). Examples of Gram-postive (+) cells, stained violet, and Gram-negative (-) cells, stained pink, are indicated by arrows. Scale bars indicate 5 µm.

15. Information processing by distribution among *Listeria innocua.*

Cocultures were established directly with an AI-2 producer under conditions that biased AI-2 accumulation such that it presided within each responder's resolution range. Having established that type A cells elicit sensitive response at low AI-2 levels and type B during high AI-2 exposure, we explored the ability of each cell type to engage in processing AI-2 information in an environmental scenario. *Listeria innocua* was used as a target AI-2-producing cell type at fixed initial proportions of AI-2 producer to processing unit in varied culturing media.

First, we tested non-ideal growth conditions in order to evaluate the ability of population A to respond to AI-2 while coresiding with *L. innocua* in a growth-minimized scenario. To create a growth-limiting environment, brain-heart infusion (BHI) was mixed with Luria-Bertani broth (LB) and diluted 10-fold in sterile PBS. A test culture showed, through plating and colony counts, that *L. innocua* proliferated by only three generations in these conditions. *L. innocua* was grown to an OD of 0.3, then rinsed and resuspended in the diluted media. The culture was serially diluted 10-fold and 100-fold, then incubated with population A, resulting in producer-to-processor ratios of 2, 20, and 200, from the most dilute to most dense initial conditions.

Alternatively, we used richer media to portray more optimal conditions for growth and proliferation, having essential nutrients readily available. We observed that cocultures with *L. innocua* proliferate by approximately six generations to reach cell densities at least 10^2 higher than the nutrient-limiting conditions (Fig. S14, bottom). Consequently, we used these conditions to support high microbial community densities and tested the capability of population B to interpret AI-2 activity. *L. innocua* was adapted to media composed of 25 % BHI in LB and probed at low exponential phase points, throughout which AI-2 accumulation occurs.[145]

Population B was distributed directly into the growing *L. innocua* culture to create inoculation ratios with the AI-2 producer between 75 and 600:1 and coincubated.

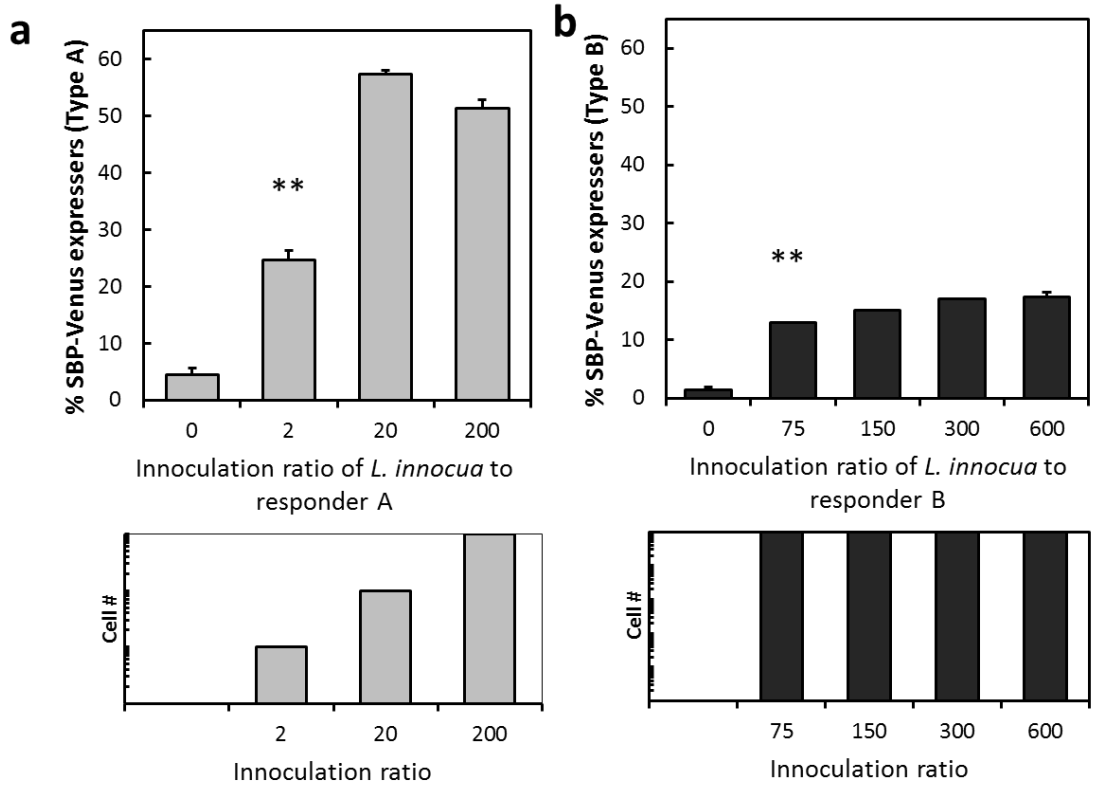


Figure S14 Effect of initial culture density on quorum sensing-surveillance of *Listeria innocua*. Upper graphs show Venus expression by populations distributed within cultures of the AI-2 producer *L. innoua*, where the inoculation ratio between the producer and responder cells was varied. Below, final coculture densities are reported below as an order of magnitude. (a) Population A assessed AI-2 prevalence in conditions of limited nutrients and (b) B in conditions of rich nutrients. Symbols denote the onset of statistical significant of triplicate data compared to OD_{600} 0. Statistical significance of data is represented by * for $p < 0.05$ and ** for $p < 0.005$.

The FACS analyses in Fig. S14 (top graphs) compare the "consensus" percentages post-culturing, or in other words, the cells in which fluorescence has been triggered after processing available AI-2. The data show that quorum sensing by *L. innocua* is detectable by population A when introduced to the thinnest population (10^6 CFUs, established from a 2:1 ratio with *L.*

innocua) (Fig. S14a). While Type B was not statistically different from the control in the dilute conditions, an emergence of Venus-expressers is observed in conditions permitting higher cell density; the trend starts at a producer-to-processor ratio of 75:1, with the resulting percentage of "on" units being statistically significant (Fig. S14b). Not reported, we observed that the medium composition, particularly BHI, interfered with induction of AI-2 processing. Previously, interference in standard AI-2 detection methods for *L. monocytogenes* has was attributed to glucose content and glucose has also been shown to favor QS repression in *E. coli* QS reporting. [163,165] Nevertheless, detection of *L. innocua* was possible in each medium here and exhibits the property of reliable performance in dynamic living environments, enabled by the use of many cell "processing units" for population-based information acquisition.

16. Dual population responses to molecular inputs that yield unique combinatorial outputs.

Table S4 Equations for parallel populations that generate red and green response patterns with input *x*.

	Red $f(x)$	**Green $f(x)$**
i	$\frac{x^2}{5}+\frac{x}{20}$	$\frac{x^2}{10}+\frac{1}{2}$
ii	$5x+50$	$5x-40$
iii	$10x$	$\frac{400x}{x+10}$
iv	$5x+50$	$x+10$
v	$\frac{1000x}{x+30}$	$5x+10$
vi	$\frac{x^3+x^2+200}{400}$	$\frac{600x}{x+3}$
vii	$\frac{300x}{2+x}$	$\frac{200x}{20+x}$
viii	$\frac{x^2+100x+30}{5}$	$2x+2$
ix	$x+2$	$\frac{x^2+50x+30}{5}$
x	$100-100\sin(\frac{x}{2})$	$2x$

17. Parameterization of two standard response curves to influence two-color output patterns.

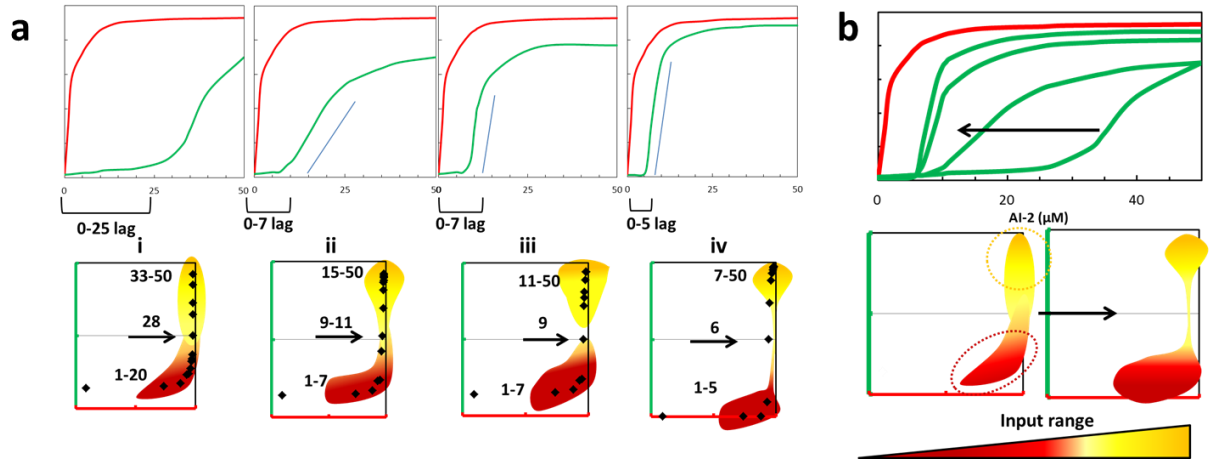


Figure S15 Hypothetical response curves to an input range, leading to red and green output. (a) Combinations of response curve pairs, and visual planes of the appearance of red (x-axis) versus green (y-axis) output levels. Input values are defined across the plane. (b) The trend of changing a green response curve with respect to red is summarized. A heat map representation of the data indicates the shift from low to high input concentrations.

References

1. Cheng AA, Lu TK: **Synthetic biology: an emerging engineering discipline**. *Annu Rev Biomed Eng* 2012, **14**:155-178.
2. Horner M, Reischmann N, Weber W: **Synthetic biology: programming cells for biomedical applications**. *Perspect Biol Med* 2012, **55**:490-502.
3. Esch MB, King TL, Shuler ML: **The Role of Body-on-a-Chip Devices in Drug and Toxicity Studies**. *Annu Rev Biomed Eng* 2010.
4. Gao W, Wang J: **The environmental impact of micro/nanomachines: a review**. *ACS Nano* 2014, **8**:3170-3180.
5. Forlin M, Lentini R, Mansy SS: **Cellular imitations**. *Current Opinion in Chemical Biology* 2012, **16**:586-592.
6. Gine LP and Akyildiz IF: **Molecular communication options for long range nanonetworks**. *Comput. Netw.* 2009, **53**:2753-2766.
7. Cobo LC, Akyildiz IF: **Bacteria-based communication in nanonetworks**. *Nano Communication Networks* 2010, **1**:244-256.
8. Adleman LM: **Molecular computation of solutions to combinatorial problems**. *Science* 1994, **266**:1021-1024.
9. Rudchenko M, Taylor S, Pallavi P, Dechkovskaia A, Khan S, Butler VP, Jr., Rudchenko S, Stojanovic MN: **Autonomous molecular cascades for evaluation of cell surfaces**. *Nat Nanotechnol* 2013, **8**:580-586.
10. Ford TJ, Silver PA: **Synthetic biology expands chemical control of microorganisms**. *Current Opinion in Chemical Biology* 2015, **28**:20-28.
11. Callura JM, Cantor CR, Collins JJ: **Genetic switchboard for synthetic biology applications**. *Proc Natl Acad Sci U S A* 2012, **109**:5850-5855.
12. Stricker J, Cookson S, Bennett MR, Mather WH, Tsimring LS, Hasty J: **A fast, robust and tunable synthetic gene oscillator**. *Nature* 2008, **456**:516-519.
13. Bonnet J, Subsoontorn P, Endy D: **Rewritable digital data storage in live cells via engineered control of**. *Proc Natl Acad Sci U S A* 2012, **109**:8884-8889.
14. Khalil AS, Collins JJ: **Synthetic biology: applications come of age**. *Nat Rev Genet* 2010, **11**:367-379.
15. Saeidi N, Wong CK, Lo TM, Nguyen HX, Ling H, Leong SS, Poh CL, Chang MW: **Engineering microbes to sense and eradicate Pseudomonas aeruginosa, a human**. *Mol Syst Biol* 2011, **7**:521.
16. Wu HC, Tsao CY, Quan DN, Cheng Y, Servinsky MD, Carter KK, Jee KJ, Terrell JL, Zargar A, Rubloff GW, et al.: **Autonomous bacterial localization and gene expression based on nearby cell receptor density**. *Mol Syst Biol* 2013, **9**:636.
17. Chen X, Goodnight D, Gao Z, Cavusoglu AH, Sabharwal N, DeLay M, Driks A, Sahin O: **Scaling up nanoscale water-driven energy conversion into evaporation-driven engines and generators**. *Nat Commun* 2015, **6**.
18. Weibel DB, Garstecki P, Ryan D, DiLuzio WR, Mayer M, Seto JE, Whitesides GM: **Microoxen: microorganisms to move microscale loads**. *Proc Natl Acad Sci U S A* 2005, **102**:11963-11967.
19. Fernandes R, Zuniga M, Sassine FR, Karakoy M, Gracias DH: **Enabling cargo-carrying bacteria via surface attachment and triggered release**. *Small* 2011, **7**:588-592.
20. Liu Y, Kim E, Ghodssi R, Rubloff GW, Culver JN, Bentley WE, Payne GF: **Biofabrication to build the biology-device interface**. In *Biofabrication*. Edited by; 2010:022002. vol 2.]
21. Wu LQ, Payne GF: **Biofabrication: using biological materials and biocatalysts to construct nanostructured assemblies**. *Trends in Biotechnology* 2004, **22**:593-599.
22. Wu LQ, Bentley WE, Payne GF: **Biofabrication with biopolymers and enzymes: potential for constructing scaffolds from soft matter**. *Int J Artif Organs* 2011, **34**:215-224.

23. Park TH, Shuler ML: **Integration of cell culture and microfabrication technology**. *Biotechnol Prog* 2003, **19**:243-253.
24. Sung JH, Yu J, Luo D, Shuler ML, March JC: **Microscale 3-D hydrogel scaffold for biomimetic gastrointestinal (GI) tract model**. *Lab Chip* 2011, **11**:389-392.
25. Palacios MA, Benito-Peña E, Manesse M, Mazzeo AD, LaFratta CN, Whitesides GM, Walt DR: **InfoBiology by printed arrays of microorganism colonies for timed and on-demand release of messages**. *Proceedings of the National Academy of Sciences* 2011.
26. Elad T, Lee JH, Gu MB, Belkin S: **Microbial cell arrays**. *Adv Biochem Eng Biotechnol* 2010, **117**:85-108.
27. Zakeri B, Lu TK: **DNA nanotechnology: new adventures for an old warhorse**. *Curr Opin Chem Biol* 2015, **28**:9-14.
28. Venter JC: *Life at the speed of light: from the double helix to the dawn of digital life*: Penguin; 2013.
29. Cox JP: **Long-term data storage in DNA**. *Trends Biotechnol* 2001, **19**:247-250.
30. Church GM, Gao Y, Kosuri S: **Next-generation digital information storage in DNA**. *Science* 2012, **337**:1628.
31. He Y, Liu DR: **Autonomous multistep organic synthesis in a single isothermal solution mediated**. *Nat Nanotechnol* 2010, **5**:778-782.
32. Pereira CS, Thompson JA, Xavier KB: **AI-2-mediated signalling in bacteria**. *FEMS Microbiol Rev* 2013, **37**:156-181.
33. Zhang J, Jensen MK, Keasling JD: **Development of biosensors and their application in metabolic engineering**. *Current Opinion in Chemical Biology* 2015, **28**:1-8.
34. Basu S, Gerchman Y, Collins CH, Arnold FH, Weiss R: **A synthetic multicellular system for programmed pattern formation**. *Nature* 2005, **434**:1130-1134.
35. Terrell JL, Gordonov T, Cheng Y, Wu HC, Sampey D, Luo X, Tsao CY, Ghodssi R, Rubloff GW, Payne GF, et al.: **Integrated biofabrication for electro-addressed in-film bioprocessing**. *Biotechnol J* 2012, **7**:428-439.
36. Schapper D, Alam MN, Szita N, Eliasson Lantz A, Gernaey KV: **Application of microbioreactors in fermentation process development: a review**. *Anal Bioanal Chem* 2009, **395**:679-695.
37. Chen A, Chitta R, Chang D, Amanullah A: **Twenty-four well plate miniature bioreactor system as a scale-down model for cell culture process development**. *Biotechnol Bioeng* 2009, **102**:148-160.
38. Diao J, Young L, Zhou P, Shuler ML: **An actively mixed mini-bioreactor for protein production from suspended animal cells**. *Biotechnol Bioeng* 2008, **100**:72-81.
39. Rao G, Moreira A, Brorson K: **Disposable bioprocessing: the future has arrived**. *Biotechnol Bioeng* 2009, **102**:348-356.
40. Duetz WA: **Microtiter plates as mini-bioreactors: miniaturization of fermentation methods**. In *Trends Microbiol*. Edited by; 2007:469-475. vol 15.]
41. Szmacinski H, Smith DS, Hanson MA, Kostov Y, Lakowicz JR, Rao G: **A novel method for monitoring monoclonal antibody production during cell culture**. *Biotechnol Bioeng* 2008, **100**:448-457.
42. Shi X, Lin LI, Chen SY, Chao SH, Zhang W, Meldrum DR: **Real-time PCR of single bacterial cells on an array of adhering droplets**. *Lab Chip* 2011, **11**:2276-2281.
43. Wu W, Kang KT, Lee NY: **Bubble-free on-chip continuous-flow polymerase chain reaction: concept and application**. *Analyst* 2011, **136**:2287-2293.
44. Miao J, Wu W, Spielmann T, Belfort M, Derbyshire V, Belfort G: **Single-step affinity purification of toxic and non-toxic proteins on a fluidics platform**. *Lab Chip* 2005, **5**:248-253.
45. Lai JJ, Nelson KE, Nash MA, Hoffman AS, Yager P, Stayton PS: **Dynamic bioprocessing and microfluidic transport control with smart magnetic nanoparticles in laminar-flow devices**. *Lab Chip* 2009, **9**:1997-2002.

46. Bareither R, Pollard D: **A review of advanced small-scale parallel bioreactor technology for accelerated process development: current state and future need**. *Biotechnol Prog* 2011, **27**:2-14.
47. Wu LQ, Payne GF: **Biofabrication: using biological materials and biocatalysts to construct nanostructured assemblies**. In *Trends Biotechnol*. Edited by; 2004:593-599. vol 22.]
48. Chen TH, Small DA, Wu LQ, Rubloff GW, Ghodssi R, Vazquez-Duhalt R, Bentley WE, Payne GF: **Nature-inspired creation of protein-polysaccharide conjugate and its subsequent assembly onto a patterned surface**. *Langmuir* 2003, **19**:9382-9386.
49. Wu HC, Shi XW, Tsao CY, Lewandowski AT, Fernandes R, Hung CW, DeShong P, Kobatake E, Valdes JJ, Payne GF, et al.: **Biofabrication of antibodies and antigens via IgG-binding domain engineered with activatable pentatyrosine pro-tag**. *Biotechnol Bioeng* 2009, **103**:231-240.
50. Shi XW, Tsao CY, Yang XH, Liu Y, Dykstra P, Rubloff GW, Ghodssi R, Bentley WE, Payne GF: **Electroaddressing of Cell Populations by Co-Deposition with Calcium Alginate Hydrogels**. *Advanced Functional Materials* 2009, **19**:2074-2080.
51. Cheng Y, Luo X, Tsao CY, Wu HC, Betz J, Payne GF, Bentley WE, Rubloff GW: **Biocompatible multi-address 3D cell assembly in microfluidic devices using spatially programmable gel formation**. *Lab Chip* 2011.
52. Koev ST, Dykstra PH, Luo X, Rubloff GW, Bentley WE, Payne GF, Ghodssi R: **Chitosan: an integrative biomaterial for lab-on-a-chip devices**. *Lab Chip* 2010, **10**:3026-3042.
53. Sun F, Zhitomirsky I: **Electrochemical deposition of composite biopolymer films**. *Surface Engineering* 2010, **26**:546-551.
54. Boccaccini AR, Keim S, Ma R, Li Y, Zhitomirsky I: **Electrophoretic deposition of biomaterials**. In *J R Soc Interface*. Edited by; 2010:S581-613. vol 7 Suppl 5.]
55. Yang XH, Kim E, Liu Y, Shi XW, Rubloff GW, Ghodssi R, Bentley WE, Pancer Z, Payne GF: **In-Film Bioprocessing and Immunoanalysis with Electroaddressable Stimuli-Responsive Polysaccharides**. *Advanced Functional Materials* 2010, **20**:1645-1652.
56. Selimoglu SM, Elibol M: **Alginate as an immobilization material for MAb production via encapsulated hybridoma cells**. *Crit Rev Biotechnol* 2010, **30**:145-159.
57. Lee GM, Palsson BO: **Immobilization can improve the stability of hybridoma antibody productivity in serum-free media**. *Biotechnol Bioeng* 1990, **36**:1049-1055.
58. Akerstrom B, Brodin T, Reis K, Bjorck L: **Protein G: a powerful tool for binding and detection of monoclonal and polyclonal antibodies**. *J Immunol* 1985, **135**:2589-2592.
59. Lee JM, Park HK, Jung Y, Kim JK, Jung SO, Chung BH: **Direct immobilization of protein g variants with various numbers of cysteine residues on a gold surface**. *Anal Chem* 2007, **79**:2680-2687.
60. Tanaka G, Funabashi H, Mie M, Kobatake E: **Fabrication of an antibody microwell array with self-adhering antibody binding protein**. In *Anal Biochem*. Edited by; 2006:298-303. vol 350.]
61. Lewandowski AT, Small DA, Chen T, Payne GF, Bentley WE: **Tyrosine-based "activatable pro-tag": enzyme-catalyzed protein capture and release**. *Biotechnol Bioeng* 2006, **93**:1207-1215.
62. Anghileri A, Lantto R, Kruus K, Arosio C, Freddi G: **Tyrosinase-catalyzed grafting of sericin peptides onto chitosan and production of protein-polysaccharide bioconjugates**. In *J Biotechnol*. Edited by; 2007:508-519. vol 127.]
63. Freddi G, Anghileri A, Sampaio S, Buchert J, Monti P, Taddei P: **Tyrosinase-catalyzed modification of Bombyx mori silk fibroin: grafting of chitosan under heterogeneous reaction conditions**. In *J Biotechnol*. Edited by; 2006:281-294. vol 125.]
64. Chen T, Embree HD, Wu LQ, Payne GF: **In vitro protein-polysaccharide conjugation: tyrosinase-catalyzed conjugation of gelatin and chitosan**. *Biopolymers* 2002, **64**:292-302.
65. Cheng Y, Luo XL, Betz J, Payne GF, Bentley WE, Rubloff GW: **Mechanism of anodic electrodeposition of calcium alginate**. *Soft Matter* 2011, **7**:5677-5684.

66. Braschler T, Johann R, Heule M, Metref L, Renaud P: **Gentle cell trapping and release on a microfluidic chip by in situ alginate hydrogel formation**. *Lab Chip* 2005, **5**:553-559.
67. Wu LQ, Yi HM, Li S, Rubloff GW, Bentley WE, Ghodssi R, Payne GF: **Spatially selective deposition of a reactive polysaccharide layer onto a patterned template**. *Langmuir* 2003, **19**:519-524.
68. Polk A, Amsden B, De Yao K, Peng T, Goosen MF: **Controlled release of albumin from chitosan-alginate microcapsules**. *J Pharm Sci* 1994, **83**:178-185.
69. Tapia C, Montezuma V, Yazdani-Pedram M: **Microencapsulation by Spray Coagulation of Diltiazem HCl in Calcium Alginate-Coated Chitosan**. *AAPS PharmSciTech* 2008, **9**:1198-1206.
70. Barbucci R, Dash M, Piras AM, Chiellini F: **Chitosan-Based Beads for Controlled Release of Proteins**. In *Hydrogels*. Edited by: Springer Milan; 2009:111-120.
71. Wu LQ, Lee K, Wang X, English DS, Losert W, Payne GF: **Chitosan-mediated and spatially selective electrodeposition of nanoscale particles**. *Langmuir* 2005, **21**:3641-3646.
72. Church GM, Regis E: *Regenesis: how synthetic biology will reinvent nature and ourselves*: Basic Books; 2014.
73. Pak Chung W, Kwong-Kwok W, Foote H: **ORGANIC DATA MEMORY Using the DNA Approach**. *Communications of the ACM* 2003, **46**:95-98.
74. Zala K: **Q&A: Poetry in the genes**. *Nature* 2009, **458**:35-35.
75. Gehani A, LaBean T, Reif J: **DNA-based Cryptography**. In *Aspects of Molecular Computing*. Edited by Jonoska N, Păun G, Rozenberg G: Springer Berlin Heidelberg; 2004:167-188. Lecture Notes in Computer Science, vol 2950.]
76. Clelland CT, Risca V, Bancroft C: **Hiding messages in DNA microdots**. *Nature* 1999, **399**:533-534.
77. Liss M, Daubert D, Brunner K, Kliche K, Hammes U, Leiherer A, Wagner R: **Embedding Permanent Watermarks in Synthetic Genes**. *PLoS ONE* 2012, **7**:e42465.
78. Sussenbach JS, Monfoort CH, Schiphof R, Stobberingh EE: **A restriction endonuclease from Staphylococcus aureus**. *Nucleic Acids Research* 1976, **3**:3193-3202.
79. Coffin SR, Reich NO: **Modulation of Escherichia coli DNA methyltransferase activity by biologically**. *J Biol Chem* 2008, **283**:20106-20116.
80. Patnaik PK, Merlin S, Polisky B: **Effect of altering GATC sequences in the plasmid ColE1 primer promoter**. *J Bacteriol* 1990, **172**:1762-1768.
81. Russell DW, Horiuchi K: **The mutH gene regulates the replication and methylation of the pMB1 origin**. *Journal of Bacteriology* 1991, **173**:3209-3214.
82. Frenkel L, Bremer H: **Increased amplification of plasmids pBR322 and pBR327 by low concentrations of chloramphenicol**. *Dna* 1986, **5**:539-544.
83. Sharp PM, Cowe E, Higgins DG, Shields DC, Wolfe KH, Wright F: **Codon usage patterns in Escherichia coli, Bacillus subtilis, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Drosophila melanogaster and Homo sapiens; a review of the considerable within-species diversity**. *Nucleic Acids Research* 1988, **16**:8207-8211.
84. Ikemura T: **Codon usage and tRNA content in unicellular and multicellular organisms**. *Molecular Biology and Evolution* 1985, **2**:13-34.
85. Servinsky MD, Terrell JL, Tsao C-Y, Wu H-C, Quan DN, Zargar A, Allen PC, Byrd CM, Sund CJ, Bentley WE: **Directed assembly of a bacterial quorum**. *ISME J* 2015.
86. Surette MG, Bassler BL: **Quorum sensing in Escherichia coli and Salmonella typhimurium**. *Proc Natl Acad Sci U S A* 1998, **95**:7046-7050.
87. Vendeville A, Winzer K, Heurlier K, Tang CM, Hardie KR: **Making 'sense' of metabolism: autoinducer-2, LuxS and pathogenic bacteria**. *Nat Rev Microbiol* 2005, **3**:383-396.
88. Lewandowski AT, Bentley WE, Yi H, Rubloff GW, Payne GF, Ghodssi R: **Towards area-based in vitro metabolic engineering: assembly of Pfs enzyme onto**. *Biotechnol Prog* 2008, **24**:1042-1051.

89. Luo X, Lewandowski AT, Yi H, Payne GF, Ghodssi R, Bentley WE, Rubloff GW: **Programmable assembly of a metabolic pathway enzyme in a pre-packaged reusable**. *Lab Chip* 2008, **8**:420-430.
90. Leduc PR, Wong MS, Ferreira PM, Groff RE, Haslinger K, Koonce MP, Lee WY, Love JC, McCammon JA, Monteiro-Riviere NA, et al.: **Towards an in vivo biologically inspired nanofactory**. *Nat Nanotechnol* 2007, **2**:3-7.
91. Fernandes R, Roy V, Wu HC, Bentley WE: **Engineered biological nanofactories trigger quorum sensing response in targeted**. *Nat Nanotechnol* 2010, **5**:213-217.
92. Marinus MG, Casadesus J: **Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more**. *FEMS Microbiol Rev* 2009, **33**:488-503.
93. Horton JR, Liebert K, Bekes M, Jeltsch A, Cheng X: **Structure and substrate recognition of the Escherichia coli DNA adenine methyltransferase**. *J Mol Biol* 2006, **358**:559-570.
94. Urig S, Gowher H, Hermann A, Beck C, Fatemi M, Humeny A, Jeltsch A: **The Escherichia coli dam DNA methyltransferase modifies DNA in a highly**. *J Mol Biol* 2002, **319**:1085-1096.
95. Schauder S, Shokat K, Surette MG, Bassler BL: **The LuxS family of bacterial autoinducers: biosynthesis of a novel quorum-sensing signal molecule**. In *Mol Microbiol*. Edited by; 2001:463-476. vol 41.]
96. Winzer K, Hardie KR, Burgess N, Doherty N, Kirke D, Holden MT, Linforth R, Cornell KA, Taylor AJ, Hill PJ, et al.: **LuxS: its role in central metabolism and the in vitro synthesis of 4-hydroxy-5-methyl-3(2H)-furanone**. *Microbiology* 2002, **148**:909-922.
97. de la Campa AG, Springhorn SS, Kale P, Lacks SA: **Proteins encoded by the DpnI restriction gene cassette. Hyperproduction and characterization of the DpnI endonuclease**. *J Biol Chem* 1988, **263**:14696-14702.
98. Holmes ML, Nuttall SD, Dyall-Smith ML: **Construction and use of halobacterial shuttle vectors and further studies on Haloferax DNA gyrase**. *J Bacteriol* 1991, **173**:3807-3813.
99. Palmer BR, Marinus MG: **The dam and dcm strains of Escherichia coli--a review**. *Gene* 1994, **143**:1-12.
100. Marinus MG, Morris NR: **Pleiotropic effects of a DNA adenine methylation mutation (dam-3) in Escherichia coli K12**. *Mutat Res* 1975, **28**:15-26.
101. Blanco L, Bernad A, Lázaro JM, Martin G, Garmendia C, Salas M: **Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication**. *Journal of Biological Chemistry* 1989, **264**:8935-8940.
102. Dean FB, Nelson JR, Giesler TL, Lasken RS: **Rapid amplification of plasmid and phage DNA using phi29 DNA polymerase and multiply-primed rolling circle amplification**. *Genome research* 2001, **11**:1095-1099.
103. Banér J, Nilsson M, Mendel-Hartvig M, Landegren U: **Signal amplification of padlock probes by rolling circle replication**. *Nucleic acids research* 1998, **26**:5073-5078.
104. O'Donoghue SI, Gavin AC, Gehlenborg N, Goodsell DS, Heriche JK, Nielsen CB, North C, Olson AJ, Procter JB, Shattuck DW, et al.: **Visualizing biological data-now and in the future**. *Nat Methods* 2010, **7**:S2-4.
105. Ghazalpour A, Bennett B, Petyuk VA, Orozco L, Hagopian R, Mungrue IN, Farber CR, Sinsheimer J, Kang HM, Furlotte N, et al.: **Comparative analysis of proteome and transcriptome variation in mouse**. *PLoS Genet* 2011, **7**:e1001393.
106. Lentini R, Santero SP, Chizzolini F, Cecchi D, Fontana J, Marchioretto M, Del Bianco C, Terrell JL, Spencer AC, Martini L, et al.: **Integrating artificial with natural cells to translate chemical messages that direct E. coli behaviour**. *Nat Commun* 2014, **5**:4012.
107. Su L, Jia W, Hou C, Lei Y: **Microbial biosensors: a review**. *Biosens Bioelectron* 2011, **26**:1788-1799.

108. Purnick PEM, Weiss R: **The second wave of synthetic biology: from modules to systems**. *Nat Rev Mol Cell Biol* 2009, **10**:410-422.
109. Ergün A, Lawrence CA, Kohanski MA, Brennan TA, Collins JJ: **A network biology approach to prostate cancer**. *Molecular Systems Biology* 2007, **3**:82-82.
110. Shong J, Collins CH: **Quorum sensing-modulated AND-gate promoters control gene expression in response to a combination of endogenous and exogenous signals**. *ACS Synth Biol* 2014, **3**:238-246.
111. Balázsi G, van Oudenaarden A, Collins JJ: **Cellular Decision-Making and Biological Noise: From Microbes to Mammals**. *Cell* 2011, **144**:910-925.
112. Elman NM, Ben-Yoav H, Sternheim M, Rosen R, Krylov S, Shacham-Diamand Y: **Towards toxicity detection using a lab-on-chip based on the integration of MOEMS and whole-cell sensors**. *Biosens Bioelectron* 2008, **23**:1631-1636.
113. Ben-Yoav H, Elad T, Shlomovits O, Belkin S, Shacham-Diamand Y: **Optical modeling of bioluminescence in whole cell biosensors**. *Biosens Bioelectron* 2009, **24**:1969-1973.
114. Kim E, Gordonov T, Bentley WE, Payne GF: **Amplified and In Situ Detection of Redox-Active Metabolite Using a Bio-based Redox-Capacitor**. *Anal Chem* 2013.
115. Nakano T, Eckford AW, Haraguchi T: *Molecular Communication*. New York: Cambridge University Press; 2013.
116. Song H, Payne S, Tan C, You L: **Programming microbial population dynamics by engineered cell-cell communication**. *Biotechnol J* 2011, **6**:837-849.
117. Pai A, Tanouchi Y, Collins CH, You L: **Engineering multicellular systems by cell-cell communication**. *Curr Opin Biotechnol* 2009, **20**:461-470.
118. Bacchus W, Fussenegger M: **Engineering of synthetic intercellular communication systems**. *Metab Eng* 2013, **16**:33-41.
119. Didovyk A, Kanakov OI, Ivanchenko MV, Hasty J, Huerta R, Tsimring L: **Distributed Classifier Based on Genetically Engineered Bacterial Cell Cultures**. *ACS Synthetic Biology* 2014.
120. Zargar A, Quan DN, Emamian M, Tsao CY, Wu H-C, Virgile CR, Bentley WE: **Rational design of 'controller cells' to manipulate protein and phenotype expression**. *Metabolic Engineering* 2015, **30**:61-68.
121. Rubenstein M, Cornejo A, Nagpal R: **Robotics. Programmable self-assembly in a thousand-robot swarm**. *Science* 2014, **345**:795-799.
122. Hong SH, Hegde M, Kim J, Wang X, Jayaraman A, Wood TK: **Synthetic quorum-sensing circuit to control consortial biofilm formation and**. *Nat Commun* 2012, **3**:613.
123. Balagadde FK, Song H, Ozaki J, Collins CH, Barnet M, Arnold FH, Quake SR, You L: **A synthetic Escherichia coli predator-prey ecosystem**. *Mol Syst Biol* 2008, **4**:187.
124. Nakano T, Suda T, Okaie Y, Moore MJ, Vasilakos AV: **Molecular communication among biological nanomachines: a layered architecture and research issues**. *IEEE Trans Nanobioscience* 2014, **13**:169-197.
125. Rao R, Kesidis G: **Purposeful mobility for relaying and surveillance in mobile ad hoc sensor networks**. *Mobile Computing, IEEE Transactions on* 2004, **3**:225-231.
126. Beyer HM, Naumann S, Weber W, Radziwill G: **Optogenetic control of signaling in mammalian cells**. *Biotechnology Journal* 2014:n/a-n/a.
127. Lee JJ, Jeong KJ, Hashimoto M, Kwon AH, Rwei A, Shankarappa SA, Tsui JH, Kohane DS: **Synthetic ligand-coated magnetic nanoparticles for microfluidic bacterial separation from blood**. *Nano Lett* 2014, **14**:1-5.
128. Miltenyi S, Müller W, Weichel W, Radbruch A: **High gradient magnetic cell separation with MACS**. *Cytometry* 1990, **11**:231-238.

129. Saliba AE, Saias L, Psychari E, Minc N, Simon D, Bidard FC, Mathiot C, Pierga JY, Fraisier V, Salamero J, et al.: **Microfluidic sorting and multimodal typing of cancer cells in self-assembled magnetic arrays**. *Proc Natl Acad Sci U S A* 2010, **107**:14524-14529.
130. Waters CM, Bassler BL: **Quorum sensing: cell-to-cell communication in bacteria**. *Annu Rev Cell Dev Biol* 2005, **21**:319-346.
131. Rutherford ST, Bassler BL: **Bacterial quorum sensing: its role in virulence and possibilities for its control**. *Cold Spring Harb Perspect Med* 2012, **2**.
132. Winzer K, Hardie KR, Williams P: **Bacterial cell-to-cell communication: sorry, can't talk now — gone to lunch!** *Current Opinion in Microbiology* 2002, **5**:216-222.
133. Hegde M, Englert DL, Schrock S, Cohn WB, Vogt C, Wood TK, Manson MD, Jayaraman A: **Chemotaxis to the quorum-sensing signal AI-2 requires the Tsr chemoreceptor and the periplasmic LsrB AI-2-binding protein**. *J Bacteriol* 2011, **193**:768-773.
134. Benz I, Schmidt MA: **Cloning and expression of an adhesin (AIDA-I) involved in diffuse adherence of enteropathogenic Escherichia coli**. *Infect Immun* 1989, **57**:1506-1511.
135. Jarmander J, Gustavsson M, Do TH, Samuelson P, Larsson G: **A dual tag system for facilitated detection of surface expressed proteins in Escherichia coli**. *Microb Cell Fact* 2012, **11**:118.
136. Keefe AD, Wilson DS, Seelig B, Szostak JW: **One-step purification of recombinant proteins using a nanomolar-affinity streptavidin-binding peptide, the SBP-Tag**. *Protein Expr Purif* 2001, **23**:440-446.
137. Bessette PH, Rice JJ, Daugherty PS: **Rapid isolation of high-affinity protein binding peptides using bacterial display**. *Protein Eng Des Sel* 2004, **17**:731-739.
138. Okaie Y, Nakano T, Hara T, Nishio S: **Autonomous mobile bionanosensor networks for target tracking: A two-dimensional model**. *Nano Communication Networks* 2014, **5**:63-71.
139. Tsao CY, Hooshangi S, Wu HC, Valdes JJ, Bentley WE: **Autonomous induction of recombinant proteins by minimally rewiring native quorum**. *Metab Eng* 2010, **12**:291-297.
140. Xavier KB, Bassler BL: **Regulation of uptake and processing of the quorum-sensing autoinducer AI-2 in Escherichia coli**. *J Bacteriol* 2005, **187**:238-248.
141. Marques JC, Oh IK, Ly DC, Lamosa P, Ventura MR, Miller ST, Xavier KB: **LsrF, a coenzyme A-dependent thiolase, catalyzes the terminal step in processing the quorum sensing signal autoinducer-2**. In *Proc Natl Acad Sci U S A*. Edited by; 2014:14235-14240. vol 111.]
142. Marques JC, Lamosa P, Russell C, Ventura R, Maycock C, Semmelhack MF, Miller ST, Xavier KB: **Processing the interspecies quorum-sensing signal autoinducer-2 (AI-2): characterization of phospho-(S)-4,5-dihydroxy-2,3-pentanedione isomerization by LsrG protein**. In *J Biol Chem*. Edited by: 2011 by The American Society for Biochemistry and Molecular Biology, Inc.; 2011:18331-18343. vol 286.]
143. Hooshangi S, Bentley WE: **LsrR quorum sensing "switch" is revealed by a bottom-up approach**. *PLoS Comput Biol* 2011, **7**:e1002172.
144. Sezonov G, Joseleau-Petit D, D'Ari R: **Escherichia coli physiology in Luria-Bertani broth**. *J Bacteriol* 2007, **189**:8746-8749.
145. Yang L, Portugal F, Bentley WE: **Conditioned medium from Listeria innocua stimulates emergence from a resting**. *Biotechnol Prog* 2006, **22**:387-393.
146. Duan F, March JC: **Engineered bacterial communication prevents Vibrio cholerae virulence in an infant mouse model**. *Proceedings of the National Academy of Sciences* 2010, **107**:11260-11264.
147. Kotula JW, Kerns SJ, Shaket LA, Siraj L, Collins JJ, Way JC, Silver PA: **Programmable bacteria detect and record an environmental signal in the mammalian gut**. *Proceedings of the National Academy of Sciences* 2014, **111**:4838-4843.
148. Alam MA: **How to break the "fundamental" limits of nanobiosensing**. In *Colloquium on Sensor Devices*; *University of Maryland, College Park*,Edited by Circuits IEDaS-S: 2013.

149. Jain A, Nair PR, Alam MA: **Flexure-FET biosensor to break the fundamental sensitivity limits of nanobiosensors using nonlinear electromechanical coupling**. *Proceedings of the National Academy of Sciences* 2012.
150. Ebrahimi A, Dak P, Salm E, Dash S, Garimella SV, Bashir R, Alam MA: **Nanotextured superhydrophobic electrodes enable detection of attomolar-scale DNA concentration within a droplet by non-faradaic impedance spectroscopy**. *Lab on a Chip* 2013, **13**:4248-4256.
151. Regev A, Shapiro E: **Cellular abstractions: Cells as computation**. *Nature* 2002, **419**:343-343.
152. Cardelli L: **Abstract Machines of Systems Biology**. In *Transactions on Computational Systems Biology III*. Edited by Priami C, Merelli E, Gonzalez P, Omicini A: Springer Berlin Heidelberg; 2005:145-168. Lecture Notes in Computer Science, vol 3737.]
153. Nakano T, Moore MJ, Fang W, Vasilakos AV, Jianwei S: **Molecular Communication and Networking: Opportunities and Challenges**. *NanoBioscience, IEEE Transactions on* 2012, **11**:135-148.
154. Fernandes R, Luo X, Tsao C-Y, Payne GF, Ghodssi R, Rubloff GW, Bentley WE: **Biological nanofactories facilitate spatially selective capture and manipulation of quorum sensing bacteria in a bioMEMS device**. *Lab on a Chip* 2010, **10**:1128-1134.
155. Gordonov T, Kim E, Cheng Y, Ben-Yoav H, Ghodssi R, Rubloff G, Yin JJ, Payne GF, Bentley WE: **Electronic modulation of biochemical signal generation**. *Nat Nanotechnol* 2014, **9**:605-610.
156. Li Y, Liu Y, Gao T, Zhang B, Song Y, Terrell JL, Barber N, Bentley WE, Takeuchi I, Payne GF, et al.: **Self-Assembly with Orthogonal-Imposed Stimuli To Impart Structure and Confer Magnetic Function To Electrodeposited Hydrogels**. *ACS Applied Materials & Interfaces* 2015, **7**:10587-10598.
157. Gupta A, Terrell JL, Fernandes R, Dowling MB, Payne GF, Raghavan SR, Bentley WE: **Encapsulated fusion protein confers "sense and respond" activity to**. *Biotechnol Bioeng* 2012.
158. Servinsky M, Allen P, Tsao C-Y, Byrd C, Sund C, Quan D, Wu H-C, Terrell JL, Bentley WE: **Directed assembly of a bacterial quorum**. Edited by. ISME Journal: U.S. Army Research Laboratory; 2014.
159. Martini L, Mansy SS: **Cell-like systems with riboswitch controlled gene expression**. *Chemical Communications* 2011, **47**:10734-10736.
160. DeLisa MP, Valdes JJ, Bentley WE: **Mapping stress-induced changes in autoinducer AI-2 production in chemostat-cultivated Escherichia coli K-12**. *Journal of Bacteriology* 2001, **183**:2918-2928.
161. Bassler BL, Wright M, Showalter RE, Silverman MR: **Intercellular signalling in *Vibrio harveyi*: sequence and function of genes regulating expression of luminescence**. *Mol Microbiol* 1993, **9**:773-786.
162. Koop AH, Hartley ME, Bourgeois S: **A low-copy-number vector utilizing beta-galactosidase for the analysis of gene control elements**. *Gene* 1987, **52**:245-256.
163. Wang L, Hashimoto Y, Tsao CY, Valdes JJ, Bentley WE: **Cyclic AMP (cAMP) and cAMP receptor protein influence both synthesis and uptake of extracellular autoinducer 2 in Escherichia coli**. *J Bacteriol* 2005, **187**:2066-2076.
164. Nhan NT, Gonzalez de Valdivia E, Gustavsson M, Hai TN, Larsson G: **Surface display of Salmonella epitopes in Escherichia coli and Staphylococcus carnosus**. *Microb Cell Fact* 2011, **10**:22.
165. Turovskiy Y, Chikindas ML: **Autoinducer-2 bioassay is a qualitative, not quantitative method influenced by glucose**. *J Microbiol Methods* 2006, **66**:497-503.