ACS SyntheticBiology
Research Article
Cite This: ACS Synth. Biol. 2019, 8, 181−190
pubs.acs.org/synthbio

Engineering the S-Layer of *Caulobacter crescentus* as a Foundation for Stable, High-Density, 2D Living Materials

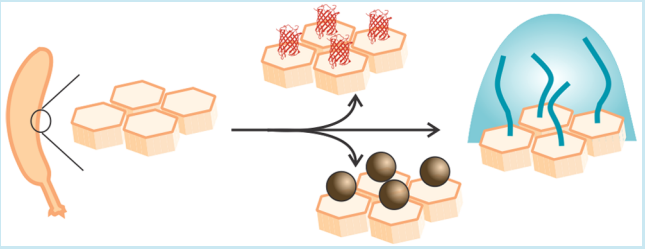
Marimikel Charrier,[†] Dong Li,[†] Victor R. Mann,[†,‡,∞] Lisa Yun,[†,∞] Sneha Jani,[†,∞] Behzad Rad,[†] Bruce E. Cohen,[†] Paul D. Ashby,[†,§] Kathleen R. Ryan,[∥,⊥] and Caroline M. Ajo-Franklin*[,†,#,▽]

[†]Molecular Foundry, [§]Materials Sciences Division, [∥]Environmental Genomics and Systems Biology Division, [#]Molecular Biophysics and Integrated Biosciences Division, and [▽]Synthetic Biology Institute, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

[⊥]Department of Plant & Microbial Biology, and [‡]Department of Chemistry, University of California Berkeley, Berkeley, California 94720, United States

Supporting Information

ABSTRACT: Materials synthesized by organisms, such as bones and wood, combine the ability to self-repair with remarkable mechanical properties. This multifunctionality arises from the presence of living cells within the material and hierarchical assembly of different components across nanometer to micron scales. While creating engineered analogues of these natural materials is of growing interest, our ability to hierarchically order materials using living cells largely relies on engineered 1D protein filaments. Here, we lay the foundation for bottom-up assembly of engineered living material composites in 2D along the cell body using a synthetic biology approach. We engineer the paracrystalline surface-layer (S-layer) of *Caulobacter crescentus* to display SpyTag peptides that form irreversible isopeptide bonds to SpyCatcher-modified proteins, nanocrystals, and biopolymers on the extracellular surface. Using flow cytometry and confocal microscopy, we show that attachment of these materials to the cell surface is uniform, specific, and covalent, and its density can be controlled on the basis of the insertion location within the S-layer protein, RsaA. Moreover, we leverage the irreversible nature of this attachment to demonstrate *via* SDS-PAGE that the engineered S-layer can display a high density of materials, reaching 1 attachment site per 288 nm^2. Finally, we show that ligation of quantum dots to the cell surface does not impair cell viability, and this composite material remains intact over a period of 2 weeks. Taken together, this work provides a platform for self-organization of soft and hard nanomaterials on a cell surface with precise control over 2D density, composition, and stability of the resulting composite, and is a key step toward building hierarchically ordered engineered living materials with emergent properties.



KEYWORDS: *RsaA, engineered living materials, quantum dots, Caulobacter, biomaterial*

Living organisms hierarchically order soft and hard components to create biominerals that have multiple exceptional physical properties.[1] For example, the hierarchical structure of nacre creates its unusual combination of stiffness, toughness, and iridescence. Genetically manipulating living cells to arrange synthesized materials into engineered living materials (ELMs)[2,3] opens a variety of applications in bioelectronics,[4] biosensing,[5] smart materials,[6] and catalysis.[3−7] Many of these approaches use surface display of 1D protein filaments[8−11] or membrane proteins[12,13] to arrange materials, while cell display methods that hierarchically order materials in 2D with controlled spatial positioning and density have yet to be fully developed. This gap limits the structural versatility and degree of control available to rationally engineer ELMs.

Surface-layer (S-layer) proteins offer an attractive platform to scaffold materials in 2D on living cells due to their dense, periodic structures, which form lattices on the outermost surface of many prokaryotes[14] and some eukaryotes.[15] These monomolecular arrays can have hexagonal (p3, p6), oblique (p1, p2),[18,19] or tetragonal (p4)[20] geometries and play critical roles in cell structure,[21,22] virulence,[23] protection,[24] adhesion,[25] and more. Recombinant S-layer proteins can replace the wild-type lattice in native hosts, or can be isolated and recrystallized *in vitro*, on solid supports, or as vesicles.[26] These have been used for a number of applications,[26,27] including bioremediation[28] and therapeutics[29] on cells.

To date, only two S-layer proteins have solved atomic structures, allowing for subnanometer precise positioning of attached materials: SbsB of *Geobacillus stearothermophilus* PV72[18] and RsaA of *Caulobacter crescentus* CB15.[30] Of these two S-layer proteins, there is a well-established toolkit for the genetic modification of *C. crescentus*, as it has been studied extensively for its dimorphic cell cycle.[31] Additionally, *C. crescentus* is a Gram-negative, oligotrophic bacterium that

Received: October 26, 2018
Published: December 21, 2018

DOI: 10.1021/acssynbio.8b00448
ACS Synth. Biol. 2019, 8, 181−190

thrives in low-nutrient conditions, and while a strict aerobe, can survive microaeration.[32] Together, this makes *C. crescentus* particularly suitable as an ELM chassis. RsaA forms a p6 hexameric lattice with a 22 nm unit cell (Figure 1a,b) at an

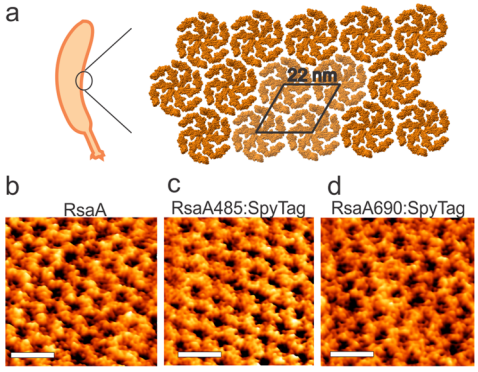


Figure 1. RsaA forms a 2D hexameric lattice on the surface of *C. crescentus*. (a) Structure of the RsaA lattice.[30] (b) High resolution AFM images of the wild-type RsaA lattice (strain MFm111), (c) RsaA485:SpyTag (strain MFm 118), (d) RsaA690:SpyTag (strain MFm 120) on the surface of *C. crescentus* cells. In all three cases, a well-ordered, hexagonal protein lattice is observed. The unit cell length (center-to-center distance between adjacent hexagons) is 22 ± 1 nm, which is the same as reported in literature. Scale bar is 40 nm. See Methods for experimental details of AFM.

estimated density of 45 000 monomers per bacterium[33] and is amenable to peptide insertions.[34] Specific protein domains have been inserted in RsaA to bind lanthanide metal ions[35] or viruses.[29] However, engineered RsaA variants currently lack the ability to assemble a variety of materials, in an irreversible fashion, and with well-characterized density—all key features needed for ELMs.

Here we engineer RsaA as a modular docking point to ligate inorganic, polymeric, or biological materials to the cell surface of *C. crescentus* without disrupting cell viability. This 2D assembly system is specific, stable, and allows for control over the density of attached materials without the use of chemical cues, achieving a maximal coverage of ~25% of all possible sites, the highest density of cell-surface displayed proteins reported to our knowledge. This work forms the foundation for a new generation of hierarchically assembled ELMs.

RESULTS AND DISCUSSION

Design and Construction of *Caulobacter crescentus* S-Layer Variants for Surface Display. To display materials on the surface of *C. crescentus* cells, we designed a genetic module that meets four criteria: (i) a solution-exposed peptide that drives (ii) specific, stable, and stoichiometric attachment (iii) with tunable occupancy and (iv) that does not disrupt RsaA coverage. We hypothesized that varying the location of the binding peptide within RsaA might affect its solution accessibility leading to strains that have a range of occupancy. Therefore, we selected a panel of locations arrayed across the entire RsaA monomer (Figure 2a) to insert the peptide. Smit and colleagues previously identified two sites, at amino acid positions 723 and 944, that allowed for surface display of peptides,[34] so we started with these positions. We then selected six additional sites that are known to be susceptible to proteolytic cleavage, presumably by the previously characterized S-layer Associated Protease (sapA).[36] We hypothesized

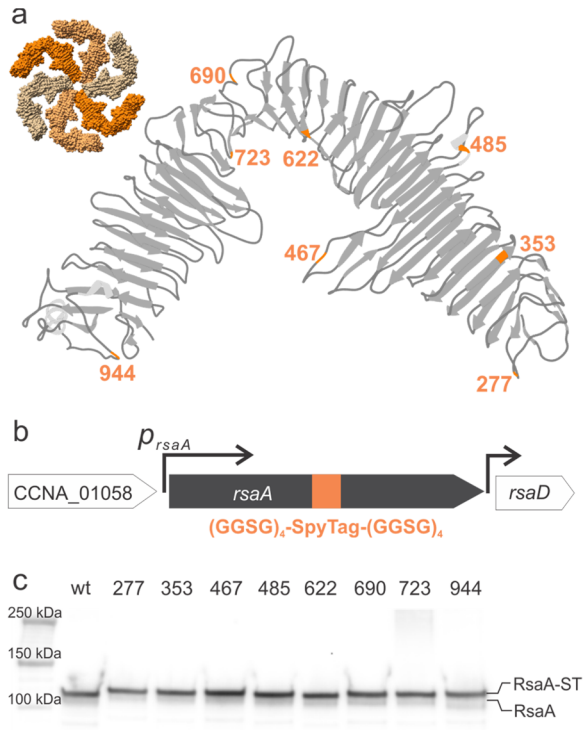


Figure 2. Design and expression of RsaA-SpyTag in *C. crescentus*. (a) Ribbon diagram of the RsaA monomer structure[30] indicating SpyTag insertion sites (orange). Inset shows a space-filling model of the RsaA hexamer. (b) Design of engineered *C. crescentus* strains expressing RsaA-SpyTag. SpyTag flanked by upstream and downstream $(GGSG)_4$ spacers was directly inserted into the genomic copy of *rsaA*. (c) Immunoblot with anti-RsaA antibodies of *C. crescentus* strains whole cell lysate. The band corresponding to RsaA increases in molecular weight from wild-type RsaA (lane 2) to RsaA-SpyTag at the each insertion sites (lanes 3−10).

these additional sites, immediately following amino acid positions 277, 353, 467, 485, 622, and 690, might be accessible in a Δ*sapA* strain which we created (abbreviated as CB15NΔ*sapA*). All subsequent engineering to the eight positions within *rsaA* was done in this background.

To achieve specific, stoichiometric, and irreversible conjugation to RsaA, we employed the split-protein system SpyTag-SpyCatcher,[37,38] which forms an isopeptide bond between the SpyCatcher protein and SpyTag peptide. The RsaA S-layer can accommodate insertion of large peptide sequences,[29,35] which suggested that the 45-mer SpyTag peptide sequence, flanked on each side by a $(GSSG)_4$ flexible linker for accessibility, may be integrated and displayed without disrupting S-layer assembly. This modified lattice should then allow the formation of a covalent isopeptide bond between any material displaying the SpyCatcher partner protein and the SpyTag on the cell surface.

Since expression of RsaA from a p4-based plasmid in an Δ*rsaA* background formed a lattice structure indistinguishable from genomically expressed RsaA (Figure 1b), we initially constructed p4-based plasmids[39] that constitutively express RsaA-SpyTag fusions (Table S2) and transformed them into *C. crescentus* JS4038.[39] Examination of the cell surface of two of these plasmid-bearing strains by AFM confirmed that RsaA-SpyTag was expressed and showed that the RsaA-SpyTag formed a S-layer lattice with the same nanoscale ordering as wild-type RsaA (Figure 1b−d). However, we observed

DOI: 10.1021/acssynbio.8b00448
ACS Synth. Biol. 2019, 8, 181−190

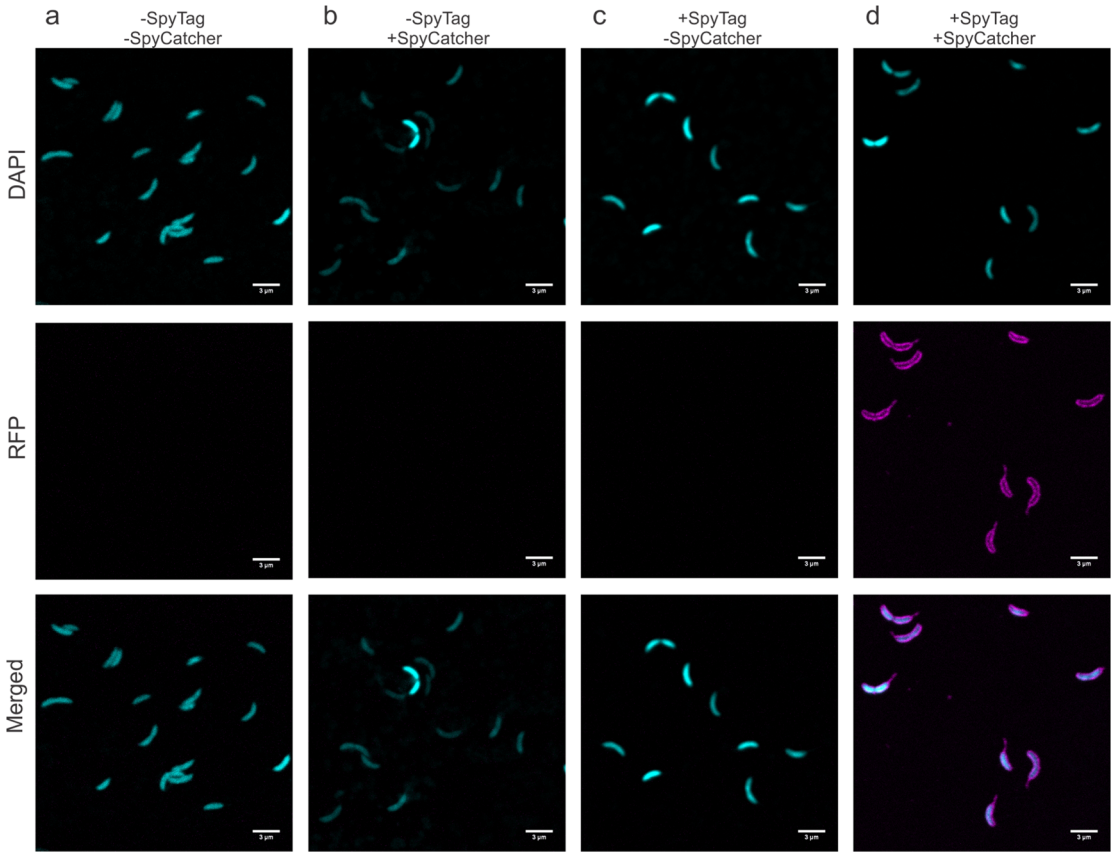


Figure 3. SpyCatcher protein fusions ligate specifically to the surface of *C. crescentus* expressing RsaA-SpyTag. (a–d) Confocal fluorescence images of *C. crescentus* cells visualized in DAPI and RFP channels. Cells expressing wild-type RsaA incubated with (a) mRFP1 or (b) SpyCatcher-mRFP1. Cells expressing RsaA690-SpyTag with (c) mRFP1 or (d) SpyCatcher-mRFP1. Only when the SpyCatcher-mRFP1 probe is introduced to cells displaying SpyTag (d) is RFP fluorescence tightly associated with the cell membrane observed, including the stalk region. Scale bar = 3 μm.

significant growth defects, morphological changes, and unstable RsaA expression in all of the plasmid-bearing strains (Figure S1). For this reason, we integrated SpyTag and its linkers directly into the genomic copy of *rsaA* (Figure 2b) in the CB15NΔ*sapA* background (Table S1). We notate these strains based on the SpyTag insertion site, *e.g.*, *rsaA*690:SpyTag denotes insertion of the SpyTag and $(GSSG)_4$ linkers immediately after amino acid 690. No growth defects or morphological changes are apparent in any of the engineered strains, implying that our genomic insertions do not affect cell viability, and therefore these strains were used for the rest of the study. These observations suggest the more regulated genomic expression of recombinant *rsaA*, a highly transcribed gene, sidesteps growth impairments.

SDS-PAGE analysis of RsaA-SpyTag expression of wild-type (CB15NΔ*sapA*) and engineered cells (*rsaA*:SpyTag variants) shows the expected band for wild-type RsaA at 110 kDa (Figure 2c), in line with the observed migration of RsaA on SDS-PAGE.[40,30] Moreover, all eight engineered proteins have comparable expression levels to wild-type RsaA and show the small increase in molecular weight associated with SpyTag and its linkers (Figure 2c). These observations demonstrate that successful expression of SpyTag within RsaA at a range of different positions does not adversely affect RsaA expression levels.

Engineered S-Layers Specifically Display Proteins Ligated to the Cell Surface. To explore accessibility of the SpyTag peptide on the *C. crescentus* cell surface, we engineered and purified a fusion of SpyCatcher and mRFP1.[41−43] We incubated the wild-type and engineered *rsaA*690:SpyTag strains with the fluorescent SpyCatcher-mRFP1 protein or mRFP1 alone, washed away unbound protein, and visualized mRFP1 attachment to individual cells *via* confocal microscopy (Figure 3). No significant mRFP1 fluorescence is apparent in controls that used *C. crescentus* expressing wild-type RsaA or mRFP1 without SpyCatcher (Figure 3a–c), indicating no significant nonspecific binding of mRFP1 to the cell surface. When SpyTag is displayed on RsaA and SpyCatcher-mRFP1 is present, bright and uniform fluorescence is observed along the morphologically normal, curved cell surface, including the stalk which is covered by the S-layer lattice[17,30,33] (Figure 3d). These observations indicate engineering SpyTag into RsaA enables specific binding of a SpyCatcher fusion protein to the extracellular surface, and furthermore illustrates that engineering SpyTag into the S-layer does not substantially affect the morphology of *C. crescentus*.

Density of Attached Materials Is Controlled by Insertion Location. Having demonstrated specific display of proteins on the surface, we turned to the hypothesis that the solution-accessibility of the eight insertion locations within the RsaA monomer would allow us to vary the density of attached materials. To quantify this relative accessibility, we again incubated the engineered strains with SpyCatcher-mRFP1, washed away unbound protein, and measured the fluorescence intensity per cell with flow cytometry. The engineered strains show a >100-fold increase in fluorescent signal (Figure 4a)

DOI: 10.1021/acssynbio.8b00448
ACS Synth. Biol. 2019, 8, 181−190

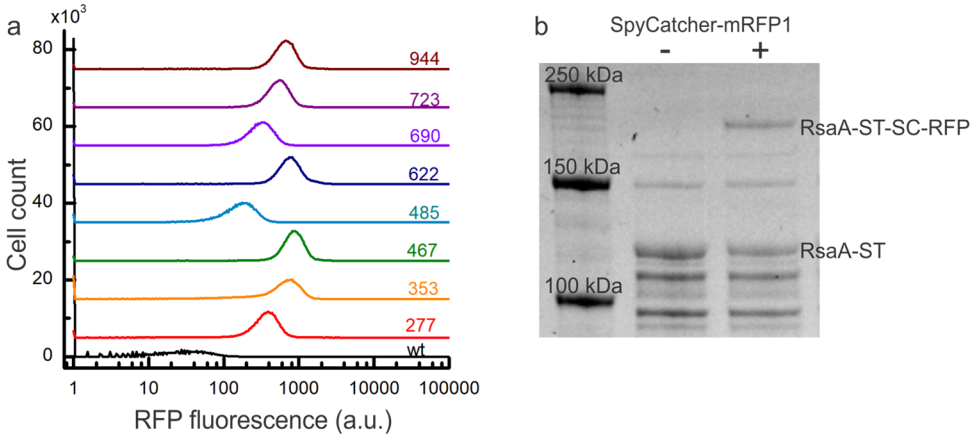


Figure 4. SpyCatcher protein fusions covalently bind to RsaA-SpyTag with variable occupancy according to the SpyTag location. (a) Flow cytometry histograms of RFP fluorescence per cell for strains expressing wild-type RsaA (black) and RsaA-SpyTag (colored lines) incubated with SpyCatcher-mRFP1 for 1 h. Baselines are offset for clarity. All eight strains displaying RsaA-SpyTag show an increase in the intensity of RFP fluorescence over the negative control with their intensity varying based on where SpyTag is inserted within RsaA. (b) SDS-PAGE of whole cell lysates from the *rsaA*467:SpyTag strain incubated for 24 h without (lane 2) and with (lane 3) SpyCatcher-mRFP1 protein. Appearance of a higher molecular weight band only in the reaction containing SpyCatcher-mRFP1 indicates covalent binding to RsaA-SpyTag.

Table 1. Normalized and Absolute Levels of SpyCatcher-mRFP1 Ligation

location of SpyTag insertion	absolute intensity of bound SpyCatcher-mRFP1, I_{loc} (mean ± SEM)	relative SpyCatcher-mRFP1 binding, $I_{loc,\,rel}$ (mean ± SEM)	percentage of RsaA-SpyTag covalently modified (%), P_{loc} (percentage, SEM)
277	$373.9 \pm 3.6 \times 10^{-01}$	$0.43 \pm 5.5 \times 10^{-04}$	9.9 ± 2.4
353	$606.8 \pm 8.5 \times 10^{-01}$	$0.70 \pm 1.1 \times 10^{-03}$	16.0 ± 4.9
467	$871.6 \pm 7.3 \times 10^{-01}$	$1.00 \pm 1.2 \times 10^{-03}$	23.0 ± 2.0
485	$170.3 \pm 2.3 \times 10^{-01}$	$0.20 \pm 3.2 \times 10^{-04}$	4.5 ± 1.3
622	$778.1 \pm 8.6 \times 10^{-01}$	$0.89 \pm 1.2 \times 10^{-03}$	20.5 ± 5.4
690	$316.3 \pm 3.3 \times 10^{-01}$	$0.36 \pm 4.9 \times 10^{-04}$	8.3 ± 2.1
723	$536.0 \pm 4.8 \times 10^{-01}$	$0.61 \pm 7.5 \times 10^{-04}$	14.1 ± 3.3
944	$668.9 \pm 6.1 \times 10^{-01}$	$0.77 \pm 9.6 \times 10^{-04}$	17.6 ± 4.2

over the wild-type control, indicating that all eight positions can ligate significant amounts of SpyCatcher fusion protein. Among the eight engineered strains, there is a ~5-fold variation in the levels of ligation (Table 1), with *rsaA*467:SpyTag and *rsaA*485:SpyTag showing the highest and lowest densities of binding, respectively. These results unveil six new permissive insertion sites within RsaA and show that the amount of protein bound to the cell surface can be controlled by utilizing these different insertion points.

To test that the fusion protein is irreversibly conjugated to RsaA-SpyTag, we incubated strain *rsaA*467:SpyTag, which showed the highest fluorescence by flow cytometry, with SpyCatcher-mRFP1, boiled the sample for 10 min with SDS and 2-mercaptoethanol, and visualized covalent attachment by SDS-PAGE (Figure 4b). The band corresponding to RsaA-SpyTag (Figure 4b) decreases in intensity while the band corresponding to the RsaA-SpyTag-SpyCatcher-mRFP1 assembly appears in as little as 1 h and increases over 24 h. Subsequent immunoblotting of this reaction with anti-RsaA polyclonal antibodies confirms that the assembly band contains RsaA (Figure S2). These observations indicate the binding is covalent.

We leveraged the formation of this covalent bond to quantify the absolute density of SpyCatcher-mRFP1 displayed on the *C. crescentus* cell surface. The density of the RsaA band decreases by 23 ± 5% (n = 6, refer to Methods for the details of this calculation), indicating that nearly a quarter of the *rsaA*467:SpyTag protein is ligated to SpyCatcher-mRFP1 after 24 h. On the basis of the estimate of 45 000 RsaA monomers per cell,[33] this translates to >11 000 copies of SpyCatcher-RFP displayed on the cell surface, an average density of 1.5 SpyCatcher-RFPs per RsaA hexamer, or 1 SpyCatcher-RFP per 288 nm^2. Combining this information with the flow cytometry data, we calculated the percentage of the RsaA lattice that is covalently modified can be controlled over a range from 4 to 23% by varying the engineered location (Table 1, refer to Methods for the details of this calculation). These results provide quantitative information on how to utilize position-dependent insertion of SpyTag in RsaA to tune the density of attached materials and thus substantially improve our ability to rationally engineer ELMs.

Arraying Hard and Soft Materials on the Surface of Engineered Cells. Next we sought to test whether engineered RsaA could assemble soft materials on the surface of *C. crescentus*. We selected elastin-like polypeptide (ELP) as our model soft material because it is well-studied, easily expressed recombinantly, and exhibits interesting temperature-dependent phase behavior.[44] We incubated a SpyCatcher-ELP-mCherry fusion protein[45] with the wild-type and *rsaA*690:SpyTag strains, washed away unbound protein, and imaged individual cells by confocal microscopy. As before, we observe no significant mCherry fluorescence from incubations lacking either SpyTag or SpyCatcher (Figure 5a,b), indicating there is no significant nonspecific binding of ELP-mCherry to the *C. crescentus* surface. When both SpyTag and SpyCatcher are present, we observe significant mCherry fluorescence that uniformly covers the cell surface (Figure 5c). This work

DOI: 10.1021/acssynbio.8b00448
ACS Synth. Biol. 2019, 8, 181−190

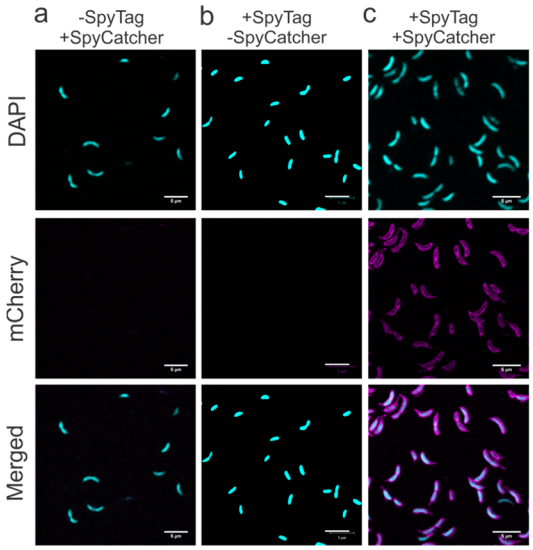


Figure 5. Engineered RsaA assembles biopolymers on the *C. crescentus* cell surface. (a–c) Confocal fluorescence images of *C. crescentus* cells incubated with ELP-mCherry fusion proteins visualized in DAPI and mCherry channels. Cells expressing (a) wild-type RsaA incubated with SpyCatcher-ELP-mCherry and (b) expressing RsaA690:SpyTag incubated with ELP-mCherry. Only the *rsaA*690:SpyTag strain incubated with SpyCatcher-ELP-mCherry (c) shows signal along the cell membrane in the mCherry channel, indicating specific assembly on the cell surface. Scale bar = 5 μm.

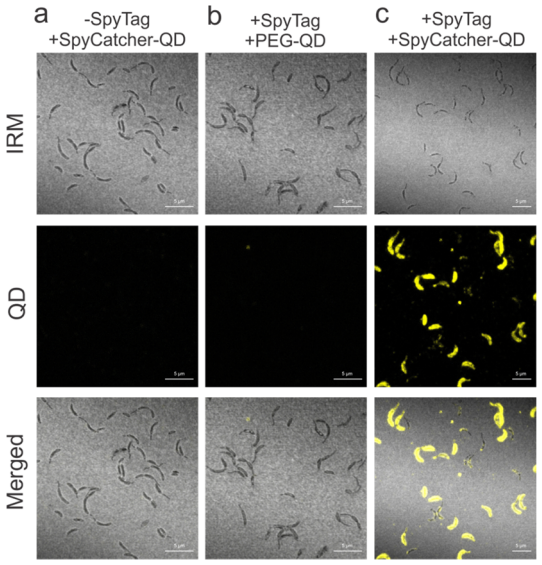


Figure 6. Engineered RsaA assembles inorganic nanocrystals on the *C. crescentus* cell surface. (a–c) Interference reflection microscopy (IRM) and confocal fluorescence images of *C. crescentus* cells incubated with QDs. Cells expressing (a) wild-type RsaA incubated with SpyCatcher-QDs and (b) expressing RsaA690:SpyTag incubated with PEG-QDs. (c) Cells expressing RsaA690:SpyTag incubated with SpyCatcher-QDs show QD fluorescence along the cell surface, indicating specific assembly of SpyCatcher-QDs by the engineered strain. Scale bar = 5 μm.

indicates that the engineered RsaA lattice can assemble polymeric materials to the cell surface (Figure 5).

To explore the diversity of structures that can be created at the cell surface using SpyCatcher-SpyTag ligation, we tested the capacity of engineered bacteria to conjugate CdSe/ZnS semiconductor quantum dots (QDs).[46,47] SpyCatcher-functionalized QDs were generated through attachment of a hetero-bifunctional PEG linker molecule to an amphiphilic polymer encapsulating the QD surface. Subsequent incubation with SpyCatcher-Ser35Cys single cysteine mutant protein yielded QDs with surface-displayed SpyCatcher protein. We incubated PEGylated QDs and SpyCatcher-conjugated QDs (see Supporting Information) with wild-type and *rsaA*690:SpyTag strains, performed a wash, and visualized individual cells *via* confocal microscopy. There is significant QD fluorescence along the cell body in samples containing SpyCatcher-QDs and the engineered strain, while there is no significant fluorescence with the wild-type strain (Figure 6) or the PEGylated QDs. This demonstrates that hard nanomaterials can also be specifically attached to the engineered RsaA lattice.

Nanoparticle Attachment Does Not Affect Cell Viability. Finally, we explored the effect of coating the surface of the *C. crescentus* cells with nanoparticles on their viability, as this is key to creating hybrid living materials that remain metabolically active over time.[48] We incubated control wild-type (CB15NΔ*sapA*) and engineered *rsaA*467:SpyTag cells with or without SpyCatcher-QDs for 2 weeks, sampled the cultures periodically, and enumerated the living cells (CFU/mL). We also imaged the samples using confocal microscopy to determine whether the SpyCatcher-QDs remained stably bound to the engineered S-layer. Under all conditions, the total cell numbers decrease over the two week duration, which is expected since nutrients are not replenished (Figure 7a). More importantly, the number of viable cells in the wild-type culture

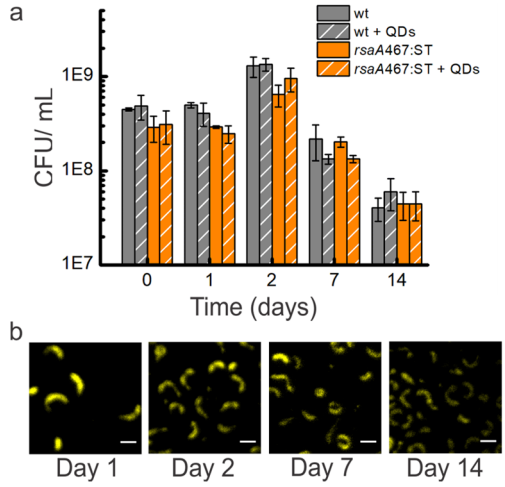


Figure 7. Engineered *C. crescentus* with ligated SpyCatcher-QDs remain viable over 2 weeks. (a) Viability of CB15NΔ*sapA* (wild-type) and CB15NΔ*sapA rsaA*467:SpyTag strains incubated without or with SpyCatcher-QDs (+ QD) was assessed by quantifying colony forming units/mL (CFU/mL) as described in the Methods section. Data shown represent mean ± standard deviation of three replicates per condition. The CFU/mL of cells with SpyCatcher-QDs is very similar to that of cells grown without SpyCatcher-QDs. (b) Confocal images of *rsaA*467:Spytag + SpyCatcherQD show QD fluorescence over the two week duration indicating sustained attachment of SpyCatcher-QDs to the engineered strain. Scale bar = 3 μm.

without QDs is not significantly different from the wild-type with unbound SpyCatcher-QDs, the *rsaA*467:SpyTag culture without QDs, or the *rsaA*467:SpyTag culture with SpyCatcher-QDs (Figure 7a). These results indicate that neither unbound SpyCatcher-QDs in the wild-type culture nor surface-bound SpyCatcher-QDs on the engineered cells affect viability, and

DOI: 10.1021/acssynbio.8b00448
ACS Synth. Biol. 2019, 8, 181–190

there is no notable difference in viability between the wild-type and engineered cells. One possible interpretation of this cell viability is that the S-layer acts as an effective barrier, preventing disruption of the outer cell membrane, fulfilling one of its key evolutionary roles.[49−51] Imaging reveals that SpyCatcher-QDs remain attached to the cell surface over 2 weeks (Figure 7B) and nonspecific binding of SpyCatcher-QDs on the surface of wild-type cells is not observed (Figure S5), once again highlighting the specificity and stability of the SpyTag-SpyCatcher system on S-layers. We do note that QD emission decreases over the course of the experiment, which may be due to nonspecific cleavage of bonds between RsaA and the QD, turnover of the RsaA protein, or slow QD quenching in biological media. SpyCatcher-QDs incubated alone in M2G buffer show a *ca.* 30% decrease in emission over 14 days (Figure S6), which suggests QD quenching is the likely cause of observed emission decrease in QD-RsaA conjugates. Nonetheless, these results demonstrate that engineered RsaA can be used to generate stable living materials that require cells to remain viable for extended periods of time.

Advancement of RsaA S-Layer as a Platform for Controlled Material Assembly. In summary, we show that the S-layer of *C. crescentus*, RsaA, is a versatile platform for cell surface attachment of proteins, biopolymers, and inorganic materials when combined with the Spy conjugation system. We demonstrate that eight sites are available for peptide insertion within RsaA and that the insertion location tunes the attachment density. Ligation to the RsaA-SpyTag lattice is highly specific and covalent, with the absolute level of density of RsaA-displayed proteins reaching ∼25% of the total RsaA, or 1 site per 288 nm^2, which is the highest density of cell-surface displayed proteins reported to our knowledge. Moreover, we show that QD-*C. crescentus* composites assembled *via* RsaA-SpyTag form engineered living materials that persist for at least 2 weeks. In the following, we discuss possible reasons for the site-dependent variation in attachment density, specific applications for cell-display using the RsaA platform, and the broader opportunities it opens in the area of ELMs.

We observed that the relative ligation efficiency varies ∼5-fold across the eight permissive sites, with *rsaA*467:SpyTag affording the densest array of SpyCatcher-mRFP1. This variance is not due to protein expression levels, which do not vary significantly between strains (Figure 2c), and is unlikely to be caused by disruption to the S-layer lattice as our findings indicate that SpyTag insertions to not alter the structure on the nanometer scale (Figure 1b−d) but may be due to solvent-accessibility within the RsaA, steric clashes between sites on nearby RsaA monomers, or a combination of these factors. The most efficient binding site, RsaA467:SpyTag, is in an unstructured loop in a gap in the hexamer (Figure 2a), potentially giving more freedom for the SpyTag peptide to access a SpyCatcher-fusion. Since RsaA485:SpyTag and RsaA690:SpyTag are in an α helix and a calcium-binding pocket, respectively (Figure 2a), these insertions could be causing local disruption in structure, leading to the lower occupancy we observe (Table 1). Additionally, position 277 is located near the pore of the hexamer, resulting in the five neighboring positions being between 1.4 and 2.8 nm away. Since the entire engineered linkage to mRFP1, *i.e.*, $(GSSG)_4$-SpyTag-SpyCatcher-mRFP1, is roughly 2.9 by 2.5 by 15 nm in dimensions, it is likely that some of the neighboring sites are sterically inaccessible once a single mRFP1 is bound. Further investigation will be required to untangle these possibilities.

Engineered *C. crescentus* Opens New Possibilities for Hierarchical Assembly of Hybrid Living Materials. The engineered S-layer system described here offers immediate opportunities for engineering enzyme cascades on cells and encapsulation in hydrogels. By eliminating the need for direct fusion of enzymes to the S-layer, we avoid potential enzyme activity inhibition caused by expressing the protein in tandem with the S-layer monomer.[52] In addition, the varied ligation density and SpyTag spatial positioning engineered in our strains provides flexibility to attach enzymes in the most ideal pattern. As another potential application, bacterial cells are frequently encapsulated in hydrogels to enhance their stability as probiotics,[53] as adjuvants to plant growth in agriculture,[54] or as biostimulants in wastewater treatment.[55] Typically no specific adherence mechanism is engineered between bacterial cells and the hydrogel, and many factors can affect gel stiffness,[56] including number or type of cells and media content. By using direct attachments between the S-layer and hydrogel polymers, we may achieve more stability and unique mechanical properties due to the sheer number of covalent cross-links the between the engineered S-layer and the hydrogel matrix.

Our work more broadly introduces several foundational aspects useful for engineering ELMs. First, our results (Figures 3, 5, 6) suggest any material on which SpyCatcher can be conjugated can be self-assembled on the modified 2D S-layer lattice, thus avoiding the labor-intensive reengineering of RsaA with peptides designed for specific targets. This makes our strain a versatile starting point for building an array of ELMs. Second, while ELMs with impressive functionality have been assembled *via* 1D curli fiber proteins and the type III secretion apparatus,[57] the 2D structure of the S-layer lattice yields another dimension of spatial control. Because hierarchical ordering underlies the exceptional physical properties of many natural biocomposites, the ability to regulate spacing of different components in multiple dimensions is key to rationally designing predictable ELMs. Third, we can attach materials densely to the cell surface; here we demonstrate ligation of ∼11 000 copies of a protein to the *C. crescentus* cell surface, or 1 attached protein per 288 nm^2. This is the highest density of surface arrayed proteins reported on a bacterium to our knowledge. Being able to access high densities is important because it ensures well-ordered structures while the ability to tune density may result in control over material properties. Lastly, the combined robustness of the covalent SpyCatcher-SpyTag system, the RsaA S-layer, and *C. crescentus* enables long-term persistence of the assembled structure and cell viability in an ELM even under low aeration and nutrient conditions. We envision this robustness will enable ELMs that can function in nutrient-poor environments with minimal intervention. Thus, the RsaA platform described here offers a modular, stable platform for assembling materials densely in 2D that opens new possibilities for constructing ELMs.

CONCLUSIONS

In closing, hierarchically ordered hybrid materials could allow for the rational design of materials with the emergent properties seen in natural materials. A bottom-up approach toward these engineered living materials is controlled attachment of materials to the cell surface in 2D, which we achieved by engineering the *C. crescentus* S-layer with the Spy

DOI: 10.1021/acssynbio.8b00448
ACS Synth. Biol. 2019, 8, 181−190

conjugation system for specific attachment of hard, soft, and biological materials at controllable densities. This modular base could lead to higher ordered materials that combine the functions of inorganic materials with the self-assembly and self-healing properties of living cells for applications that span medicine, infrastructure, and devices.

METHODS

Strains. All strains used in this study are listed in Table S1. *C. crescentus* strains were grown in PYE media (0.2% peptone, 0.1% yeast extract, 1 mM $MgSO_4$, 0.5 mM $CaCl_2$) at 30 °C with aeration. *E. coli* strains were grown in LB media (1% tryptone, 0.5% yeast extract, 1% NaCl) at 37 °C with aeration. When required, antibiotics were included at the following concentrations: For *E. coli*, 50 μg/mL ampicillin, 20 μg/mL chloramphenicol, 30 μg/mL kanamycin. For *C. crescentus*, 10 μg/mL (liquid) or 50 mg/mL (plate) ampicillin, 2 μg/mL (liquid) or 1 μg/mL (plate) chloramphenicol, 5 μg/mL (liquid) or 25 μg/mL (plate) kanamycin. Diaminopimelic acid (DAP) was supplemented at 300 μM and sucrose at 3% w/v for conjugation and recombination methods, respectively. All chemicals were purchased from Sigma-Aldrich or VWR.

Plasmid Construction. A list of all strains, plasmids, and primers used in this study is available in Tables S1−S3. Details on construction of p4B expression plasmids, pNPTS138 integration plasmids, and protein purification plasmids can be found in Supporting Information. Plasmids were introduced to *E. coli* using standard transformation techniques with chemically competent or electrocompetent cells, and to *C. crescentus* using conjugation *via E. coli* strain WM3064.

Genome Engineering of *C. crescentus*. The $(GGSG)_4$-SpyTag-$(GGSG)_4$ sequence was integrated into the genomic copy of rsaA using a 2-step recombination technique and sucrose counterselection. The pNPTS-rsaA(SpyTag) integration plasmids were conjugated into *C. crescentus CB15NΔsap* and plated on PYE with kanamycin to select for integration of the plasmid. Successful integrants were incubated in liquid media overnight and plated on PYE supplemented with 3% (w/v) sucrose to select for excision of the plasmid and *sacB* gene, leaving the SpyTag sequence behind. Colonies were then spotted on PYE with kanamycin plates to confirm loss of plasmid-borne kanamycin gene. Integration of the SpyTag sequence and removal of the *sacB* gene was confirmed by colony PCR with OneTaq Hot Start Quick-Load 2× Master Mix with GC buffer (New England BioLabs) using a Touchdown thermocycling protocol with an annealing temperature ranging from 72°-62 °C, decreasing 1° per cycle.

Successful RsaA-SpyTag protein expression was confirmed by band shift in whole cell lysate in Laemmli buffer and 0.05% 2-mercaptoethanol on a BioRad Criterion Stain-free 4−20% SDS-PAGE. The gel was UV-activated for 5 min before imaging on a ProteinSimple FluorChem E system. As RsaA was migrating higher than expected, Western blot was performed for confirmation. A Bio-Rad Trans-Blot Turbo system with nitrocellulose membrane was used to transfer protein from the SDS-PAGE gel and the membrane incubated in Thermo-Fisher SuperBlock buffer for 1 h. The protein of interest was first labeled during a 30 min incubation with Rabbit-C Terminal Anti-RsaA polyclonal antibody[58] (Courtesy of the Smit lab; 1:5000 in TBST, Tris-Buffered Saline with 0.1−0.05% Tween-20), followed by another 30 min incubation with Goat-Anti Rabbit-HRP (Sigma-Aldrich; 1:5000 in TBST). BioRad Precision Plus Protein Standards (Bio-Rad) were labeled with Precision Protein StrepTactin-HRP conjugate antibodies (Bio-Rad; 1:5000 in TBST). HRP fluorescence was activated with Thermo-Fisher SuperSignal West Pico Chemiluminescent Substrate and imaged in chemiluminescent mode. TBST washes were performed between each incubation step. The relative molecular weight of bands quantified against the BioRad Precision Plus Protein Standards using ProteinSimple's AlphaView software.

Monitoring Ligation of SpyCatcher-Fusions to *C. crescentus*. For flow cytometry experiments, cells were grown at 25 °C to mid log phase and cells containing the pBXMCS-2-RFP plasmid were induced for 1−2 h with 0.03% xylose to serve as a positive control. A population of ∼10^8 cells (determined by optical density measurement where OD_{600} of 0.05 contains 10^8 cells) were harvested by centrifugation at 8000 RCF for 5−10 min and resuspended in PBS + 0.5 mM $CaCl_2$. Using the cell density as determined by OD_{600} and assuming 4.5 × 10^4 RsaA monomers/cell,[33] we added SpyCatcher-mRPF1 to a final molar ratio of 1:20 RsaA protein to SpyCatcher-mRFP1. The reaction was then incubated for 1 h at room temperature with rotation. All samples were protected from light with aluminum foil during the procedure and washed twice with 1 mL of Phosphate-Buffered Saline (PBS) + 0.5 mM $CaCl_2$ buffer prior to imaging to remove any unbound protein. Cells were diluted to 10^6 cells/mL and analyzed on a BD LSR Fortessa. Data on forward scatter (area and height), side scatter, and PE Texas Red (561 mm laser, 600 LP 610/20 filter) was collected. A total of 150 000 events for each strain was measured over three experiments.

For each strain, the total population was gated using scatter measurements to remove events corresponding to aggregates and debris. All events from the resulting main population were used to create histograms of the fluorescence intensity of bound SpyCatcher-mRPF1 for each strain expressing RsaA (wild-type control) or RsaA-SpyTag (Figure 4A). These fluorescence intensity values were also used to calculate the absolute intensity of bound of SpyCatcher-mRFP1 for RsaA-SpyTag insertion location (I_{loc}) shown in Table 1. The relative SpyCatcher-mRFP1 binding ($I_{loc,rel}$) was calculated by normalizing (I_{loc}) by the absolute intensity at location 467:

$$I_{loc,rel} = \frac{I_{loc}}{I_{467}} \quad (1)$$

For confocal microscopy, cells were grown to mid log phase and ∼10^8 cells (again determined by OD_{600} measurement) harvested by centrifugation at 8000 RCF for 5−10 min. They were then resuspended in PBS + 0.5 mM $CaCl_2$ and, as in the flow cytometry experiments, a 1:20 ratio of RsaA protein to fluorescent probe, *i.e.*, mRFP1, SpyCatcher-mRFP1, SpyCatcher-ELP-mCherry, or ELP-mCherry was added. The reaction was then incubated for 1 h at room temperature with rotation for the mRFP1 probes and 24 h at 4 °C for the ELP-mCherry probes. 2 × 10^7 cells were incubated with 100 nM QDs for 24 h at 4 °C with rotation. All samples were protected from light with aluminum foil during the procedure and washed twice with 1 mL of buffer prior to imaging to remove any unbound protein. After the wash, cells with fluorescent probes were stained with 1 μM of DAPI (4′,6-diamidino-2-phenylindole). All samples were spotted onto agarose pads (1.5% w/v agarose in distilled water) and mounted between glass slides and glass coverslips. Immersol 518F immersion oil with a refractive index of 1.518 was placed between the sample and the 100× oil

DOI: 10.1021/acssynbio.8b00448
ACS Synth. Biol. 2019, 8, 181−190

immersion objective (Plan-Apochromat, 1.40 NA) prior to imaging. Fluorescence and IRM images were collected using a Zeiss LSM 710 confocal microscope (Carl Zeiss Micro Imaging, Thornwood, NY) with the Zen Black software. For fluorescent imaging, a 561 nm laser was used for RFP/mCherry excitation and 405 nm for DAPI. For IRM, a 514 nm laser was reflected into the sample using a mBST80/R20 plate, and then the reflected light was collected and imaged onto the detector. Images were false colored and brightness/contrast adjusted using ImageJ.[59]

To quantify the binding of SpyCatcher-mRFP1 to RsaA-SpyTag, the same procedure was used as above except with a 1:2 ratio of RsaA to RFP was used and incubation for 24 h at 4 °C with rotation. The reaction was visualized on a BioRad Criterion Stain-free 7.5% SDS-PAGE in Laemmli buffer with 0.05% 2-mercaptoethanol and the molecular weight of bands quantified against BioRad Precision Plus Protein Standards using Protein Simple's AlphaView software. The measurements were made in triplicate on two separate occasions, and all six results were averaged for the final percentage reported. For each experiment, the density of bands was measured using ImageJ.[59] Background subtraction was applied to the entire image, and the background-subtracted integrated density within an equal area was determined for each RsaA-SpyTag protein band. The integrated density of the bands from triplicate reactions lacking SpyCatcher were averaged to give I_{unreact}. To calculate the percentage of RsaA_{467}:SpyTag ligated to SpyCatcher-mRFP1 for each experiment, we calculated the difference in density between each RsaA-SpyTag band from reactions with SpyCatcher-mRFP1 ($I_{\text{react},1}$, $I_{\text{react},2}$, $I_{\text{react},3}$) relative to the unreacted control (I_{unreact}) and normalized this value by the unreacted control (I_{unreact}):

$$P_{467} = \frac{1}{3}\sum_{i=1}^{3}\frac{I_{\text{unreact}} - I_{\text{react},i}}{I_{\text{unreact}}} \quad (2)$$

The reported value (P_{467}) is an average of the two experiments. We then used the absolute percentage of ligation at location 467 (P_{467}) and the relative binding of SpyCatcher-mRFP1 at each location to calculate the percentage of ligation for all the insertion positions.

$$P_{\text{loc}} = I_{\text{loc,rel}} \times P_{467} \quad (3)$$

The values of P_{loc} are shown in Table 1.

Cell Viability Assay. Cell viability in the presence of SpyCatcher-QDs was determined using the viable plate count method. Approximately 4×10^8 mid log phase cells (day 0) (cell number determined by OD_{600} measurement) were first incubated with 100 nM SpyCatcher-QDs in M2G buffer (1× M2 salts without NH_4Cl to prevent extensive cell growth, 1 mM MgSO_4, 0.5 mM CaCl_2, 2% glucose) for 24 h at 4 °C with rotation to allow QD ligation to the cell surface. Post binding (day 1), the cultures were transferred to a 25 °C humidified incubator and left stationary for 2 weeks. Cultures were sampled at different time points (days 2, 7, and 14), serially diluted, and titered on PYE agar plates (0.2% peptone, 0.1% yeast extract, 1 mM MgSO_4, 0.5 mM CaCl_2, 1.5% agar), which were incubated at 30 °C for 2 days. Colonies on the plates were counted and cell viability was quantified by enumeration of Colony Forming Units/mL as follows:

$$\text{CFU/mL} = \text{number of colonies/dilution} \times \text{volume plated (mL)}$$

At the specified time points, 30 μL of culture was removed from the tube and centrifuged at 16 000*g* for 1 min. The supernatant was discarded and the cells were resuspended in 3 μL M2G buffer. A 1.5% agarose pad was prepared on a 25 × 75 mm glass slide and 0.6 μL of the resuspended culture was placed on it. An 18 × 18 mm coverslip was then placed on the pad, and the trapped cells were imaged and processed using as outlined above.

***In Situ* Atomic Force Microscope (AFM) Imaging.** Mid log cultures of *C. crescentus* JS4038 carrying p4B-rsaA_{600}, p4B-rsaA_{600}467: $(\text{GGSG})_4$-spytag-$(\text{GGSG})_4$, or p4B-rsaA_{600}690: $(\text{GGSG})_4$-spytag-$(\text{GGSG})_4$, were harvested at 8000 rpm for 5 min and the pellet resuspended in PBS + 5 mM CaCl_2 buffer. The JS4038 strain is defective in capsular polysaccharide synthesis. This is necessary for AFM imaging as the capsular polysaccharide layer obscures the S-layer lattice. The cells were washed three times to remove any debris. 100 μL of the washed cell culture was applied to a poly-L-lysine coated glass coverslip (12 mm cover glasses, BioCoat from VWR), which was premounted onto a metal puck. The sample was incubated at room temperature for 1 h to allow sufficient cell attachment, and then 1 mL of PBS + 5 mM CaCl_2 buffer was used to wash away unbound cells from the glass surface. 50 μL of PBS + 5 mM CaCl_2 buffer was added to the resulting glass surface, and the sample was transferred onto the sample stage for imaging.

In situ AFM imaging was performed on a Bruker Multimode AFM using PeakForce Tapping mode in liquid. An Olympus Biolever-mini cantilever (BL-AC40TS) was used for high resolution imaging. The following set of parameters was normally employed to ensure the best image quality: 0.2 to 0.5 Hz scanning rate, 512 × 512 scanning lines, 15 nm peak force amplitude, and 50 to 100 pN peak force set point.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.8b00448.

Further methods on plasmid construction, protein purification, and QD synthesis (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Tel: 510-486-4299. E-mail: cajo-franklin@lbl.gov.

ORCID

Marimikel Charrier: 0000-0003-4264-5985
Dong Li: 0000-0002-5282-6918
Victor R. Mann: 0000-0002-0907-1978
Lisa Yun: 0000-0001-7940-4902
Sneha Jani: 0000-0002-4218-5336
Behzad Rad: 0000-0001-6531-9168
Bruce E. Cohen: 0000-0003-3655-3638
Paul D. Ashby: 0000-0003-4195-310X
Kathleen R. Ryan: 0000-0003-3657-8069
Caroline M. Ajo-Franklin: 0000-0001-8909-6712

Author Contributions

∞These authors contributed equally.

DOI: 10.1021/acssynbio.8b00448
ACS Synth. Biol. 2019, 8, 181−190

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are indebted to Prof. John Smit and Dr. John Nomellini for helpful conversations and starting materials. We also thank Dr. Francesca Manea for her contributions to the S-layer research at the Molecular Foundry and Dr. Andrew Hagen for the gift of the pARH356 plasmid. This work was supported by the Defense Advanced Research Projects Agency (Engineered Living Materials Program, C.M.A-F.) and National Institutes of Health award R01NS096317 (B.E.C.) Work at the Molecular Foundry was supported by the Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

■ ABBREVIATIONS

DAPI, 4′,6-diamidino-2-phenylindole; ELP, elastin-like polypeptide; ELM, engineered living material; QD, quantum dot; RFP, red fluorescent protein; S-layer, surface-layer; sapA, S-layer associated protease; AFM, atomic force microscopy; IRM, interference reflection microscopy; ST, SpyTag; SC, SpyCatcher.

■ REFERENCES

(1) Mann, S. (2001) *Biomineralization: Principles and Concepts in Bioinorganic Materials Chemistry*, Oxford University Press.
(2) Chen, A. Y., Zhong, C., and Lu, T. K. (2015) Engineering Living Functional Materials. *ACS Synth. Biol. 4* (1), 8−11.
(3) Nguyen, P. Q., Courchesne, N.-M. D., Duraj-Thatte, A., Praveschotinunt, P., and Joshi, N. S. (2018) Engineered Living Materials: Prospects and Challenges for Using Biological Systems to Direct the Assembly of Smart Materials. *Adv. Mater. 30* (19), e1704847.
(4) Chen, A. Y., Deng, Z., Billings, A. N., Seker, U. O. S., Lu, M. Y., Citorik, R. J., Zakeri, B., and Lu, T. K. (2014) Synthesis and Patterning of Tunable Multiscale Materials with Engineered Cells. *Nat. Mater. 13*, 515.
(5) Zhou, A. Y., Baruch, M., Ajo-Franklin, C. M., and Maharbiz, M. M. (2017) A Portable Bioelectronic Sensing System (BESSY) for Environmental Deployment Incorporating Differential Microbial Sensing in Miniaturized Reactors. *PLoS One 12* (9), e0184994.
(6) Gerber, L. C., Koehler, F. M., Grass, R. N., and Stark, W. J. (2012) Incorporating Microorganisms into Polymer Layers Provides Bioinspired Functional Living Materials. *Proc. Natl. Acad. Sci. U. S. A. 109* (1), 90−94.
(7) Lee, S. Y., Choi, J. H., and Xu, Z. (2003) Microbial Cell-Surface Display. *Trends Biotechnol. 21* (1), 45−52.
(8) Nussbaumer, M. G., Nguyen, P. Q., Tay, P. K. R., Naydich, A., Hysi, E., Botyanszki, Z., and Joshi, N. S. (2017) Bootstrapped Biocatalysis: Biofilm-Derived Materials as Reversibly Functionalizable Multienzyme Surfaces. *ChemCatChem 9* (23), 4328−4333.
(9) Nguyen, P. Q., Botyanszki, Z., Tay, P. K. R., and Joshi, N. S. (2014) Programmable Biofilm-Based Materials from Engineered Curli Nanofibres. *Nat. Commun. 5*, 4945.
(10) Seker, U. O. S., Chen, A. Y., Citorik, R. J., and Lu, T. K. (2017) Synthetic Biogenesis of Bacterial Amyloid Nanomaterials with Tunable Inorganic−Organic Interfaces and Electrical Conductivity. *ACS Synth. Biol. 6* (2), 266−275.
(11) Zhong, C., Gurry, T., Cheng, A. A., Downey, J., Deng, Z., Stultz, C. M., and Lu, T. K. (2014) Strong Underwater Adhesives Made by Self-Assembling Multi-Protein Nanofibres. *Nat. Nanotechnol. 9*, 858.
(12) Nicolay, T., Vanderleyden, J., and Spaepen, S. (2015) Autotransporter-Based Cell Surface Display in Gram-Negative Bacteria. *Crit. Rev. Microbiol. 41* (1), 109−123.
(13) Daugherty, P. S. (2007) Protein Engineering with Bacterial Display. *Curr. Opin. Struct. Biol. 17* (4), 474−480.
(14) Messner, P., Schäffer, C., Egelseer, E.-M., and Sleytr, U. B. (2010) Occurrence, Structure, Chemistry, Genetics, Morphogenesis, and Functions of S-Layers. In *Prokaryotic Cell Wall Compounds: Structure and Biochemistry* (König, H., Claus, H., and Varma, A., Eds.) pp 53−109, Springer, Berlin, Heidelberg.
(15) Roberts, K., Grief, C., Hills, G. J., and Shaw, P. J. (1985) Cell Wall Glycoproteins: Structure and Function. *J. Cell Sci. 2*, 105−127.
(16) Lembcke, G., Baumeister, W., Beckmann, E., and Zemlin, F. (1993) Cryo-Electron Microscopy of the Surface Protein of Sulfolobus Shibatae. *Ultramicroscopy 49* (1), 397−406.
(17) Smit, J., Engelhardt, H., Volker, S., Smith, S. H., and Baumeister, W. (1992) The S-Layer of Caulobacter Crescentus: Three-Dimensional Image Reconstruction and Structure Analysis by Electron Microscopy. *J. Bacteriol. 174* (20), 6527−6538.
(18) Baranova, E., Fronzes, R., Garcia-Pino, A., Van Gerven, N., Papapostolou, D., Péhau-Arnaudet, G., Pardon, E., Steyaert, J., Howorka, S., and Remaut, H. (2012) SbsB Structure and Lattice Reconstruction Unveil Ca2+ Triggered S-Layer Assembly. *Nature 487* (7405), 119−122.
(19) Messner, P., Pum, D., and Sleytr, U. B. (1986) Characterization of the Ultrastructure and the Self-Assembly of the Surface Layer of Bacillus Stearothermophilus Strain NRS 2004/3a. *J. Ultrastruct. Mol. Struct. Res. 97* (1−3), 73−88.
(20) Norville, J. E., Kelly, D. F., Knight, T. F., Jr, Belcher, A. M., and Walz, T. (2007) 7A Projection Map of the S-Layer Protein sbpA Obtained with Trehalose-Embedded Monolayer Crystals. *J. Struct. Biol. 160* (3), 313−323.
(21) Wildhaber, I., and Baumeister, W. (1987) The Cell Envelope of Thermoproteus Tenax: Three-Dimensional Structure of the Surface Layer and Its Role in Shape Maintenance. *EMBO J. 6* (5), 1475−1480.
(22) Pum, D., Messner, P., and Sleytr, U. B. (1991) Role of the S Layer in Morphogenesis and Cell Division of the Archaebacterium Methanocorpusculum Sinense. *J. Bacteriol. 173* (21), 6865−6873.
(23) Kern, J., and Schneewind, O. (2010) BslA, the S-Layer Adhesin of B. Anthracis, Is a Virulence Factor for Anthrax Pathogenesis. *Mol. Microbiol. 75* (2), 324−332.
(24) Koval, S. F., and Hynes, S. H. (1991) Effect of Paracrystalline Protein Surface Layers on Predation by Bdellovibrio Bacteriovorus. *J. Bacteriol. 173* (7), 2244−2249.
(25) Åvall-Jääskeläinen, S., Lindholm, A., and Palva, A. (2003) Surface Display of the Receptor-Binding Region of the Lactobacillus Brevis S-Layer Protein in Lactococcus Lactis Provides Nonadhesive Lactococci with the Ability to Adhere to Intestinal Epithelial Cells. *Appl. Environ. Microbiol. 69* (4), 2230−2236.
(26) Sleytr, U. B., Schuster, B., Egelseer, E.-M., and Pum, D. (2014) S-Layers: Principles and Applications. *FEMS Microbiol. Rev. 38* (5), 823−864.
(27) Ilk, N., Egelseer, E. M., and Sleytr, U. B. (2011) S-Layer Fusion Proteins—construction Principles and Applications. *Curr. Opin. Biotechnol. 22* (6), 824−831.
(28) Patel, J., Zhang, Q., McKay, R. M. L., Vincent, R., and Xu, Z. (2010) Genetic Engineering of Caulobacter Crescentus for Removal of Cadmium from Water. *Appl. Biochem. Biotechnol. 160* (1), 232−243.
(29) Farr, C., Nomellini, J. F., Ailon, E., Shanina, I., Sangsari, S., Cavacini, L. A., Smit, J., and Horwitz, M. S. (2013) Development of an HIV-1 Microbicide Based on Caulobacter Crescentus: Blocking Infection by High-Density Display of Virus Entry Inhibitors. *PLoS One 8* (6), e65965.
(30) Bharat, T. A. M., Kureisaite-Ciziene, D., Hardy, G. G., Yu, E. W., Devant, J. M., Hagen, W. J. H., Brun, Y. V., Briggs, J. A. G., and Löwe, J. (2017) Structure of the Hexagonal Surface Layer on Caulobacter Crescentus Cells. *Nat. Microbiol 2*, 17059.
(31) Laub, M. T., Shapiro, L., and McAdams, H. H. (2007) Systems Biology of Caulobacter. *Annu. Rev. Genet. 41*, 429−441.

DOI: 10.1021/acssynbio.8b00448
ACS Synth. Biol. 2019, 8, 181−190

(32) Crosson, S., McGrath, P. T., Stephens, C., McAdams, H. H., and Shapiro, L. (2005) Conserved Modular Design of an Oxygen Sensory/signaling Network with Species-Specific Output. *Proc. Natl. Acad. Sci. U. S. A. 102* (22), 8018−8023.
(33) Amat, F., Comolli, L. R., Nomellini, J. F., Moussavi, F., Downing, K. H., Smit, J., and Horowitz, M. (2010) Analysis of the Intact Surface Layer of Caulobacter Crescentus by Cryo-Electron Tomography. *J. Bacteriol. 192* (22), 5855−5865.
(34) Bingle, W. H., Nomellini, J. F., and Smit, J. (1997) Cell-Surface Display of a Pseudomonas Aeruginosa Strain K Pilin Peptide within the Paracrystalline S-Layer of Caulobacter Crescentus. *Mol. Microbiol. 26* (2), 277−288.
(35) Park, D. M., Reed, D. W., Yung, M. C., Eslamimanesh, A., Lencka, M. M., Anderko, A., Fujita, Y., Riman, R. E., Navrotsky, A., and Jiao, Y. (2016) Bioadsorption of Rare Earth Elements through Cell Surface Display of Lanthanide Binding Tags. *Environ. Sci. Technol. 50* (5), 2735−2742.
(36) Gandham, L., Nomellini, J. F., and Smit, J. (2012) Evaluating Secretion and Surface Attachment of SapA, an S-Layer-Associated Metalloprotease of Caulobacter Crescentus. *Arch. Microbiol. 194* (10), 865−877.
(37) Zakeri, B., Fierer, J. O., Celik, E., Chittock, E. C., Schwarz-Linek, U., Moy, V. T., and Howarth, M. (2012) Peptide Tag Forming a Rapid Covalent Bond to a Protein, through Engineering a Bacterial Adhesin. *Proc. Natl. Acad. Sci. U. S. A. 109* (12), E690−E697.
(38) Reddington, S. C., and Howarth, M. (2015) Secrets of a Covalent Interaction for Biomaterials and Biotechnology: SpyTag and SpyCatcher. *Curr. Opin. Chem. Biol. 29*, 94−99.
(39) Lau, J. H. Y., Nomellini, J. F., and Smit, J. (2010) Analysis of High-Level S-Layer Protein Secretion in Caulobacter Crescentus. *Can. J. Microbiol. 56* (6), 501−514.
(40) Nomellini, J. F., Duncan, G., Dorocicz, I. R., and Smit, J. (2007) S-Layer-Mediated Display of the Immunoglobulin G-Binding Domain of Streptococcal Protein G on the Surface of Caulobacter Crescentus: Development of an Immunoactive Reagent. *Appl. Environ. Microbiol. 73* (10), 3245−3253.
(41) Hagen, A., Sutter, M., Sloan, N., and Kerfeld, C. A. (2018) Programmed Loading and Rapid Purification of Engineered Bacterial Microcompartment Shells. *Nat. Commun. 9* (1), 2881.
(42) Bedbrook, C. N., Kato, M., Ravindra Kumar, S., Lakshmanan, A., Nath, R. D., Sun, F., Sternberg, P. W., Arnold, F. H., and Gradinaru, V. (2015) Genetically Encoded Spy Peptide Fusion System to Detect Plasma Membrane-Localized Proteins In Vivo. *Chem. Biol. 22* (8), 1108−1121.
(43) Campbell, R. E., Tour, O., Palmer, A. E., Steinbach, P. A., Baird, G. S., Zacharias, D. A., and Tsien, R. Y. (2002) A Monomeric Red Fluorescent Protein. *Proc. Natl. Acad. Sci. U. S. A. 99* (12), 7877−7882.
(44) Muiznieks, L. D., Reichheld, S. E., Sitarz, E. E., Miao, M., and Keeley, F. W. (2015) Proline-Poor Hydrophobic Domains Modulate the Assembly and Material Properties of Polymeric Elastin: Role of Domain 30 in Elastin Assembly and Material Properties. *Biopolymers 103* (10), 563−573.
(45) Sun, F., Zhang, W.-B., Mahdavi, A., Arnold, F. H., and Tirrell, D. A. (2014) Synthesis of Bioactive Protein Hydrogels by Genetically Encoded SpyTag-SpyCatcher Chemistry. *Proc. Natl. Acad. Sci. U. S. A. 111* (31), 11269−11274.
(46) Wichner, S. M., Mann, V. R., Powers, A. S., Segal, M. A., Mir, M., Bandaria, J. N., DeWitt, M. A., Darzacq, X., Yildiz, A., and Cohen, B. E. (2017) Covalent Protein Labeling and Improved Single-Molecule Optical Properties of Aqueous CdSe/CdS Quantum Dots. *ACS Nano 11* (7), 6773−6781.
(47) Mann, V. R., Powers, A. S., Tilley, D. C., Sack, J. T., and Cohen, B. E. (2018) Azide-Alkyne Click Conjugation on Quantum Dots by Selective Copper Coordination. *ACS Nano 12* (5), 4469−4477.
(48) Sakimoto, K. K., Wong, A. B., and Yang, P. (2016) Self-Photosensitization of Nonphotosynthetic Bacteria for Solar-to-Chemical Production. *Science 351* (6268), 74−77.
(49) Rothfuss, H., Lara, J. C., Schmid, A. K., and Lidstrom, M. E. (2006) Involvement of the S-Layer Proteins Hpi and SlpA in the Maintenance of Cell Envelope Integrity in Deinococcus Radiodurans R1. *Microbiology 152* (9), 2779−2787.
(50) Gerbino, E., Carasi, P., Mobili, P., Serradell, M. A., and Gómez-Zavaglia, A. (2015) Role of S-Layer Proteins in Bacteria. *World J. Microbiol. Biotechnol. 31* (12), 1877−1887.
(51) Schultze-Lam, S., Harauz, G., and Beveridge, T. J. (1992) Participation of a Cyanobacterial S Layer in Fine-Grain Mineral Formation. *J. Bacteriol. 174* (24), 7971−7981.
(52) Küpcü, S., Mader, C., and Sára, M. (1995) The Crystalline Cell Surface Layer from Thermoanaerobacter Thermohydrosulfuricus L111−69 as an Immobilization Matrix: Influence of the Morphological Properties and the Pore Size of the Matrix on the Loss of Activity of Covalently Bound Enzymes. *Biotechnol. Appl. Biochem. 21* (3), 275−286.
(53) Martín, M. J., Lara-Villoslada, F., Ruiz, M. A., and Morales, M. E. (2015) Microencapsulation of Bacteria: A Review of Different Technologies and Their Impact on the Probiotic Effects. *Innovative Food Sci. Emerging Technol. 27*, 15−25.
(54) Schoebitz, M., López, M. D., and Roldán, A. (2013) Bioencapsulation of Microbial Inoculants for Better Soil−plant Fertilization. A Review. *Agron. Sustainable Dev. 33*, 751.
(55) Herrero, M., and Stuckey, D. C. (2015) Bioaugmentation and Its Application in Wastewater Treatment: A Review. *Chemosphere 140*, 119−128.
(56) Kandemir, N., Vollmer, W., Jakubovics, N. S., and Chen, J. (2018) Mechanical Interactions between Bacteria and Hydrogels. *Sci. Rep. 8* (1), 10893.
(57) Azam, A., and Tullman-Ercek, D. (2016) Type-III Secretion Filaments as Scaffolds for Inorganic Nanostructures. *J. R. Soc., Interface 13* (114), 20150938.
(58) Toporowski, M. C., Nomellini, J. F., Awram, P., and Smit, J. (2004) Two Outer Membrane Proteins Are Required for Maximal Type I Secretion of the Caulobacter Crescentus S-Layer Protein. *J. Bacteriol. 186* (23), 8000−8009.
(59) Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 Years of Image Analysis. *Nat. Methods 9* (7), 671−675.

DOI: 10.1021/acssynbio.8b00448
ACS Synth. Biol. 2019, 8, 181−190

Sector: Transforming Cellular Factories with Synthetic Biology

Use Case for Cells as Factories: Temporally and Spatially Controlled Production of Structural Materials

End Product: Various materials produced and exported by cells on queue

Organism(s) if applicable: Microbial cultures including archaea and extremophiles, mixed cultures

The rapid pace of increasing capabilities for bioengineering and synthetic biology has enabled dramatic improvements in our ability to design microorganisms as microscale factories to efficiently produce high value small molecules including pharmaceuticals, chemicals and commodity precursors. The value of capabilities for robust synthesis of diverse small molecules, including molecules not found in nature, is significant and warrants continued development. However, the synthetic capabilities of biology far exceed small molecules. Natural biological systems produce a dizzying array of higher order molecular assemblies at the mesoscale and beyond, often "cheating" thermodynamics to access disfavored material phases that require extreme processing conditions or are completely inaccessible via traditional materials synthesis routes. If we can harness the synthetic capabilities of biology to produce complex molecular assemblies and materials, we can revolutionize the production of structural materials. Engineered biological systems could be designed in which a single organism is programmed to produce a variety of products, each dependent on a specific input. This would enable a tunable materials synthesis platform in which different inputs drive the production of different output materials**.**

Engineered microbes have potential to produce mesoscale materials of desired function beyond molecules. Microbes already produce structures such as carboxysomes. Once infected by phages, microbes produce virus or virus-like particles that could be potentially be co-opted to perform novel structural function. Exported mesoscale materials maybe engineered to interface with non-biological materials such as synthetic polymers, metals, minerals, semi-conductors, or carbonaceous materials. Engineered microbes themselves may be able to serve as a structural component of high-ordered material or devices beyond serving as biochemical factories. Significant effort is underway to engineer microbial ecosystem that can produce specific molecules and compounds but the challenges remain in using cells to produce mesoscale/supramolecular structures that may lead to direct interfacing with abiotic materials in temporally and spatially controlled manner. Phenotypes borrowed from extremophiles (such as chemolithotrophy, biostasis) may endow engineered microbes to perform new chemistry on queue while minimizing the loss of resources to biomass once optimal population for production has been achieved.

Desired outcome(s) that stretch current capabilities

- Tools for temporal control of cellular production
- Engineered microbes that can produce and export variety of structural materials (e.g. magnetosomes, gas vesicles, virus-like particles)
- Engineered extremophiles that are genetically facile or engineered lab strains to withstand extreme conditions
- Tools for controlling multi-step reactions using co-cultures or community of microbes
- Tools for spatial control of microbes

Sector: Transforming Cellular Factories with Synthetic Biology

Use Case for Cells as Factories: Temporally and Spatially Controlled Production of Structural Materials using Artificial Cells and Cell Free Systems

End Product: Structural materials produced via temporally and spatially controlled biological synthesis of mesoscale molecular structures

Organism(s) if applicable: Various microbial organisms including archaea and extremophiles, mixed cultures, cell-free systems produced from these organisms, or artificial cells assembled from the bottom-up with isolated components from these organisms

The rapid pace of increasing capabilities for bioengineering and synthetic biology has enabled dramatic improvements in our ability to design microorganisms as microscale factories to efficiently produce high value small molecules. However, the synthetic capabilities of biology far exceed small molecules. While the engineering of living cells has already expanded our manufacturing capabilities and holds promise to revolutionize how we produce molecules and materials that impact multiple sectors spanning from medicine to textiles, living cells are inherently complex structures that have evolved to achieve outcomes that benefit the cell (reproduction, protection against threats, etc.). Engineering living cells has proven exceedingly difficult because it requires the modification of a system that is not fully understood, that is extraordinarily complex, that can function autonomously, that can exchange material with its environment, and that can evolve in response to external influences or threats. Living cells are also optimized to function in their natural environmental niche and are often unable to withstand dramatic shifts in temperature or environmental chemistry.

Artificial cells, non-living structures that are compatible with biological cellular machinery, or cell-free systems, either cell extracts or a collection of purified proteins/biomolecules, could provide a robust, predictable, non-autonomous platform for the execution of human-designed biological programs to realize novel material synthesis capabilities. Artificial cells and cell-free systems could serve as non-biological reaction platforms that (i) sequester human-designed biological programs and the molecules needed to execute these programs, (ii) execute programs only when triggered, with no spurious activity and high signal to noise, (iii) can be stored in an inactive state for extended periods of time with highly reliable “re-boot” when needed, and (iv) are able to execute their programs over a wide range of environmental conditions.

For synthesis of complex biomolecular assemblies and mesoscale materials, artificial cells and cell-free systems may provide distinct advantages over living cells as manufacturing platforms. These systems may more readily accommodate production of complexes and materials of increased physical dimensions, as cell-free systems have no external bounds and the physical scale of artificial cells could be tuned based on the desired product. Moreover, concerns related to cell toxicity and cellular resource allocation are dramatically simplified in artificial cell and cell-free systems.

Desired outcome(s) that stretch current capabilities

- Integration of discrete artificial cell functions into coordinated artificial cell platforms
- Optimization of cell-free systems to enable a robust, versatile manufacturing capability for mesoscale molecular structures and materials.

nature COMMUNICATIONS

ARTICLE

DOI: 10.1038/s41467-018-07473-7 OPEN

Communication and quorum sensing in non-living mimics of eukaryotic cells

Henrike Niederholtmeyer [1], Cynthia Chaggan[1] & Neal K. Devaraj[1]

Cells in tissues or biofilms communicate with one another through chemical and mechanical signals to coordinate collective behaviors. Non-living cell mimics provide simplified models of natural systems; however, it has remained challenging to implement communication capabilities comparable to living cells. Here we present a porous artificial cell-mimic containing a nucleus-like DNA-hydrogel compartment that is able to express and display proteins, and communicate with neighboring cell-mimics through diffusive protein signals. We show that communication between cell-mimics allows distribution of tasks, quorum sensing, and cellular differentiation according to local environment. Cell-mimics can be manufactured in large quantities, easily stored, chemically modified, and spatially organized into diffusively connected tissue-like arrangements, offering a means for studying communication in large ensembles of artificial cells.

[1] Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla CA 92093, USA. Correspondence and requests for materials should be addressed to N.K.D. (email: ndevaraj@ucsd.edu)

In communities of single-celled and multicellular organisms, cell–cell communication enables cells to organize in space, distribute tasks, and to coordinate collective responses. Synthetic biologists have engineered living, communicating cells to form cellular patterns[1,2] and synchronize gene expression[3] but living systems are inherently challenging to study and engineer. Chemically constructed cell-mimics, as non-living, biochemically simplified and engineerable systems, could serve as models to study mechanisms of pattern formation and collective responses, and lead to the development of novel sensors and self-organizing materials. Important biochemical processes like protein synthesis[4,5], DNA replication[6], metabolism[7], and cytoskeletal functions[8] have been reconstituted and studied in single synthetic cell-mimics. While biochemical reactions in microfluidic chambers[9–11], in droplets[12,13] and on beads[14] can emulate aspects of intercellular communication, studies on systems that structurally resemble natural cells with their semi-permeable membranes have been limited in scope by the availability of communication channels and assembly methods. Addressing the scalable assembly of artificial cells, microfluidic methods have been developed to mass-produce highly homogeneous populations of phospholipid vesicles encapsulating active biomolecules[15–18]. Recent studies have demonstrated communication between synthetic microcompartments to induce gene expression[5,13,19,20] or chemical reactions[21–23] using small molecule signals. To implement communication, signaling molecules must travel between compartments. Some small molecules diffuse freely between compartments[5,13,19–22], phospholipid vesicles can be permeabilized by inserting alpha-hemolysin pores[5,23], and other synthetic microcompartments such as gel-shell beads[24], polymersomes[21], proteinosomes[23], and colloidosomes[22] can be assembled with permeable membranes. Signaling molecules for communication between artificial cell-mimics have so far been limited to small molecules. In contrast, signaling in multicellular organisms often involves secretion of proteins serving as growth factors or morphogens that provide cells with the information they need to develop into functional tissues[25].

Here, we aim to expand the communication capabilities of artificial cells by developing a cellular mimic that produces and releases diffusive protein signals that travel in and get interpreted by large populations of cell-mimics. We describe the microfluidic production of cell-mimics with a porous polymer membrane containing an artificial hydrogel compartment, which resembles a eukaryotic cell's nucleus in that it contains the cell-mimics' genetic material for protein synthesis and can sequester transcription factors. Cell-mimics are able to communicate through diffusive protein signals, activate gene expression in neighboring cell-mimics, and display collective responses to cell-mimic density similar to bacterial quorum sensing.

Results

Porous cell-mimics containing artificial nuclei. We prepared porous cell-mimics capable of gene expression and communication via diffusive protein signals using a microfluidic method (Fig. 1a, b). First, water-in-oil-in-water double emulsion droplets were formed in a polydimethylsiloxane (PDMS) device (Supplementary Figure 1, Supplementary Movie 1). The droplets had a middle organic phase consisting of a 1-decanol and acrylate monomer solution and encapsulated DNA and clay minerals. Second, double emulsion droplets were collected and polymerized using UV light, inducing a phase separation of the inert 1-decanol to form porous microcapsules[26]. Third, following polymerization, we simultaneously permeabilized the polymer membrane and induced formation of a clay-hydrogel in their interior by adding a solution of ethanol and HEPES buffer. Membrane pores had diameters of 200–300 nm (Fig. 1a, Supplementary Figure 2). Polymer membranes were permeable to macromolecules up to 2 MDa but excluded 220 nm nanoparticles from about 90% of the microcapsules (Supplementary Figure 3). Like in similarly prepared porous microcapsules[26,27], polymer membranes were mechanically stable and rigid. Microcapsules could be centrifuged at high speeds, and only broke under high stress from a razor blade (Supplementary Figure 2). The encapsulated clay minerals

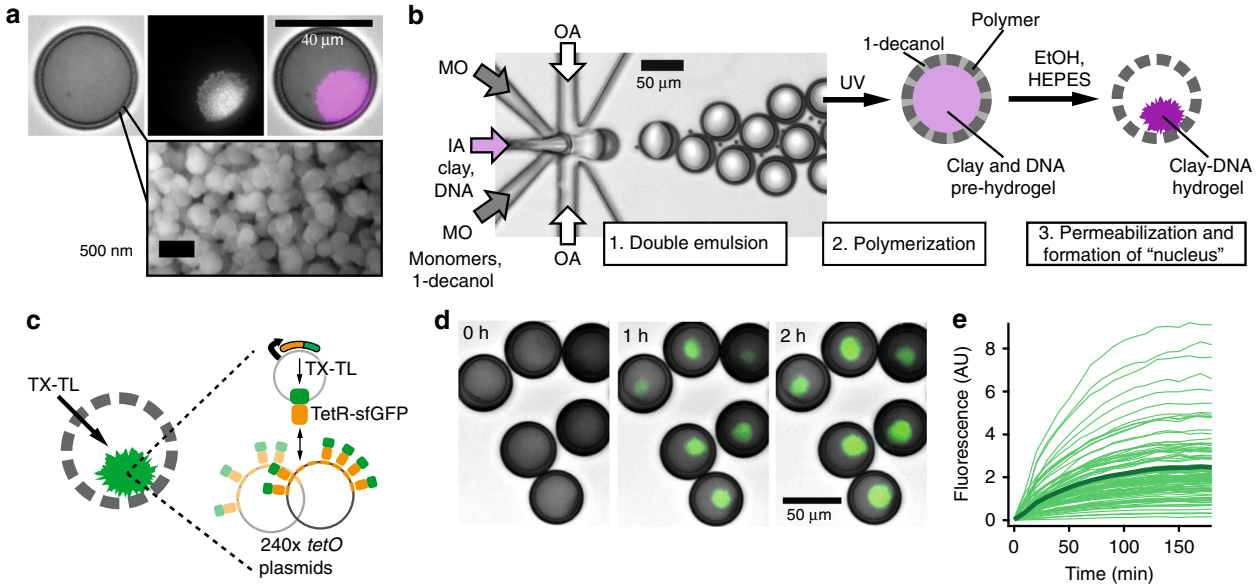


Fig. 1 Formation of cell-mimics with artificial nuclei capable of gene expression. **a** Optical micrographs (top) of a cell-mimic with GelRed stained hydrogel nucleus (brightfield, red fluorescence, merge) and scanning electron microscopy of porous cell-mimic membrane (bottom). **b** Microfluidic production of double emulsion droplets encapsulating a pre-hydrogel in a photocurable middle layer, and schematic of subsequent processing steps. IA: Inner aqueous, MO: middle organic, OA: Outer aqueous phase. **c** Schematic and timelapse images **d** of expression and capture of TetR-sfGFP in hydrogel nuclei (green, merged with brightfield images). TX-TL reagents were added at 0 h. **e** Dynamics of fluorescence signal increase in the hydrogel nuclei of 100 cell-mimics with average shown in bold

had a large capacity for binding and capturing DNA from solution, and retained DNA in the clay-DNA hydrogel that formed after electrolyte addition (Supplementary Figure 4)[28]. During permeabilization of the polymer membrane with a solution of HEPES buffer and ethanol, the encapsulated clay minerals and DNA condensed into a round hydrogel structure in the microcapsules' interior that is analogous to the cell nucleus (Fig. 1c). In at least 95% of microcapsules, only one clay-DNA hydrogel nucleus formed that measured approximately half of the microcapsules' diameter in size. In solution, not templated by a surrounding microcapsule, clay minerals form irregularly shaped and sized hydrogel particles (Supplementary Figure 5). DNA adsorbs to the surfaces of clay minerals through electrostatic interactions, cation bridges, hydrogen bonding and ligand exchange[28–30]. Clay-DNA hydrogel nuclei had a porous structure, allowing macromolecules of up to 500 kDa to diffuse into the hydrogel while 2000 kDa dextran was partially excluded from hydrogel nuclei (Supplementary Figure 3). The permeability of the hydrogel nuclei to large macromolecules suggested that the DNA would be accessible to the transcription machinery. As a final step in the preparation of cell-mimics, we passivated their polymer membrane with polyethylene glycol (PEG) to prevent non-specific binding of proteins (Supplementary Figure 6).

Gene expression in cell-mimics. The porous structure of the polymer membrane allowed supply of cell-free transcription and translation (TX-TL) reagents from the outside to induce synthesis of proteins encoded by the DNA in the cell-mimics' hydrogel nuclei. Even ribosomes, the largest components of TX-TL reagents, were able to diffuse into cell-mimics through their porous membranes (Supplementary Figure 7). To capture protein products within cell-mimics, we expressed a fusion protein of the tetracycline repressor TetR and sfGFP (TetR-sfGFP) as a fluorescent reporter. TetR binds the *tet* operator sequence (*tetO*). A co-encapsulated 240 × *tetO* array plasmid localized the reporter protein to the hydrogel nucleus (Fig. 1c), which increased in fluorescence after TX-TL addition (Fig. 1d, Supplementary Movie 2). Localization of TetR-sfGFP to the hydrogel nucleus was reversible and due to the specific interaction of TetR with *tetO* sites. Addition of anhydrotetracyline, which prevents TetR from binding DNA, caused a substantial unbinding of TetR-sfGFP. Without the *tetO* plasmid, fluorescence increased in solution but not in hydrogel nuclei (Supplementary Figure 8). In *tetR*-sfGFP/*tetO* cell-mimics, fluorescence increased substantially in almost all cell-mimics (Fig. 1e). Variations in intensity were likely due to differences in DNA capture during formation of hydrogel nuclei (Supplementary Fig. 5). Cell-mimics retained full expression capabilities after 2 years of storage, and separate batches showed comparable expression levels and dynamics (Supplementary Figure 9).

Protein exchange between cell-mimics. Due to their porosity, cell-mimics likely released mRNA and protein products that diffused into neighboring cell-mimics. To demonstrate that neighboring cell-mimics exchanged protein products with each other, we prepared sender cell-mimics, fluorescently labeled in their polymeric membranes and containing the *tetR*-sfGFP expression plasmid, and receiver cell-mimics containing the *tetO* array plasmid to capture the reporter protein. When both cell-mimic types were mixed at approximately a one to one ratio, only the nuclei of the receiver cell-mimics increased in fluorescence (Fig. 2a). To explore how far TetR-sfGFP protein originating from a given sender cell-mimic travelled, we used a large excess of receivers and spread them densely into a circular 3.5 mm wide colony. Under these conditions, TetR-sfGFP spread from sender to surrounding receiver cell-mimics but stayed localized in patches around individual sender cell-mimics (Fig. 2b, Supplementary Figure 10a). This pattern of captured protein around source cell-mimics persisted for 24 h after expression ended, demonstrating that TetR-sfGFP was essentially trapped in the hydrogel nuclei once it was bound in the high local density of *tetO* sites. Assuming free diffusion, we would expect protein gradients to have disappeared within 5 h in similar geometries (Supplementary Figure 10b–d).

To test the preference of a given cell-mimic to bind protein originating from its own DNA, we prepared *tetR*-mCherry / *tetO* cell-mimics that accumulated red fluorescence in their hydrogel nuclei (Supplementary Fig. 10). When mixed with *tetR*-sfGFP / *tetO* cell-mimics (Fig. 1c), there was essentially no difference in relative fluorescence in either channel between the cell-mimic types, indicating that in close proximity, neighboring cell-mimics completely exchanged protein products (Supplementary Figure 11). While transcription occurred in the hydrogel nuclei where DNA was localized, these results indicate that translation was likely not localized to the cell-mimic a given mRNA originated from. However, because mRNA lifetime in TX-TL reagents is short, and mRNA thus has a limited diffusion range, we expected the localization of TetR-sfGFP and TetR-mCherry to depend strongly on distance between cell-mimics. We distributed the two cell-mimic types in a reaction chamber so that they mixed in the center but remained separate on either side. Cell-mimics in the center showed mixed fluorescence while cell-mimics on the sides fluoresced primarily in one channel (Fig. 2c), demonstrating that locally, on the order of few cell-mimic lengths, proteins exchanged with little hindrance by the polymer membranes, whereas exchange of protein with distant cell-mimics was limited by diffusion.

Communication through a diffusive transcriptional activator. Communication in vesicle-based cell-mimics has so far been limited to small molecule signals such as quorum sensing molecules[19,20] or IPTG and glucose, combined with membrane pores, like alpha-hemolysin[5,23]. Our porous cell-mimics exchanged proteins with their neighbors, suggesting they are able to communicate with each other directly through genetic regulators. To demonstrate this we constructed a two-stage activation cascade and distributed the network into two separate cell-mimic types. T3 RNA polymerase (T3 RNAP) served as a diffusive signaling molecule transmitting the instruction to express a reporter gene from activator to reporter cell-mimics. Activator cell-mimics contained the template for the expression of T3 RNAP. Reporter cell-mimics contained the template for the T3 RNAP-driven synthesis of the TetR-sfGFP reporter as well as *tetO* array plasmids to capture the reporter protein. When both cell-mimic types were mixed, reporter cell-mimics expressed and bound the fluorescent reporter (Fig. 3, Supplementary Movie 3), while activator cell-mimics alone did not increase in fluorescence (Supplementary Figure 12).

Artificial quorum sensing. We hypothesized that T3 RNAP could serve as a soluble signaling molecule providing cell-mimics with information about population density. Indeed, cell-mimics containing both the activation circuit and reporter constructs (Fig. 4a) underwent a collective response where fluorescence accumulated in cell-mimics only at high densities. At low cell-mimic densities, signals from the hydrogel nuclei were not detectably different from background fluorescence (Fig. 4b). We titrated the density of cell-mimics in a fixed volume and found a sharp transition from off to on, which resembled bacterial quorum sensing responses to cell density[31] even though, unlike in

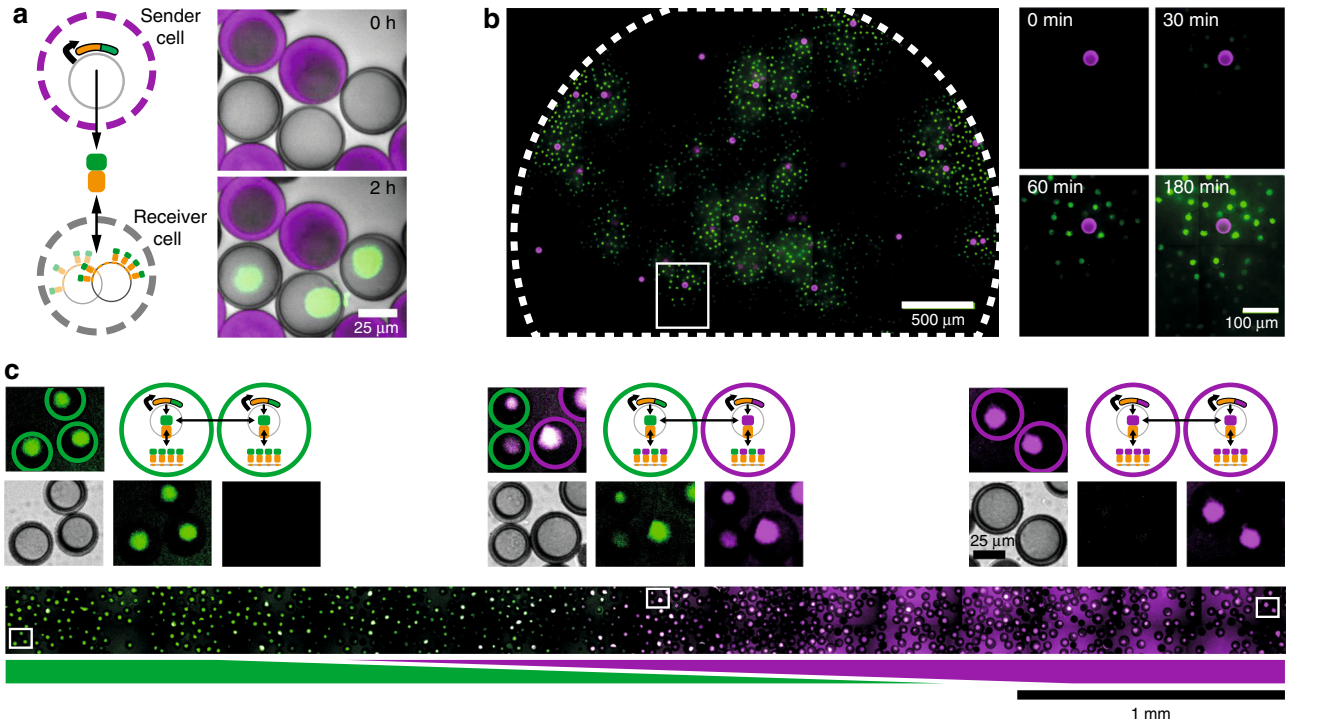


Fig. 2 Protein exchange between neighboring and distant cell-mimics. **a** Schematic of diffusive TetR-sfGFP exchange between neighboring sender and receiver cell-mimics and timelapse images of neighboring senders (rhodamine B stained membranes) and receivers (unstained membranes). Merge of brightfield and fluorescence channels (sender membranes, magenta; TetR-sfGFP, green). **b** Distribution of TetR-sfGFP (green) in a dense droplet of receivers and sparse senders (magenta) after 3 h of expression. A small region around a sender (white box) is magnified and spreading of fluorescence is shown at different time points. **c** Inhomogeneous mix of two types of cell-mimics producing and binding different color reporter proteins. *tetR*-sfGFP / *tetO* (green) and *tetR*-mCherry / *tetO* cell-mimics (magenta) were distributed in a channel to stay separate at the sides and mix in the center. Bottom image shows the distribution of sfGFP and mCherry fluorescence after 5 h. Merge of the two channels results in a white signal (middle). Magnified images from indicated positions along the channel are shown above. Merged image with cell-mimic types indicated by colored, dashed circles (top), and brightfield, sfGFP and mCherry signals shown separately (below)

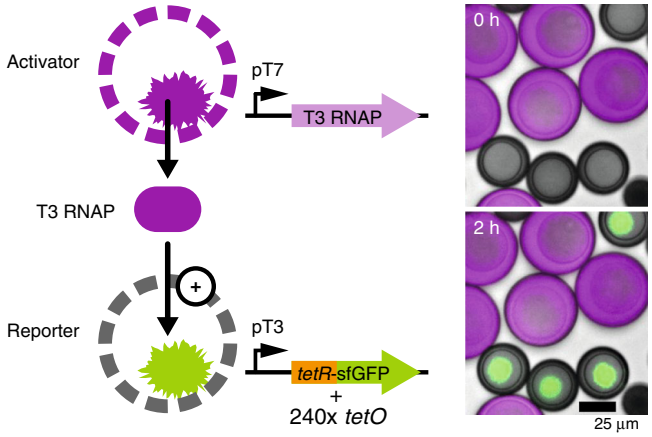


Fig. 3 Communication between cell-mimics via a diffusive genetic activator. Schematic of the two types of cell-mimics communicating through a distributed genetic activation cascade. Micrographs show a merge of brightfield images with rhodamine B fluorescence in the membranes of activators (magenta) and fluorescence of TetR-sfGFP (green) in the hydrogel nuclei of reporters directly after addition of TX-TL and after 2 h of expression

bacterial quorum sensing mechanisms, the T3 activation circuit did not contain positive feedback. The threshold cell-mimic density at which expression of the reporter turned on was 400 cell-mimics in 4.5 µl TX-TL (Fig. 4c). Cell-mimics that constitutively expressed the reporter (Fig. 1c) accumulated fluorescence in their hydrogel nuclei regardless of their density (Supplementary Figure 13). The collective response to density can be explained by T3 RNAP release from cell-mimics. At low densities, T3 RNAP is diluted in the comparably large volume of the sample, while at high density a sufficient concentration of transcriptional activator accumulates to turn on expression of the reporter. Titrating the T3 RNAP template DNA in TX-TL reactions, we found a steep transition from low to high expression with a half-maximal activation at 10 pM (Supplementary Figure 14). The calculated bulk concentration of T3 RNAP template in an artificial quorum sensing experiment at the threshold density of 400 cell-mimics per droplet is 12.5 pM, similar to the activation threshold in bulk solution.

During development, cells interpret signals secreted by their neighbors to differentiate into specialized cell-types that express different sets of genes[25]. We aimed to emulate cellular differentiation according to local environment by combining the artificial quorum sensing network with a constitutively expressed *tetR*-mCherry reporter that turns on irrespective of cell-mimic density (Fig. 4d). We distributed cell-mimics unevenly in a long narrow reaction chamber (Fig. 4e), and analyzed the fluorescence of individual hydrogel nuclei according to their location in the density gradient. While absolute fluorescence intensities and background fluorescence increased with cell-mimic density, hydrogel nuclei from the high density area displayed visibly higher sfGFP:mCherry ratios than hydrogel nuclei in the dilute region that primarily displayed mCherry fluorescence (Fig. 4f). In the continuous density gradient, we

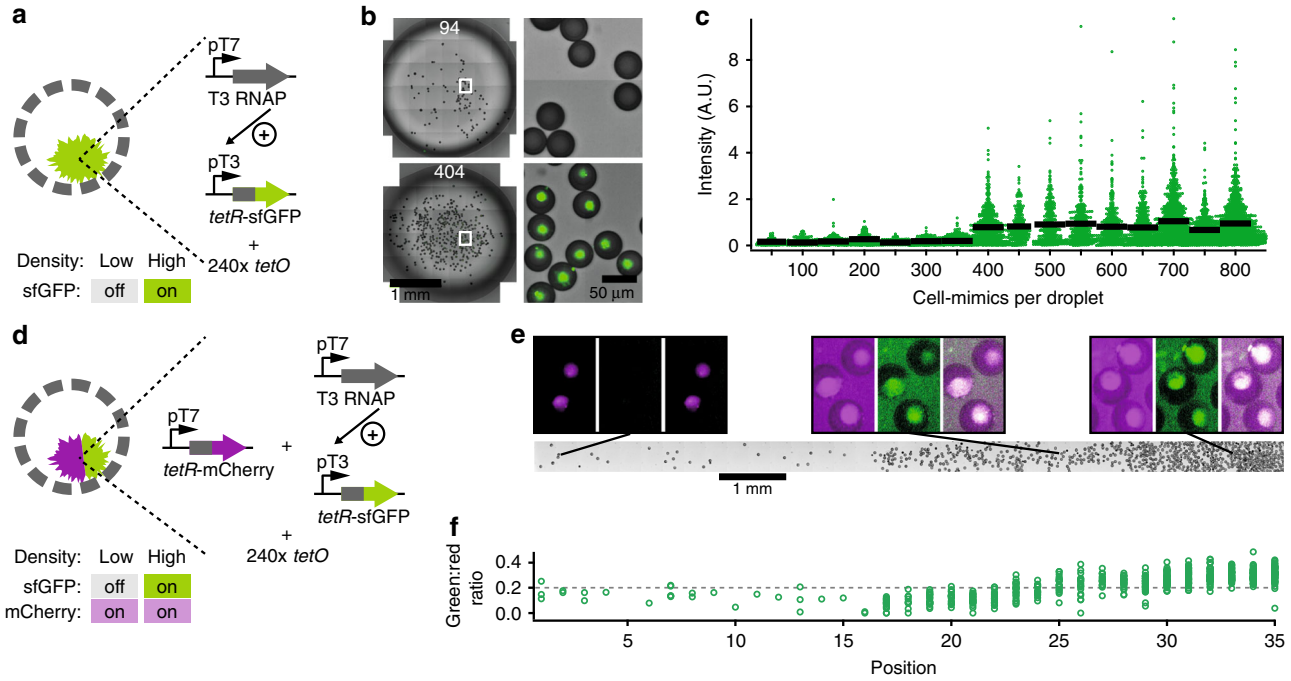


Fig. 4 Density sensing in populations of cell-mimics. **a** Artificial quorum sensing cell-mimics contain T3 activation cascade DNA templates and 240 × *tetO* plasmids. **b** Micrographs of cell-mimics in 4.5 µl droplets of TX-TL (left). The number of cell-mimics is indicated. Enlarged regions (indicated by white boxes) show presence and absence of fluorescence (green) in hydrogel nuclei after 3 h of expression. **c** Scatter dot plot of fluorescence intensities in individual cell-mimics at different densities. Each density category combines data in increments of 50 cell-mimics per droplet and contains data from at least 156 cell-mimics (Methods). Black bars show average fluorescence. **d** A 2-color response to density is achieved by adding a constitutively expressed reporter (pT7-*tetR*-mCherry), which is on independent of density. **e** 2-color density sensors were spread at increasing density in an elongated chamber (brightfield image, bottom). Panels above show magnified fluorescence images of indicated regions (mCherry fluorescence: magenta, left; sfGFP fluorescence: green, middle; merge of fluorescence channels: right). Images in each channel are window leveled to the same values and have a width of 70 µm. **f** Ratio between sfGFP to mCherry fluorescence in individual hydrogel nuclei along the chamber. Positions correspond to tile regions of the image above

observed some graded responses in the center of the chamber at medium density. However, plotting sfGFP against mCherry fluorescence for individual hydrogel nuclei revealed two distinct populations of fluorescence signals according to position in the chamber (Supplementary Figure 15).

Discussion

We developed porous cell-mimics capable of gene expression and communication via diffusive protein signals. The clay-DNA hydrogel in the cell-mimic's interior resembles a eukaryotic cell's nucleus, and represents a different strategy to compartmentalize artificial cells[17,32,33]. Our study demonstrates that clay minerals and clay-DNA hydrogels are useful hybrid materials for synthetic biology and the assembly of artificial cell-mimics. In fact, clay minerals have been proposed as favorable environments for prebiotic evolution because a wide variety of organic molecules adsorbs to their surface. Clay minerals have been shown to possess catalytic properties[34] and to enhance protein synthesis in cell-free expression systems[28]. Unlike lipid vesicles that require careful matching of osmolarities and gentle separation techniques, the cell-mimics reported here with their porous polymer membranes were physically highly stable, easily transferred into new media by centrifugation, and retained full expression capabilities after 2 years of storage. Microfluidic production of highly homogeneous cell-mimics will facilitate studies requiring large quantities of cell-mimics as our results on spatially arrayed, communicating cell-mimics demonstrate. Passivating cell-mimics' polymer membranes with PEG, we showed that their membranes could be chemically modified, which will allow further functionalization, for example, for immobilization on substrates, to target specific proteins to the membrane or to tune membrane permeability.

So far, communication in synthetic, non-living cell-mimics has been limited to small molecule signals[5,19,20,22,23]. The porous cell-mimics developed here expand the communication capabilities of artificial cells to large macromolecules like RNAs and proteins. We showed that genetic circuits could be distributed into separate cell-mimics, which allowed them to share tasks. Such modularity might facilitate the assembly of functional, distributed gene circuits by titrating cell-mimics containing different parts of a network. Furthermore, individual components of circuits located in different cell-mimics can be spatially organized to generate spatiotemporal expression patterns. Protein signals play an important role for cell–cell communication in multicellular organisms, where cells release and receive protein signals in the form of hormones, growth factors, and morphogens[25]. We believe that reaction–diffusion models explaining developmental processes in multicellular organisms may be tested and emulated in artificial, tissue-like arrangements of cell-mimics even when they are composed of unnatural materials, operate on different scales and with different molecular mechanisms. Clearly, a major difference between an artificial tissue assembled from porous cell-mimics and natural tissues is that translation of proteins takes place both inside and outside of the cell-mimic a given mRNA originates from, and that mRNAs and proteins can freely diffuse. In this regard, our system emulates features of the syncytium stage during *Drosophila* embryogenesis, when thousands of nuclei accumulate in the unseparated cytoplasm of the oocyte[35]. In another analogy to this developmental stage, we showed trapping

of the TetR transcription factor in nuclei containing binding sites. Nuclear trapping is a mechanism that is responsible for establishing sharp gradients of phosphorylated ERK/MAPK (dpERK) across the syncytical *Drosophila* embryo by limiting diffusion of dpERK[36]. In future work, diffusion rates of proteins through tissue-like arrangements of cell-mimics may be tuned by using nuclear binding sites that restrict diffusion of DNA binding proteins or by modifying polymer membrane permeability.

A genetic activation circuit led to a remarkable collective response to cell-mimic density, which resembled bacterial quorum sensing. In contrast to bacterial quorum sensing[31], the mechanism of our artificial quorum sensing involved no positive feedback loop and employed a protein instead of a small molecule, showing that artificial cells can emulate biological phenomena using unnatural parts and mechanisms. Several molecular mechanisms can produce cooperative or ultrasensitive behaviors[37]. Polyvalent interactions enhance functional affinities between binding partners by favoring ligand rebinding after an initial binding event occurred[37–39]. We believe the switch-like behavior of the artificial quorum sensing response is due to cooperative binding of the T3 RNAP to highly polyvalent hydrogel nuclei containing high local concentrations of T3 RNAP promoters, combined with large distances between cell-mimics at low densities (Supplementary Figure 16), and an already sigmoidal response curve to T3 RNAP template concentration (Supplementary Figure 14). Collective responses can lead to greater accuracy and reduce noise[3], which will be particularly helpful for the assembly of reliably functioning cell-mimics, which often suffer from variability in gene expression[5,40].

In conclusion, our system has a number of potential uses, including programming cell-mimics to collectively sense and respond to their environment. Indeed, artificial cell-mimics could be used to develop sensors and self-organizing materials, as well as being arrayed into synthetic tissues of artificial cells, which could serve as simplified models for reaction–diffusion processes.

Methods

Fabrication of microfluidic chips. PDMS devices were prepared by standard soft lithography methods to produce devices with a design as shown in Supplementary Figure 1. A silicon wafer patterned with SU-8 photoresist served as a mold for PDMS devices. The mold was prepared following standard photolithography procedures to produce a feature height of 43 µm. Design of the flow-focusing junction was adapted from Desphande et al.[16] (Supplementary Figure 1). PDMS (Sylgard, Dow Corning) was prepared at a 1:10 ratio and cured for 1 h at 80 °C. PDMS devices were bonded to PDMS spincoated cover glass using oxygen plasma (50 watt for 30 s at 0.45 torr). After bonding, devices were baked overnight at 120 °C to recover hydrophobicity. Channels downstream of the flow-focusing junction were rendered hydrophilic by a treatment with a 5% (wt/vol) polyvinyl alcohol (PVA) solution[16], which was flowed in the inlet for the outer aqueous solution for 5 min while blowing air into the other inlets to avoid contact of the PVA-solution with device regions upstream of the flow-focusing junction. The remaining PVA-solution was removed by applying vacuum to the outlet. Following PVA-treatment devices were baked again for 2 h at 120 °C and used immediately or stored for up to 2 months.

Production of porous cell-mimics with DNA-hydrogel nuclei. 2% (wt/vol) Laponite XLG (BYK Additives) clay stock was prepared by mixing 10 ml of ultrapure H_2O on a magnetic stir plate to create a vortex. 200 mg of Laponite XLG were slowly added into the vortex and left to stir for 2 h until clear. The dispersion was then stored at 4 °C and used for up to a week. Photoinitiator 2,2-dimethoxy-2-phenylacetophenone was dissolved at 5% (wt/vol) in 1-decanol and in Trimethylolpropane ethoxylate triacrylate (ETPTA, Sigma-Aldrich, Mn 428). ETPTA with photoinitiator was stored at 4 °C and used for up to a week. Double emulsion droplets were prepared with an inner aqueous solution (IA) containing 0.4% (wt/vol) laponite XLG, 15% (vol/vol) glycerol, 50 mg/ml poloxamer 188, 20 µM sulfo-Cy5 and up to 300 ng/µl plasmid or linear DNA. The middle organic phase (MO) was composed of glycidyl methacrylate (GMA, Sigma-Aldrich), ETPTA, and 1-decanol at a 48:32:20 ratio and contained 2.6% (wt/vol) photoinitiator and 0.25% (vol/vol) Span-80 to produce porous microcapsules[26]. For fluorescently labeled microcapsule membranes, the MO phase contained 0.1 mg/ml methacryloxyethyl thiocarbamoyl rhodamine B. The outer aqueous phase (OA) was 15% (vol/vol) glycerol with 50 mg/ml poloxamer 188.

Using syringe pumps, the three phases were flowed through the microfluidic device at speeds of 3 to 12 µl/h for the IA, 30 to 70 µl/h for the MO and 250 to 500 µl/h for the OA phase. Flow rates were adjusted to produce a stable formation of double emulsion droplets and then left unchanged for collection of droplets. Typically, about 200 µl of double emulsion were collected from the chip. The emulsion was then placed in a 2 mm thick chamber built from cover glass and exposed to 350 nm UV light for 30 s using a UV reactor (Rayonet). The dispersion of polymerized microcapsules was then added to 2 ml solution of 70% Ethanol containing 200 mM HEPES pH 8 to permeabilize the shell and to form the DNA-clay-hydrogel nucleus. This stock was stored at −20 °C until use.

To prevent non-specific binding of proteins to porous polymer membranes, microcapsules were treated with polyethylene glycol (PEG). We coupled amino-PEG12-alcohol to the epoxide functionalities on the polymer shells. First, microcapsules were washed with 200 mM sodium carbonate buffer pH 10 by centrifugation. All supernatant was removed from the capsule pellet and a solution of 250 mM amino-PEG12-alcohol in 50% ethanol pH 10 was added to the pellet. Microcapsules were incubated at 37 °C for reaction overnight and then washed with 100 mM HEPES pH 8. These PEGylated cell-mimics were either used directly or stored in 70% ethanol 200 mM HEPES pH 8 at −20 °C.

DNA templates. Plasmids used in this study are listed in Supplementary Table 1. Plasmid DNA was purified using the NucleoBond Xtra Midi kit (Macherey-Nagel), followed by an isopropanol precipitation and resuspension of the DNA pellet in ultrapure H_2O to prepare highly concentrated plasmid stocks and maximize expression in TX-TL reactions. The template for T3 RNA polymerase was on linear DNA prepared by PCR from a plasmid template. DNA template concentrations encapsulated in cell-mimics are listed in Supplementary Table 2.

Cell-free transcription and translation reactions. *E. coli* lysate for TX-TL reactions was prepared by freeze-thawing from *E. coli* BL21-Gold (DE3)/pAD-LyseR[41]. To induce T7 RNA polymerase activity besides *E. coli* RNA polymerase activity the main culture was induced with 0.5 mM IPTG an optical density measured at 600 nm (OD_{600}) of 0.5–0.6. Cells were harvested by centrifugation at an OD_{600} of 1.4 and resuspended in ice cold S30A buffer (14 mM magnesium glutamate, 60 mM potassium glutamate, 50 mM Tris, pH 7.7). Cells were again pelleted by centrifugation, and after determining pellet mass, cells were resuspended in two volumes (relative to cell mass) of S30A buffer containing 2 mM DTT and frozen at −80 °C. For lysate preparation, cells were thawed, vigorously vortex mixed for 3 min and incubated on an orbital shaker at 300 rpm at 37 °C for 45 min. Cell suspensions were again vortex mixed and incubated for another 45 min under shaking at 37 °C. Following incubation, cells were again vigorously mixed and the lysate was cleared by centrifugation at 50,000×*g* for 45 min. Cell debris-free, clear supernatant was collected and frozen in aliquots at −80 °C until use. For the final composition of TX-TL reactions, cell lysate was diluted 2.5-fold with reaction buffer, microcapsules or DNA, and other additions as needed. This resulted in the following concentrations in the TX-TL reaction: 4.6 mg/ml protein, 7 mM Mg-glutamate, 60 mM K-glutamate, 3.5 mM DTT, 0.75 mM each amino acid except leucine, 0.63 mM leucine, 50 mM HEPES, 1.5 mM ATP and GTP, 0.9 mM CTP and UTP, 0.2 mg/ml tRNA, 0.26 mM CoA, 0.33 mM NAD, 0.75 mM cAMP, 0.068 mM folinic acid, 1 mM spermidine, 30 mM 3-PGA, 3.5% PEG-8000. When linear DNA templates were used, they were stabilized by adding 4 µM of chi6 duplex DNA to the TX-TL reaction[42]. TX-TL reactions were incubated at 29 °C for expression.

Gene expression in cell-mimics. For cell-mimics, expression reactions typically consisted of 1 µl concentrated cell-mimics in 100 mM HEPES pH 8 and TX-TL reagents for a final volume of 5 µl. Droplets of 4.5 µl of this mixture were pipetted onto a 35 mm Lumox dish (Sarstedt). The gas permeable substrate ensured homogeneous sfGFP expression in the sample. The cell-mimic droplet was covered with cover glass and sealed with a ring of vacuum grease to prevent evaporation and provide a spacer. The reaction volume was scaled up for experiments in larger samples, and was 20 µl in Fig. 2b and 35 µl for long, narrow reaction chambers in Figs. 2c and 4e. Long, narrow reaction chambers were made from two parallel 20 mm lines of vacuum grease with a gap of 2 mm, which was filled with TX-TL and cell-mimics and then sealed with cover glass.

Preparation of labeled ribosomes. *E. coli* ribosomes (New England Biolabs) were incubated in labeling buffer (50 mM HEPES pH 8.2, 100 mM KCl, 10 mM magnesium acetate) with a molar excess of Alexa Fluor 488 NHS Ester (ThermoFisher Scientific) for 90 min at room temperature. Free dye was removed by washing with labeling buffer in centrifugal filter devices with a molecular weight cut-off of 100 kDa. Each ribosome contained approximately eleven Alexa Fluor 488 labels.

Plate reader reactions. Plate reader reactions were performed in 384-well plates using a 10 µl reaction volume covered with 10 µl light mineral oil in a Tecan infinite F200 plate reader. GFP fluorescence was read every 5 min using a 485 nm ± 20 nm excitation filter and a 550 nm ± 35 nm emission filter, followed by 1 min of shaking.

Fluorescence intensity measurements were calibrated using purified sfGFP-His_6 to determine absolute concentrations.

Imaging and image analysis. Images were acquired using a spinning disk confocal microscope consisting of a Yokagawa spinning disk system (Yokagawa, Japan) built around an Axio Observer Z1 motorized inverted microscope (Carl Zeiss Microscopy GmbH, Germany) with a 20 × 1.42 NA objective. Large regions were imaged as tiles and stitched using ZEN Blue software. Further image processing and analyses were done in Fiji/ImageJ[43]. Fluorescence traces or endpoint intensities of individual hydrogel nuclei were extracted from timelapse movies by measuring fluorescence in manually selected oval regions in nuclei, using a non-fluorescent region in each cell-mimic for background subtraction. Artificial quorum sensing data was analyzed using the colony counter plugin in Fiji/ImageJ to segment and count cell-mimics and polymer beads in the stitched brightfield image of a droplet. When necessary, regions of interest were manually added or deleted. Fluorescence values of individual cell-mimics were mean fluorescence values of the individual segmented cell-mimics. For background correction, fluorescence values of segmented regions from each droplet were sorted. The lowest fluorescence intensities were from solid polymer beads, and we used this property for background correction. We removed the lowest 34% of values, which was the percentage of polymer beads in the sample, and used the highest of the removed values for background subtraction of the reduced list. Experiments (droplets with different amounts of cell-mimics) were performed for densities between 25 and 800 cell-mimics per droplet, which were binned every 50 cell-mimics. Each bin contained data from at least 156 analyzed cell-mimics.

Scanning electron microscopy was performed with a Zeiss SIGMA VP field emission scanning electron microscope using air-dried cell-mimics. To image cross sections of microcapsule polymer membranes, cell-mimics were cut using a razor blade.

Data availability

The authors declare that all relevant data supporting the findings of this study are available within the paper and its Supplementary information files. Additional data are available from the corresponding author upon request.

Received: 16 August 2018 Accepted: 1 November 2018
Published online: 28 November 2018

References

1. Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H. & Weiss, R. A synthetic multicellular system for programmed pattern formation. *Nature* **434**, 1130–1134 (2005).
2. Morsut, L. et al. Engineering customized cell sensing and response behaviors using synthetic notch receptors. *Cell* **164**, 780–791 (2016).
3. Danino, T., Mondragón-Palomino, O., Tsimring, L. & Hasty, J. A synchronized quorum of genetic clocks. *Nature* **463**, 326–330 (2010).
4. Noireaux, V. & Libchaber, A. A vesicle bioreactor as a step toward an artificial cell assembly. *Proc. Natl Acad. Sci. USA* **101**, 17669–17674 (2004).
5. Adamala, K. P., Martin-Alarcon, D. A., Guthrie-Honea, K. R. & Boyden, E. S. Engineering genetic circuit interactions within and between synthetic minimal cells. *Nat. Chem.* **9**, 431–439 (2017).
6. van Nies, P. et al. Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. *Nat. Commun.* **9**, 1583 (2018).
7. Lee, K. Y. et al. Photosynthetic artificial organelles sustain and control ATP-dependent reactions in a protocellular system. *Nat. Biotechnol.* **36**, 530–535 (2018).
8. Keber, F. C. et al. Topology and dynamics of active nematic vesicles. *Science* **345**, 1135–1139 (2014).
9. Tayar, A. M., Karzbrun, E., Noireaux, V. & Bar-Ziv, R. H. Propagating gene expression fronts in a one-dimensional coupled system of artificial cells. *Nat. Phys.* **11**, 1037–1041 (2015).
10. Tayar, A. M., Karzbrun, E., Noireaux, V. & Bar-Ziv, R. H. Synchrony and pattern formation of coupled genetic oscillators on a chip of artificial cells. *Proc. Natl Acad. Sci. USA* **114**, 11609–11614 (2017).
11. Zadorin, A. S. et al. Synthesis and materialization of a reaction-diffusion French flag pattern. *Nat. Chem.* **9**, 990–996 (2017).
12. Villar, G., Graham, A. D. & Bayley, H. A tissue-like printed material. *Science* **340**, 48–52 (2013).
13. Schwarz-Schilling, M., Aufinger, L., Mückl, A. & Simmel, F. C. Chemical communication between bacteria and cell-free gene expression systems within linear chains of emulsion droplets. *Integr. Biol.* **8**, 564–570 (2016).
14. Gines, G. et al. Microscopic agents programmed by DNA circuits. *Nat. Nanotechnol.* **12**, 351–359 (2017).
15. van Swaay, D. & deMello, A. Microfluidic methods for forming liposomes. *Lab. Chip.* **13**, 752 (2013).
16. Deshpande, S., Caspi, Y., Meijering, A. E. C. & Dekker, C. Octanol-assisted liposome assembly on chip. *Nat. Commun.* **7**, 10447 (2016).
17. Deng, N.-N., Yelleswarapu, M., Zheng, L. & Huck, W. T. S. Microfluidic assembly of monodisperse vesosomes as artificial cell models. *J. Am. Chem. Soc.* **139**, 587–590 (2017).
18. Stachowiak, J. C., Richmond, D. L., Li, T. H., Brochard-Wyart, F. & Fletcher, D. A. Inkjet formation of unilamellar lipid vesicles for cell-like encapsulation. *Lab. Chip.* **9**, 2003–2008 (2009).
19. Lentini, R. et al. Two-way chemical communication between artificial and natural cells. *ACS Cent. Sci.* **3**, 117–123 (2017).
20. Rampioni, G. et al. Synthetic cells produce a quorum sensing chemical signal perceived by *Pseudomonas aeruginosa*. *Chem. Commun.* **54**, 2090–2093 (2018).
21. Kuiper, S. M. et al. Enzymes containing porous polymersomes as nano reaction vessels for cascade reactions. *Org. Biomol. Chem.* **6**, 4315–4 (2008).
22. Sun, S. et al. Chemical signaling and functional activation in colloidosome-based protocells. *Small* **12**, 1920–1927 (2016).
23. Tang, T. Y. D. et al. Gene-mediated chemical communication in synthetic protocell communities. *ACS Synth. Biol.* **7**, 339–346 (2018).
24. Fischlechner, M. et al. Evolution of enzyme catalysts caged in biomimetic gel-shell beads. *Nat. Chem.* **6**, 791–796 (2014).
25. Basson, M. A. Signaling in cell differentiation and morphogenesis. *Cold Spring Harb. Perspect. Biol.* **4**, a008151 (2012).
26. Kim, B., Jeon, T. Y., Oh, Y.-K. & Kim, S.-H. Microfluidic production of semipermeable microcapsules by polymerization-induced phase separation. *Langmuir* **31**, 6027–6034 (2015).
27. Loiseau, E. et al. Strong microcapsules with permeable porous shells made through phase separation in double emulsions. *Langmuir* **33**, 2402–2410 (2017).
28. Yang, D. et al. Enhanced transcription and translation in clay hydrogel and implications for early life evolution. *Sci. Rep.* **3**, 3165 (2013).
29. Yu, W. H. et al. Adsorption of proteins and nucleic acids on clay minerals and their interactions: a review. *Appl. Clay Sci.* **80-81**, 443–452 (2013).
30. Franchi, M. et al. Clay-nucleic acid complexes: characteristics and implications for the preservation of genetic material in primeval habitats. *Orig. Life. Evol. Biosph.* **29**, 297–315 (1999).
31. Ng, W.-L. & Bassler, B. L. Bacterial Quorum-sensing network architectures. *Annu. Rev. Genet.* **43**, 197–222 (2009).
32. Peters, R. J. R. W. et al. Cascade reactions in multicompartmentalized polymersomes. *Angew. Chem. Int. Ed.* **53**, 146–150 (2013).
33. Deng, N.-N. & Huck, W. T. S. Microfluidic formation of monodisperse coacervate organelles in liposomes. *Angew. Chem. Int. Ed.* **56**, 9736–9740 (2017).
34. Hanczyc, M. M., Fujikawa, S. M. & Szostak, J. W. Experimental models of primitive cellular compartments: encapsulation, growth, and division. *Science* **302**, 618–622 (2003).
35. Johnston, D. S. & Nüsslein-Volhard, C. The origin of pattern and polarity in the Drosophila embryo. *Cell* **68**, 201–219 (1992).
36. Coppey, M., Boettiger, A. N., Berezhkovskii, A. M. & Shvartsman, S. Y. Nuclear trapping shapes the terminal gradient in the Drosophila embryo. *Curr. Biol.* **18**, 915–919 (2008).
37. Klein, P., Pawson, T. & Tyers, M. Mathematical modeling suggests cooperative interactions between a disordered polyvalent ligand and a single receptor site. *Curr. Biol.* **13**, 1669–1678 (2003).
38. Mammen, M., Choi, S. K. & Whitesides, G. M. Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors. *Angew. Chem. Int. Ed.* **37**, 2754–2794 (1998).
39. Whitty, A. Cooperativity and biological complexity. *Nat. Chem. Biol.* **4**, 435–439 (2008).
40. Caschera, F. & Noireaux, V. Compartmentalization of an all- *E. coli* cell-free expression system for the construction of a minimal cell. *Artif. Life* **22**,185–195 (2016).
41. Didovyk, A., Tonooka, T., Tsimring, L. & Hasty, J. Rapid and scalable preparation of bacterial lysates for cell-free gene expression. *ACS Synth. Biol.* **6**, 2198–2208 (2017).
42. Marshall, R., Maxwell, C. S., Collins, S. P., Beisel, C. L. & Noireaux, V. Short DNA containing χ sites enhances DNA stability and gene expression in *E. coli* cell-free transcription-translation systems. *Biotechnol. Bioeng.* **114**, 2137–2141 (2017).
43. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Meth.* **9**, 676–682 (2012).

Acknowledgements

This work was supported by the Department of Defense (Army Research Office) through the Multidisciplinary University Research Initiative (MURI), award W911NF-13-1-0383. H. N. was supported by Early and Advanced Postdoc.Mobility fellowships from the Swiss National Science Foundation (P2ELP3_161844, P300PA_174346). We thank Lev Tsimring, Partho Ghosh and Andrew Rudd for helpful comments on the manuscript, Ahanjit Bhattacharya for labeled ribosomes, Andriy Didovyk for *E. coli* strain BL21-Gold (DE3)/pAD-LyseR, and Prof. Jeff Hasty and Ryan Johnson for their collaboration in

microfluidic chip fabrication. We thank the UCSD School of Medicine Microscopy Core with grant NS047101 and Jennifer Santini for assistance with scanning electron microscopy, and the Waitt Advanced Biophotonics Core Facility of the Salk Institute with funding from NIH-NCI CCSG: P30 014195, NINDS Neuroscience Core Grant: NS072031 and the Waitt Foundation.

Author contributions

H.N. and N.K.D. conceived the study and designed experiments. H.N. and C.C. performed experiments. H.N. analyzed the data. H.N. and N.K.D wrote the manuscript. All authors read and accepted the manuscript.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-07473-7.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature COMMUNICATIONS

ARTICLE

DOI: 10.1038/s41467-018-03469-5 OPEN

Cell-free protein synthesis from genomically recoded bacteria enables multisite incorporation of noncanonical amino acids

Rey W. Martin[1,2,3], Benjamin J. Des Soye[2,3,4], Yong-Chan Kwon[1,2,3,5], Jennifer Kay[1,2,3], Roderick G. Davis[6], Paul M. Thomas[6,7,8], Natalia I. Majewska[1], Cindy X. Chen[1], Ryan D. Marcum[3,4], Mary Grace Weiss[1], Ashleigh E. Stoddart[1], Miriam Amiram[9,10], Arnaz K. Ranji Charna[1,2,3], Jaymin R. Patel[9,10], Farren J. Isaacs[9,10], Neil L. Kelleher[4,6,7,8,11], Seok Hoon Hong[12] & Michael C. Jewett[1,2,3,4,8,13]

Cell-free protein synthesis has emerged as a powerful approach for expanding the range of genetically encoded chemistry into proteins. Unfortunately, efforts to site-specifically incorporate multiple non-canonical amino acids into proteins using crude extract-based cell-free systems have been limited by release factor 1 competition. Here we address this limitation by establishing a bacterial cell-free protein synthesis platform based on genomically recoded *Escherichia coli* lacking release factor 1. This platform was developed by exploiting multiplex genome engineering to enhance extract performance by functionally inactivating negative effectors. Our most productive cell extracts enabled synthesis of 1,780 ± 30 mg/L superfolder green fluorescent protein. Using an optimized platform, we demonstrated the ability to introduce 40 identical *p*-acetyl-L-phenylalanine residues site specifically into an elastin-like polypeptide with high accuracy of incorporation (≥98%) and yield (96 ± 3 mg/L). We expect this cell-free platform to facilitate fundamental understanding and enable manufacturing paradigms for proteins with new and diverse chemistries.

[1] Department of Chemical and Biological Engineering, Northwestern University, Evanston, Illinois 60208, USA. [2] Chemistry of Life Processes Institute, Northwestern University, Evanston, Illinois 60208, USA. [3] Center for Synthetic Biology, Northwestern University, Evanston, Illinois 60208, USA. [4] Interdisciplinary Biological Sciences Program, Northwestern University, Evanston, Illinois 60208, USA. [5] Department of Biological and Agricultural Engineering, Louisiana State University, Baton Rouge, Louisiana 70803, USA. [6] Proteomics Center of Excellence, Northwestern University, Evanston, Illinois 60208, USA. [7] Department of Molecular Biosciences, Northwestern University, Evanston, Illinois 60208, USA. [8] Robert H. Lurie Comprehensive Cancer Center, Northwestern University, Chicago, Illinois 60611, USA. [9] Department of Molecular Cellular, and Developmental Biology, Yale University, New Haven, Connecticut 06520, USA. [10] Systems Biology Institute, Yale University, New Haven, Connecticut 06516, USA. [11] Department of Chemistry, Northwestern University, Evanston, Illinois 60208, USA. [12] Department of Chemical and Biological Engineering, Illinois Institute of Technology, Chicago, Illinois 60616, USA. [13] Simpson Querrey Institute, Northwestern University, Chicago, Illinois 60611, USA. Correspondence and requests for materials should be addressed to M.C.J. (email: m-jewett@northwestern.edu)

Cell-free synthetic biology is emerging as a transformative approach to understand, harness, and expand the capabilities of natural biological systems[1]. The foundational principle is that complex biomolecular transformations are conducted without using intact cells. Instead, crude cell lysates (or extracts) are used, which provides a unique freedom of design to control biological systems for a wide array of applications. For example, cell-free protein synthesis (CFPS) systems have been used to decipher the genetic code[2], prototype genetic circuits and metabolic pathways[3–7], enable portable diagnostics[8], facilitate on-demand biomolecular manufacturing[9,10], and produce antibody therapeutics at the commercial scale[11]. The recent surge of applications has revitalized interest in cell-free systems, especially in areas where limits imposed by the organism may impede progress. One such area is expanding the genetic code to incorporate non-canonical amino acids (ncAAs) into proteins, where the extent of engineering can be limited by the fitness of the organism[12–15].

Pioneering efforts by Schultz and others have demonstrated it is possible to genetically encode more than 150 ncAAs into proteins, and that this encoding can be a powerful tool[15,16]. For example, site-specific incorporation of ncAAs at single positions in proteins have provided new ways to study protein structure, dynamics, and posttranslational modifications[17], as well as manufacture protein–drug conjugates[18,19]. However, inefficiencies associated with the engineered orthogonal translation (TL) machinery (e.g., TL elements that specifically use a ncAA and do not interact with the cell's natural TL apparatus) have limited the ability to incorporate multiple ncAAs into proteins with high purity and yields[20,21]. A key constraint is that codon reassignment strategies typically rely on amber suppression[22], where the amber UAG stop codon is re-assigned to encode a ncAA and the orthogonal transfer RNA anticodon is mutated to CUA. The orthogonal ncAA-tRNA$_{CUA}$ must then outcompete essential TL machinery (e.g., release factor 1, RF1) for the UAG codon. Historically, this competition has led to poor protein expression yields, as premature termination by RF1 exponentially increases with the number of amber codons in the coding sequence[23]. Poor protein expression yields limit applications in both basic and applied science.

Recently, a genomically recoded *Escherichia coli* strain was developed (*C321.ΔA*) in which all 321 occurrences of the UAG stop codon were reassigned to the synonymous UAA codon[24] using multiplex automated genome engineering (MAGE)[25] and conjugative assembly genome engineering (CAGE)[26]. This allowed for the genomic deletion of RF1 (i.e., *ΔprfA* or *ΔA*) without affecting cellular physiology, thus freeing the UAG codon for dedicated ncAA incorporation[24]. Precursor RF1-deficient strains in which only a small set of essential genes were recoded have already shown the potential to produce proteins with improved ncAA incorporation efficiencies as compared with strains with RF1[24,27]; however, the upregulation of natural suppression mechanisms (e.g., *ssrA*) is problematic, because they promote the formation of truncation products, especially for tens of incorporation events[24,27,28]. The fully recoded *C321.ΔA* strain avoids these problems and we recently showed the possibility of using *C321.ΔA* coupled with extensively engineered synthetases for multi-site incorporation of up to 30 ncAAs into a single biopolymer in vivo[20]. Based on these results, we hypothesized that the fully recoded *C321.ΔA* strain would serve as an ideal chassis strain for the development of crude extract-based cell-free systems capable of highly efficient, multi-site ncAA incorporation into biopolymers. Such a system would complement in vivo manufacturing strategies, with some advantageous features[29,30]. For example, the open reaction environment means the supply of orthogonal TL system (OTS) components and their substrates necessary for high-level ncAA incorporation can be provided and controlled at precise ratios as a way to overcome enzyme inefficiencies. In addition, cell-free systems are not limited by viability requirements, thus avoiding constraints arising from toxic OTS components[27]. Our proposed approach based on the *C321.ΔA* strain might also provide cost and ease of use advantages over other cell-free systems that have tried to reduce the effects of RF1 competition by using reconstituted systems[31], antibody inhibitors[32], RF1 depletion by subtractive chromatography[33], or partially recoded *E. coli* strains with elevated natural suppression mechanisms[34].

Here we describe the development of a CFPS platform from the genomically recoded *C321.ΔA* to manufacture proteins with tens of identical site specifically introduced ncAAs. Specifically, we use MAGE to improve protein production capacity by inactivating negative effectors in the host strain such that they are not present in the lysate. By testing tens of strain variants, we isolate a CFPS platform capable of synthesizing 1,780 ± 30 mg/L of superfolder green fluorescent protein (sfGFP), as well as modified sfGFP containing up to five *p*-acetyl-L-phenylalanine (pAcF) residues at high purity (≥ 98%). Using an optimized CFPS platform, we test the ability to synthesize elastin-like polypeptides (ELPs) that

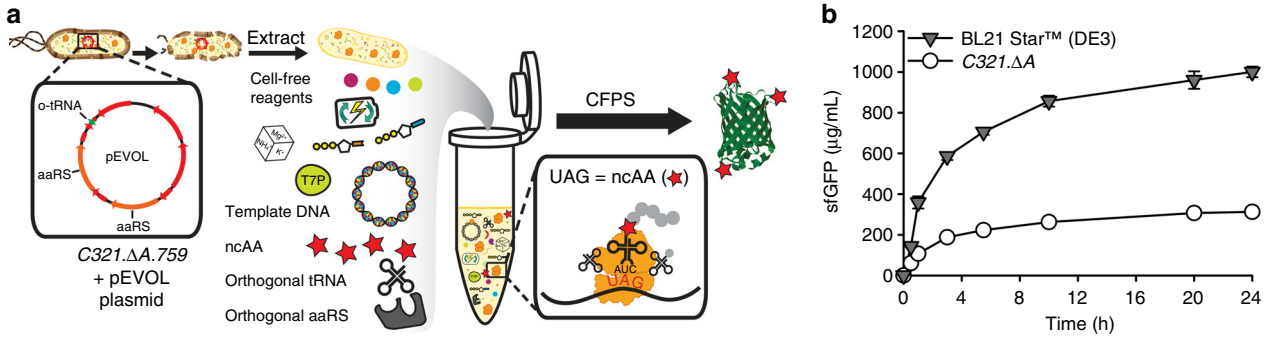


Fig. 1 CFPS from extracts of a genomically recoded organism. **a** Schematic of the production and utilization of crude extract from genomically recoded organisms with plasmid overexpression of orthogonal translation components for cell-free protein synthesis (CFPS). CFPS reactions are supplemented with the necessary substrates (e.g., amino acids, NTPs, etc.) required for in vitro transcription and translation as well as purified orthogonal translation system (OTS) components to help increase the ncAA incorporation efficiency. aaRS, aminoacyl tRNA synthetase; ncAA, non-canonical amino acid; T7P, T7 RNA polymerase; UAG, amber codon. **b** Time course of superfolder green fluorescent protein (sfGFP) synthesis catalyzed by extracts derived from a genomically recoded organism, *C321.ΔA*, and a commercial strain, BL21 Star (DE3). Three independent batch CFPS reactions ($n = 3$) were performed at 30 °C for each time point over 24 h. Error bar = 1 SD

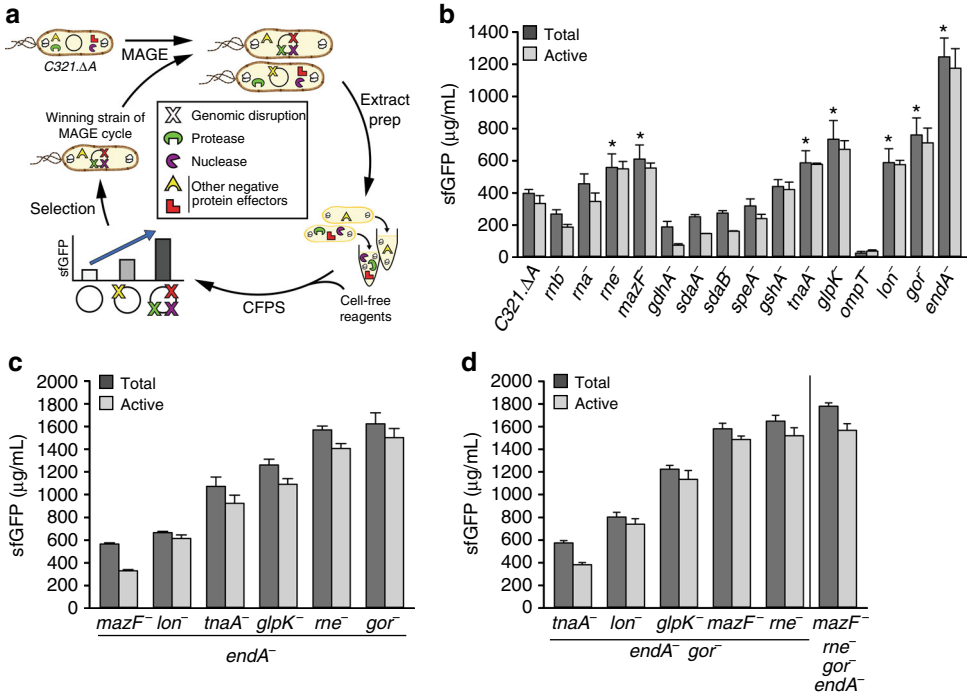


Fig. 2 Engineering *C321.ΔA* variants for enhanced CFPS. **a** Schematic of design-build-test cycles employing multiplex automated genome engineering (MAGE) to disrupt putative negative protein effectors (Supplementary Table 1) in engineered *C321.ΔA* strains for producing extracts with enhanced cell-free protein synthesis (CFPS) yields. **b** Cell extracts derived from *C321.ΔA* and genomically engineered strains containing a single putative negative effector inactivation were screened for sfGFP yields. Beneficial mutations that increase active yields ≥ 50% relative to C321.ΔA are highlighted with an *($p < 0.01$, Student's *t*-test). **c** *C321.ΔA.542* (*endA*$^-$) was chosen as the next base strain and the following beneficial disruptions were pursued in combination: *rne*, *mazF*, *tnaA*, *glpK*, *lon*, and *gor*. **d** *C321.ΔA.709* (*endA*$^-$ *gor*$^-$) was selected as the subsequent base strain for triple and quadruple mutant construction. *C321.ΔA.759* (*endA*$^-$ *gor*$^-$ *rne*$^-$ *mazF*$^-$) yielded the highest level of CFPS production. Total sfGFP concentration was measured by counting radioactive ^{14}C-Leucine incorporation and active protein was measured using fluorescence. Three independent batch CFPS reactions were performed for each sample at 30 °C for 20 h ($n = 3$). Error bar = 1 SD

contain up to 40 UAG codons. We demonstrate incorporation of 40 ncAAs per ELP protein with high yields (~ 100 mg/L) and high fidelity (≥ 98%) of site-specific ncAA incorporation.

Results

CFPS from extracts of a genomically recoded organism. To benchmark CFPS activity, we first compared sfGFP yields in extracts from *C321.ΔA* and BL21 Star (DE3), the standard commercial protein expression strain (Fig. 1a). Combined transcription (TX)–TL reactions were carried out in 15 μL volumes for 24 h at 30 °C. Protein yields from BL21 Star (DE3) extracts were > 3-fold higher than those from *C321.ΔA* (Fig. 1b), highlighting the need to improve protein synthesis yields to take advantage of the benefits of RF1 removal for making modified proteins with ncAAs for preparative purposes.

Previously, genomic modifications to the extract source strain to stabilize DNA template[35], amino acid supply[36], and protein degradation[37] have improved CFPS yields from other source strains. For example, we engineered a partially recoded strain of *E. coli* (*rEc.E13.ΔA*) by disrupting genes encoding nucleases (MCJ.559 (*endA*$^-$ *csdA*$^-$)) to improve protein synthesis yields > 4-fold relative to the parent strain[34]. Building on this knowledge, we hypothesized that the genomic disruption of negative protein effectors in *C321.ΔA* extracts would help stabilize essential substrates in cell-free reactions, extend reaction durations, and increase CFPS yields.

Strain engineering for improved CFPS performance. We targeted the functional inactivation of five nucleases (*rna*, *rnb*, *mazF*, *endA*, and *rne*), two proteases (*ompT* and *lon*), and eight targets shown previously to negatively impact amino acid, energy, and redox stability (*gdhA*, *gshA*, *sdaA*, *sdaB*, *speA*, *tnaA*, *glpK*, and *gor*) in *C321.ΔA* individually and in combination (Supplementary Table 1). Our effort followed a five-step approach. First, we generated a library of single mutant strains in which we used MAGE to insert an early TL termination sequence into the open reading frames of gene targets that would functionally inactivate them, as we have done before[34] (Fig. 2a and Supplementary Tables 2 and 3). Second, we confirmed gene disruptions using multiplex allele specific PCR and DNA sequencing. Third, we measured the growth rate for each of the MAGE-modified strains, noting that average doubling time increased 9 ± 9% above the parent strain (Supplementary Table 4). Fourth, cell extracts from each strain were generated using a high-throughput and robust extract generation procedure[38]. Fifth, we tested the strains in CFPS to assess their overall protein synthesis capability. We observed that seven single functional inactivation mutations increased CFPS yields more than 50% relative to the wild type *C321.ΔA* strain; namely, *rne*$^-$, *mazF*$^-$, *tnaA*$^-$, *glpK*$^-$, *lon*$^-$, *gor*$^-$, and *endA*$^-$ (Fig. 2b). These results suggested that some of the protein effectors targeted for inactivation were deleterious to CFPS activity. They also demonstrated the difficulty associated with predicting CFPS productivities from engineered strains. For example, some mutations identified in previous screens (e.g., *rnb*$^-$ in *rEc.E13.ΔA*)[34] were not beneficial in the *C321.ΔA* context, others which reduced cellular fitness enhanced CFPS activity (e.g., *lon*$^-$), and yet others with no impact on cell growth (e.g., *ompT*$^-$) led to poor extract performance (Fig. 2b).

With improvements in hand from single mutant strains, we next set out to identify synergistic benefits to CFPS productivity

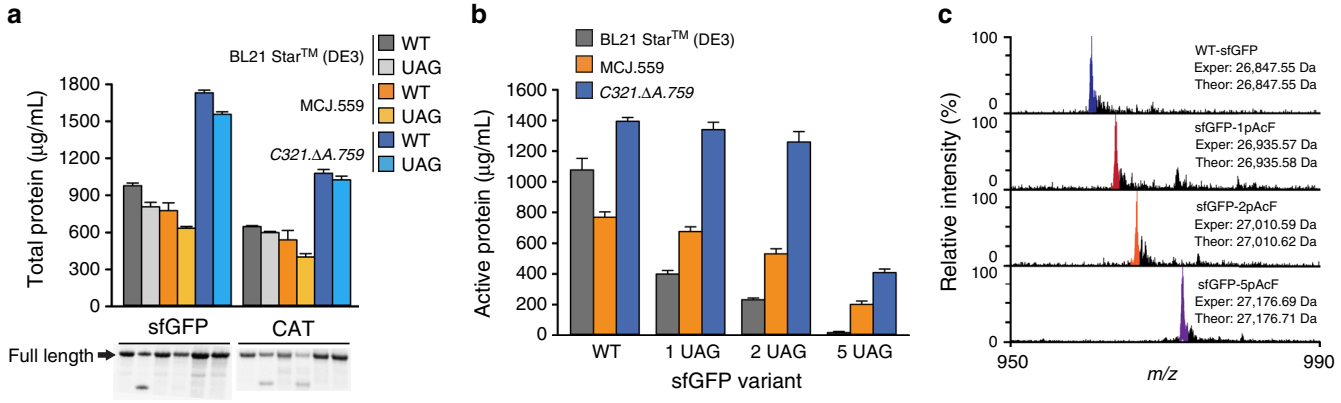


Fig. 3 Multi-site incorporation of pAcF into proteins. Cell-free *p*-acetyl-L-phenylalanine (pAcF) incorporation was compared using extracts derived from BL21 Star (DE3), MCJ.559, and *C321.ΔA.759* strains containing the pEVOL-pAcF vector. The pEVOL-pAcF vector harbors the orthogonal translation machinery necessary for pAcF incorporation. **a** Total protein yields for wild-type (WT) and 1 UAG versions of superfolder green fluorescent protein (sfGFP) and chloramphenicol acetyl transferase (CAT) are shown along with an autoradiogram of the resulting protein product. Supplementary Fig. 8 shows the entirety of the autoradiogram along with a molecular weight marker. **b** Multi-site incorporation of pAcF into sfGFP variants as quantified by active protein produced. The sfGFP variants used were wild-type (WT), sfGFP containing a single pAcF corresponding to the position of T216 (1 UAG), sfGFP containing sfGFP containing two pAcFs (2 UAG), and sfGFP containing five pAcFs (5 UAG). Three independent batch CFPS reactions were performed for each sample at 30 °C for 20 h ($n = 3$). Error bar = 1 SD. **c** Spectrum of the 28 + charge state of sfGFP, obtained by top-down mass spectrometry and illustrating site-specific incorporation of pAcF at single and multiple sites. Experimental (Exper) and theoretical (Theor) mass peaks for each sfGFP variant are shown. Major peaks (color) in each spectrum coincide with the theoretical peaks for each species (see also Supplementary Fig. 11). Smaller peaks immediately to the right of the major peaks are due to oxidation of the protein, a common electrochemical reaction occurring during electrospray ionization. Experimentally determined masses are ≤1 p.p.m. in comparison of theoretical mass calculations. Owing to the size of pAcF, misincorporation would result in peaks present at lower m/z values relative to the colored theoretical peak

by combining highly productive mutations. We introduced the *rne*$^-$, *mazF*$^-$, *tnaA*$^-$, *glpK*$^-$, *lon*$^-$, and *gor*$^-$ mutations to the best performing strain from our initial screen, strain *C321.ΔA.542* (*endA*$^-$) (Fig. 2c). The combination of *endA*$^-$ and *gor*$^-$ mutations resulted in an extract capable of synthesizing 1,620 ± 10 mg/L of active sfGFP (strain *C321.ΔA.709*). We then used *C321.ΔA.709* to generate six additional strains with combined mutations. Although we did not observe synergistic enhancements, our top performing extract chassis strain (*C321.ΔA.759* (*endA*$^-$ *gor*$^-$ *rne*$^-$ *mazF*$^-$)) resulted in total yields of 1,780 ± 30 mg/L (Fig. 2d), representing a 4.5-fold increase in sfGFP yield relative to the progenitor strain (*C321.ΔA*). In addition, we tested 12 combinatorial mutants generated throughout our MAGE screening, and although a few demonstrated CFPS yields > 1 g/L of active sfGFP, none surpassed the CFPS yields observed from *C321.ΔA.759* (Supplementary Table 5). Lastly, we determined that CFPS improvements seen in *C321.ΔA.759* brought on by genomic modifications could not be obtained by simply supplementing *C321.ΔA*-based reactions with RNAse inhibitors (Supplementary Fig. 1). Final strains were fully sequenced to verify functional targeted modifications in the genome. Whole-genome sequences for strains *C321.ΔA*, *C321.ΔA.542*, *C321.ΔA.705*, *C321.ΔA.709*, *C321.ΔA.740*, and *C321.ΔA.759* have been deposited in the NCBI SRA collection under accession code PRJNA361365. Each of the targeted mutations were achieved. MAGE has been shown to induce mutations throughout the genome before, and we observed a number of accumulated polymorphisms in the extract chassis strains. These polymorphisms, along with a specific list of protein-coding genes bearing mutations, are shown in Supplementary Tables 6 and 7. In the future, we seek to better understand the systems impact of the non-targeted mutations.

Based on our previous studies using *rEc.E13.ΔA*[34], we hypothesized that the beneficial mutations in *C321.ΔA.759* reduced messenger RNA degradation and stabilized the DNA template. To test mRNA stability, we performed TL-only reactions using extracts derived from *C321.ΔA.759* and *C321.ΔA*. Purified mRNA template coding for sfGFP was used to direct protein synthesis. We observed a twofold increase in mRNA and ~90% increase of active sfGFP using *C321.ΔA.759* extracts relative to *C321.ΔA* extracts after a 120 min cell-free reaction (Supplementary Fig. 2). To test DNA stability, TX-only reactions were used. Specifically, plasmid DNA containing the modified red fluorescent protein–Spinach aptamer gene (Supplementary Table 3) was pre-incubated with cell extract and a fluorophore molecule, 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI), for 0, 60, and 180 min. Then, CFPS reagents were added and mRNA was synthesized, then quantified by measuring the fluorescence of DFHBI-bound Spinach aptamer mRNA. After 180 min of pre-incubation, nearly 50% of Spinach aptamer mRNA was synthesized in *C321.ΔA.759* (*endA*$^-$) extracts relative to the 0 min control. In contrast, the extract with endonuclease I (*C321.ΔA*) decreased the maximum mRNA synthesis level by ~75% (Supplementary Fig. 3). Together, our data support the hypothesis that inactivating nucleases in the extract chassis strain stabilized DNA and mRNA to improve CFPS yields.

In addition to confirming added DNA and mRNA stability, we also assessed potential changes in energy and amino acid substrate stability that may have occurred in *C321.ΔA.759*–relative to *C321.ΔA*–based CFPS. Similar trends in ATP levels (Supplementary Fig. 4), adenylate charge (Supplementary Fig. 5), and amino acid concentrations (Supplementary Fig. 6) were observed in CFPS reactions derived from both strains. Supplemental feeding with the amino acids found to be most rapidly depleted did not improve yields (Supplementary Fig. 6f). The similar amino acid and energy stability profiles in *C321.ΔA.759* compared with *C321.ΔA* suggest that our strain engineering efforts did not modulate the availability of these substrates.

To generalize CFPS improvements in *C321.ΔA.759*, we next expressed four model proteins that have been previously synthesized in CFPS systems and compared productivities to BL21 Star (DE3). We observed a 31–63% increase in soluble and total protein synthesis of sfGFP, chloramphenicol

acetyltransferase (CAT), dihydrofolate reductase (DHFR), and modified murine granulocyte-macrophage colony-stimulating factor (mGM-CSF) in our engineered *C321.ΔA.759* extracts as compared to BL21 Star (DE3) extracts (Supplementary Fig. 7a). Autoradiograms of proteins produced using *C321.ΔA.759* extract show production of full-length sfGFP, CAT, DHFR, and mGM-CSF (Supplementary Fig. 7b and 7c). In addition, we observed disulfide bond formation in the model mGM-CSF under an oxidizing CFPS environment (– DTT), as has been previously shown (Supplementary Fig. 7c)[39,40]. In sum, the development of enhanced extract source strains by MAGE enabled a general and high-yielding CFPS platform.

Multi-site ncAA incorporation into proteins in CFPS. We next aimed to test site-specific ncAA incorporation into proteins using our high-yielding CFPS platform from *C321.ΔA.759*-derived extracts and compare these results to reactions using extracts from BL21 Star (DE3) (containing RF1) and a partially recoded RF1-deficient engineered strain MCJ.559 based on *rEc.E13.ΔA*. To do so, we transformed each organism with pEVOL-pAcF plasmid that expresses both orthogonal pAcF synthetase (pAcFRS) and tRNA (o-tRNA^{opt})[41]. Then, we quantitatively assessed the incorporation of pAcF into sfGFP variants with up to five in-frame amber codons. CFPS reactions were supplemented with additional OTS components based on our previous work[27]. Specifically, we added 10 μg/mL of linear DNA encoding optimized orthogonal tRNA in the form of a transzyme (o-tRNA^{opt}) for in situ synthesis of the tRNA. The orthogonal pAcFRS was overproduced, purified as previously described, and added at a level of 0.5 mg/mL. The ncAA, in this case pAcF, was supplied at a level of 2 mM in each CFPS reaction. Total protein yields were quantified by ^{14}C-leucine radioactive incorporation. Production of wild-type and modified sfGFP containing one UAG codon (sfGFP-UAG) was increased 77% and 92% in *C321.ΔA.759* extracts as compared with BL21 Star (DE3), and 120% and 145% as compared with MCJ.559, respectively (Fig. 3a and Supplementary Fig. 8a). Moreover, we observed that sfGFP-UAG was expressed at 90% the level of wild-type sfGFP. Owing to the absence of RF1 competition, the major protein produced was full-length sfGFP using extracts derived from *C321.ΔA.759* and MCJ.559, whereas truncated sfGFP was visible in reactions catalyzed by BL21 Star (DE3) extract, presumably due to RF1 competition (Fig. 3a and Supplementary Fig. 8a)[30,42]. Similar results were obtained with a second model protein, CAT with an in-frame amber codon at position 112 (CAT-UAG) (Fig. 3a and Supplementary Fig. 8b). When expressing CAT-UAG using MCJ.559 extract, similar levels of truncated CAT relative to BL21 Star (DE3) were observed; however, this is most likely due to an upregulation of rescue mechanisms for ribosome stalling in the partially recoded strain[34]. Single pAcF incorporation into CAT-UAG using *C321.ΔA.759* lysate demonstrated only full-length product. Therefore, our completely recoded, genomically engineered *C321.ΔA.759* strain provides benefits for efficient ncAA incorporation without detectable levels of truncation product.

We then evaluated the ability of our high-yielding CFPS platform to facilitate incorporation of up to five identical ncAAs into sfGFP. For ease of analysis, a fluorescence assay was used, which indicated increased production of sfGFP in extracts from *C321.ΔA.759* (Fig. 3b). Results for BL21 Star (DE3) extract displayed an exponential decrease in active sfGFP synthesized with an increasing presence of UAG, leading to the production of no detectable active protein for sfGFP-5UAG. Active protein produced by *C321.ΔA.759* extract were ~ 2-fold greater than that produced by MCJ.559 extract, suggesting that benefits observed in increased yield can be extended to multi-site ncAA incorporation for our enhanced, fully recoded strain. Furthermore, we examined the ability to incorporate consecutive pAcFs into single protein. Protein gel and autoradiogram analysis of sfGFP with eight and nine consecutive amber codons indicated that this is possible, with the percent of full-length product being ~ 75% and 60%, respectively (Supplementary Fig. 9).

In addition, batch reactions catalyzed by *C321.ΔA.759* extracts could also be scaled 17-fold without loss of productivity provided that a proper ratio of surface area to volume ratio is maintained (Supplementary Fig. 10)[43]. Of note, we believe our reactions could be further scaled to a wide range of volumes to produce larger amounts of protein if accounting for surface area to volume effects. For example, Sutro Biopharma has applied *E. coli*-based CFPS platforms to clinical manufacturing of therapeutics at the 100 L scale[44], with an expansion factor of 10^6. In terms of cost, although we use a phosphoenolpyruvate (PEP)-based CFPS system here, cellular metabolism could be used to fuel cost effective, high-level protein synthesis suitable for manufacturing applications[45,46].

After demonstrating benefits for protein expression, we carried out top-down mass spectrometry (i.e., MS analysis of whole intact proteins) to detect and provide semi-quantitative data for the incorporation efficiency of pAcF into sfGFP using extract derived from *C321.ΔA.759*. Figure 3c shows the 28 + charge state of sfGFP and clearly illustrates mass shifts corresponding to the incorporation of one, two, and five pAcF residues. Site-specific incorporation of pAcF, as detected by MS, was ≥ 98% in all samples, with ≤ 1 p.p.m. difference between experimental and theoretical protein masses (Supplementary Fig. 11). In other words, efficient and high yielding site-specific pAcF incorporation into sfGFP was observed when using *C321.ΔA.759* extract. We went on to further show that extracts generated from *C321.ΔA.759* are compatible with multiple OTSs, showing the incorporation of *p*-propargyloxy-L-phenylalanine and *p*-azido-L-phenylalanine (pAzF) (Supplementary Fig. 12).

Multi-site ncAA incorporation into ELPs. We next explored the synthesis of sequence-defined biopolymers containing tens of site specifically introduced ncAAs using our efficient and tunable CFPS system. As a model biopolymer, we selected ELPs. ELPs are biocompatible and stimuli-responsive biopolymers that can be applied for drug delivery and tissue engineering[47,48]. Typically, ELPs consist of repeats of the pentapeptide sequence VPGVG, which is known to be a key component in elastin and exhibits interesting self-assembly behavior (random coil to helix) above its transition temperature. The structure and function of elastin is maintained as long as the glycine and proline residues are present; however, the second valine residue is permissive for any amino acid except proline and is therefore also permissive to ncAAs[20]. Previously, ncAAs have been introduced into ELPs by substituting natural amino acids with structurally similar ncAAs in CFPS systems[49]. Conticello and colleagues[50] have also previously produced imperfect ELPs containing up to 22 ncAAs in vivo using an *E. coli* strain with an attenuated activity of RF1. We previously incorporated up to 30 ncAAs into ELPs by evolving orthogonal synthetases in vivo with enhanced specificities[20]. In this study, we constructed and tested in CFPS three ELP constructs containing 20, 30, and 40 UAG codons, as well as control proteins with tyrosine codons substituted for UAGs.

Before characterizing ELP yields, we first carried out a series of optimization experiments to enhance CFPS yields of sfGFP with 5 UAG codons, as expression yields for this construct were reduced in our initial studies (Fig. 3b). By testing total and soluble protein yields, we determined that the reduction in yield was a result of loss in sfGFP solubility and activity (Supplementary Fig. 13).

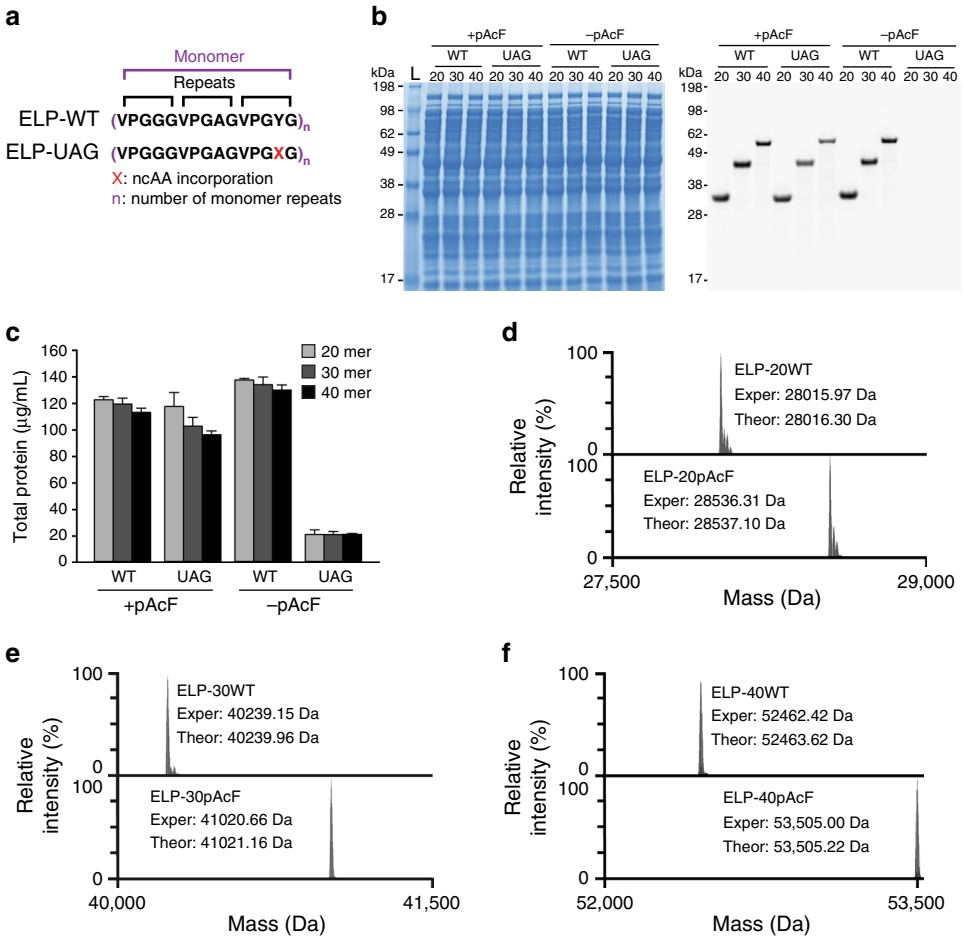


Fig. 4 Multi-site ncAA incorporation at high yield and purity. **a** Schematic of the protein sequences for wild-type ELPs containing three pentapeptide repeats per monomer unit (ELP-WT) and ELPs containing 1 ncAA per monomer unit (ELP-UAG). **b** SDS-PAGE and autoradiogram analysis of cell-free produced ELP-WT and ELP-UAG 20-, 30-, and 40-mers in the presence (+) and absence (-) of *p*-acetyl-L-phenylalanine (pAcF). Numbers next to the molecular weight ladder (L) represent the approximate kilodalton (kDa) size of the band. **c** Total protein yields of cell-free synthesized ELPs (20-, 30, and 40-mers) after incubation at 30 °C for 20 h are shown. Three independent batch CFPS reactions were performed for each sample ($n = 3$). Error bar = 1 SD. **d**–**f** Deconvoluted mass spectra of ELPs obtained by top-down mass spectrometry illustrate complete, site-specific incorporation of pAcF at **d** 20, **e** 30, and **f** 40 sites. Deconvoluted average masses for the major peaks in each spectrum (Exper) match the theoretical average mass (Theor) for each species within 1.2 Da. Smaller peaks next to the major peaks arise from minor oxidation of the protein during electrospray ionization

However, a 31% increase in sfGFP-5UAG production was observed upon increasing pAcFRS levels 2-fold, pAcF levels 2.5-fold, and o-tz-tRNA^{opt} 3-fold (Supplementary Fig. 14). Upon application of these optimized conditions, called OTS^{opt}, to the synthesis of ELP-UAGs containing 20, 30, and 40-mers, total yields increased by 40%, 33%, and 26%, respectively, as compared with supplementing with OTS levels optimized for 1 ncAA incorporation (Supplementary Fig. 15). ELP-UAG products were visualized using an autoradiogram, which demonstrated the high percentage of full-length protein and whose band intensities corroborate total yields measured (Supplementary Fig. 15b).

We next applied OTS^{opt} to the synthesis of ELP-UAGs with 20, 30, and 40-mers in the presence and absence of pAcF to demonstrate specificity of incorporation. ELP-UAGs were only synthesized in the presence of pAcF without any clear indication of truncation products, whereas no protein was observed in the absence of pAcF (Fig. 4). We anticipated that yields would decrease as the number of UAG codons increased due to the higher demand of pAcF-charged o-tRNA. In contrast, near wild type yields of ~ 100 mg/L were obtained for all UAG constructs. We then carefully examined the efficiency of multi-site ncAA incorporation using top-down liquid chromatography (LC)-MS of intact ELPs. LC-MS analysis showed ≥ 98% site-specific pAcF incorporation in ELP-UAG constructs of 20, 30, and 40-mers (Fig. 4).

Discussion

We present a crude extract-based CFPS platform based on the fully recoded *C321.ΔA* strain that is capable of high-level protein expression. This platform was generated using MAGE to create libraries of improved extract chassis strains by targeting the functional inactivation of multiple negative effectors. A combinatorial disruption of the genes *endA*$^{-}$, *gor*$^{-}$, *rne*$^{-}$, and *mazF*$^{-}$ (*C321.ΔA.759*) increased total CFPS yields from 397 ± 24 mg/L in the parent strain to 1,780 ± 30 mg/L of sfGFP, which is the highest reported protein yield from RF1-deficient extracts. These improvements translated to the enhanced yield of proteins harboring site specifically introduced ncAAs.

By optimizing the cell-free environment for multiple-identical ncAA incorporation, we were able to achieve multi-site ncAA incorporation into multiple model proteins with high yields (~ 99% wild-type sfGFP expression yields for up to 2 ncAAs, ~ 95% wild-type ELP expression yields for up to 20 ncAAs, and ~ 85%

wild-type ELP expression yields for up to 30 and 40 ncAAs) and purity (≥98% accuracy of ncAA incorporation) due to the absence of RF1 in our system. To our knowledge, these are the purest polymers with this many site-specifically introduced ncAAs (i.e., 40) synthesized to date. This exceeds our previous effort in cells that could synthesize ELP constructs with 30 UAG codons with 71% of the proteins having the desired 30 pAcF residues[20]. As such, our approach opens opportunities to site-specifically modify the dominant physical and biophysical properties of biopolymers. This will allow researchers to go beyond tag-and-modify approaches that have historically been the focus of ncAA incorporation efforts, because the field was previously limited to only one or a few instances of site-specific incorporation. Notably, our protein expression yields of ~1,700 mg/L and 99% suppression efficiency for sfGFP with 2 ncAAs outperform the best expression of proteins with single or multiple ncAAs in vivo, to our knowledge (Fig. 2d and Supplementary Fig. 13). For example, previous in vivo experiments using an RF1 knockout strain demonstrated the synthesis of enhanced GFP containing one, two, and three pAcFs at 3.5, 3.5, and 5.4 mg/L, respectively, corresponding to amber suppression efficiencies of 23%, 23%, and 36%[51]. Also in a separate report, an *E. coli* strain with attenuated RF1 activity produced 21, 17, and 27 mg/L of sfGFP with 3, ELP with 12, and ELP with 22 ncAAs incorporated, respectively[50]. However, increasing the number of amber codons (i.e., 12 and 22 ncAA incorporation) resulted in numerous truncation products. Here we demonstrate a cell-free system from engineered recoded bacteria that enables high yields of proteins containing up to 40 ncAAs with no observable truncation products. Thus, our cell-free system will serve as a complement to in vivo methods and be a useful technology for developing OTSs for robust synthesis of modified proteins.

Looking forward, incorporating our discovered mutations into a recently published optimized strain, *C321.ΔA.opt*, might further increase the protein expression yields.[52] In addition, as new genomically recoded organisms with free codons are constructed[24,53–57], the development of extract chassis strains enabled by our MAGE-guided approach could aid the generation of highly efficient CFPS systems capable of incorporating two or more distinct ncAAs into a single protein or sequence-defined polymer. We envision that the generalized CFPS platform described here will be applied to on-demand biomanufacturing and biomolecular prototyping to transform biochemical engineering and expand the range of genetically encoded chemistry of biological systems.

Methods

Strains and plasmids. The bacterial strains and plasmids used in this study are listed in Supplementary Table 3. Details for strain construction, plasmid construction, verification, and culture growth are provided in the Supplementary Methods.

Cell extract preparation. For prototyping engineered strains, cells were grown in 1 L of 2 × YTPG media (pH 7.2) in a 2.5 L Tunair shake flask and incubated at 34 °C and 220 r.p.m. to OD_{600} of 3.0. Cells were pelleted by centrifuging for 15 min at 5000 × *g* and 4 °C, washed three times with cold S30 buffer (10 mM tris-acetate pH 8.2, 14 mM magnesium acetate, 60 mM potassium acetate, 2 mM dithiothreitol)[58], and stored at − 80 °C. To make cell extract, the thawed cells were suspended in 0.8 mL of S30 buffer per 1 g of wet cell mass and processed as reported by Kwon and Jewett[59]. Full details for cell growth, collection, and CFPS extract preparation are provided in the Supplementary Methods.

CFPS reactions. The PANOx-SP system was utilized for CFPS reactions[46]. Briefly, a 15 µL CFPS reaction in a 1.5 mL microcentrifuge tube was prepared by mixing the following components: 1.2 mM ATP; 0.85 mM each of GTP, UTP, and CTP; 34 µg/mL folinic acid; 170 µg/mL of *E. coli* tRNA mixture; 13.3 µg/mL plasmid; 16 µg/mL T7 RNA polymerase; 2 mM for each of the 20 standard amino acids; 0.33 mM nicotinamide adenine dinucleotide; 0.27 mM coenzyme-A; 1.5 mM spermidine; 1 mM putrescine; 4 mM sodium oxalate; 130 mM potassium glutamate; 10 mM ammonium glutamate; 12 mM magnesium glutamate; 57 mM HEPES pH 7.2; 33 mM PEP); and 27% v/v of cell extract. For ncAA incorporation, 2 mM pAcF, 0.5 mg/mL pAcFRS, and 10 µg/mL of o-tz-$tRNA^{opt}$ linear DNA were supplemented to cell-free reactions. For multi-site and consecutive ncAA incorporation, OTS^{opt} levels were increased to 5 mM pAcF, 1 mg/mL pAcFRS, and 30 µg/mL o-tz-$tRNA^{opt}$. o-$tRNA^{opt}$ linear DNA was amplified from pY71-T7-tz-o-$tRNA^{opt}$ plasmid as described previously and transcribed during the cell-free reaction[27]. Furthermore, the o-$tRNA^{opt}$ was expressed in the source strain via a plasmid prior to extract preparation. Techniques for purifying aminoacyl tRNA synthetases are described in the Supplementary Methods. When testing the effect of RNAse inhibitor, 1 µL (4U) of inhibitor (Qiagen, Valencia, CA) was added into each 15 µL reaction as per the manufacturer's suggestion. Each CFPS reaction was incubated for 20 h at 30 °C unless noted otherwise. Fed-batch and scale-up reaction formats are described in the Supplementary Methods.

Protein quantification. Protein quantification was performed using fluorescence detection and radioactive ^{14}C-Leucine incorporation and scintillation counting as described in the Supplementary Methods.

Whole-genome sequencing. As the gene encoding MutS is inactivated in *C321.ΔA*, we chose to fully sequence the genomes of six key strains produced during our screening efforts (*C321.ΔA*, *C321.ΔA.542*, *C321.ΔA.705*, *C321.ΔA.709*, *C321.ΔA.740*, and *C321.ΔA.759*). Whole-genome sequencing was performed by the Yale Center for Genome Analysis for library prep and analysis as described previously[60] (see Supplementary Methods). Genomes have been deposited to NCBI SRA collection, accession number PRJNA361365 (Individual accession numbers: SRX2511757-SRX2511762).

Nucleotide and amino acid quantitation using HPLC. Amino acid and nucleotide concentrations were measured via high-performance LC (HPLC). Cell-free reactions were clarified by precipitation with an equal volume 5% w/v trichloroacetic acid. Samples were centrifuged at 12,000 × *g* for 15 min at 4 °C and the supernatant stored at − 80 °C until analyzed using an Agilent 1290 series HPLC system (Agilent, Santa Clara, CA). For amino acid analysis, a Poroshell HPH-C18 column (4.6 × 100 mm, 2.7 µm particle size; Agilent) was used with an automatic pre-column derivatization method using *o*-pthalaldehyde and fluorenylmethyl chloroformate[61]. Nucleotides were analyzed using a BioBasic AX column (4.6 × 150 mm, 5 µm particle size; Thermo Scientific, West Palm Beach, FL). Full methods are described in the Supplementary Methods.

Full-length sfGFP and ELP purification and MS analysis. To confirm pAcF incorporation at corresponding amber sites, LC-MS analysis was performed on purified sfGFP and ELP reporter protein constructs with pAcF putatively incorporated. MS procedures and sample preparation details are given in the Supplementary Methods.

Data availability. All data generated or analyzed during this study are included in this published article (and its supplementary files) or are available from the corresponding authors on reasonable request. Genome sequences can be found at: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA361365/. Individual accession codes are: *C321.ΔA* (https://www.ncbi.nlm.nih.gov/sra/SRX2511762), *C321.ΔA.542* (https://www.ncbi.nlm.nih.gov/sra/SRX2511761), *C321.ΔA.705* (https://www.ncbi.nlm.nih.gov/sra/SRX2511760), *C321.ΔA.709* (https://www.ncbi.nlm.nih.gov/sra/SRX2511759), *C321.ΔA.740* (https://www.ncbi.nlm.nih.gov/sra/SRX2511758), and *C321.ΔA.759* (https://www.ncbi.nlm.nih.gov/sra/SRX2511757).

Received: 13 September 2017 Accepted: 13 February 2018
Published online: 23 March 2018

References

1. Perez J. G., Stark J. C., Jewett M. C. Cell-free synthetic biology: engineering beyond the cell. *Cold Spring Harb Perspect Biol* **8**, a023853 (2016).
2. Nirenberg, M. W. & Matthaei, J. H. The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl Acad. Sci. USA* **47**, 1588–1602 (1961).
3. Dudley, Q. M., Anderson, K. C. & Jewett, M. C. Cell-free mixing of *Escherichia coli* crude extracts to prototype and rationally engineer high-titer mevalonate synthesis. *ACS Synth. Biol.* **5**, 1578–1588 (2016).
4. Karim, A. S. & Jewett, M. C. A cell-free framework for rapid biosynthetic pathway prototyping and enzyme discovery. *Metab. Eng.* **36**, 116–126 (2016).
5. Kay, J. E. & Jewett, M. C. Lysate of engineered Escherichia coli supports high-level conversion of glucose to 2,3-butanediol. *Metab. Eng.* **32**, 133–142 (2015).

6. Noireaux, V., Bar-Ziv, R. & Libchaber, A. Principles of cell-free genetic circuit assembly. *Proc. Natl Acad. Sci. USA* **100**, 12672–12677 (2003).
7. Takahashi, M. K. et al. Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription-translation (TX-TL) systems. *ACS Synth. Biol.* **4**, 503–515 (2015).
8. Pardee, K. et al. Rapid, low-cost detection of zika virus using programmable biomolecular components. *Cell* **165**, 1255–1266 (2016).
9. Pardee, K. et al. Portable, on-demand biomolecular manufacturing. *Cell* **167**, 248–259 e212 (2016).
10. Salehi, A. S. et al. Cell-free protein synthesis of a cytotoxic cancer therapeutic: onconase production and a just-add-water cell-free system. *Biotechnol. J.* **11**, 274–281 (2016).
11. Yin, G. et al. Aglycosylated antibodies and antibody fragments produced in a scalable in vitro transcription-translation system. *MAbs* **4**, 217–225 (2012).
12. Park, H. S. et al. Expanding the genetic code of Escherichia coli with phosphoserine. *Science* **333**, 1151–1154 (2011).
13. Iwane, Y. et al. Expanding the amino acid repertoire of ribosomal polypeptide synthesis via the artificial division of codon boxes. *Nat. Chem.* **8**, 317–325 (2016).
14. Lang, K. & Chin, J. W. Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chem. Rev.* **114**, 4764–4806 (2014).
15. Liu, C. C. & Schultz, P. G. Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* **79**, 413–444 (2010).
16. Dumas, A., Lercher, L., Spicer, C. D. & Davis, B. G. Designing logical codon reassignment - expanding the chemistry in biology. *Chem. Sci.* **6**, 50–69 (2015).
17. Young, T. S. & Schultz, P. G. Beyond the canonical 20 amino acids: expanding the genetic lexicon. *J. Biol. Chem.* **285**, 11039–11044 (2010).
18. Cho, H. et al. Optimized clinical performance of growth hormone with an expanded genetic code. *Proc. Natl Acad. Sci. USA* **108**, 9060–9065 (2011).
19. Currier, N. V. et al. Targeted drug delivery with an integrin-binding knottin-Fc-MMAF conjugate produced by cell-free protein synthesis. *Mol. Cancer Ther.* **15**, 1291–1300 (2016).
20. Amiram, M. et al. Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. *Nat. Biotechnol.* **33**, 1272–1279 (2015).
21. O'Donoghue, P., Ling, J., Wang, Y. S. & Soll, D. Upgrading protein synthesis for synthetic biology. *Nat. Chem. Biol.* **9**, 594–598 (2013).
22. Chin, J. W. Reprogramming the genetic code. *Embo J.* **30**, 2312–2324 (2011).
23. Albayrak, C. & Swartz, J. R. Direct polymerization of proteins. *ACS Synth. Biol.* **3**, 353–362 (2014).
24. Lajoie, M. J. et al. Genomically recoded organisms expand biological functions. *Science* **342**, 357–360 (2013).
25. Wang H.H. et al. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894-898 (2009).
26. Isaacs F. J. et al. Precise manipulation of chromosomes *in vivo* enables genome-wide codon replacement. *Science* **333**, 348-353 (2011).
27. Hong, S. H. et al. Cell-free protein synthesis from a release factor 1 deficient *Escherichia coli* activates efficient and multiple site-specific nonstandard amino acid incorporation. *ACS Synth. Biol.* **3**, 398–409 (2014).
28. Pirman, N. L. et al. A flexible codon in genomically recoded *Escherichia coli* permits programmable protein phosphorylation. *Nat. Commun.* **6**, 8130 (2015).
29. Carlson, E. D., Gan, R., Hodgman, C. E. & Jewett, M. C. Cell-free protein synthesis: applications come of age. *Biotechnol. Adv.* **30**, 1185–1194 (2012).
30. Hong, S. H., Kwon, Y. C. & Jewett, M. C. Non-standard amino acid incorporation into proteins using *Escherichia coli* cell-free protein synthesis. *Front. Chem.* **2**, 34 (2014).
31. Josephson, K., Hartman, M. C. & Szostak, J. W. Ribosomal synthesis of unnatural peptides. *J. Am. Chem. Soc.* **127**, 11727–11735 (2005).
32. Agafonov, D. E., Huang, Y., Grote, M. & Sprinzl, M. Efficient suppression of the amber codon in *E. coli* in vitro translation system. *FEBS Lett.* **579**, 2156–2160 (2005).
33. Loscha, K. V. et al. Multiple-site labeling of proteins with unnatural amino acids. *Angew. Chem. Int Ed. Engl.* **51**, 2243–2246 (2012).
34. Hong, S. H. et al. Improving cell-free protein synthesis through genome engineering of *Escherichia coli* lacking release factor 1. *Chembiochem* **16**, 844–853 (2015).
35. Michel-Reydellet, N., Woodrow, K. & Swartz, J. Increasing PCR fragment stability and protein yields in a cell-free system with genetically modified *Escherichia coli* extracts. *J. Mol. Microbiol. Biotechnol.* **9**, 26–34 (2005).
36. Michel-Reydellet, N., Calhoun, K. & Swartz, J. Amino acid stabilization for cell-free protein synthesis by modification of the *Escherichia coli* genome. *Metab. Eng.* **6**, 197–203 (2004).
37. Jiang, X. P. et al. Reduction of protein degradation by use of protease-deficient mutants in cell-free protein synthesis system of *Escherichia coli*. *J. Biosci. Bioeng.* **93**, 151–156 (2002).
38. Kwon, Y. C. & Jewett, M. C. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Sci. Rep.* **5**, 8663 (2015).
39. Heinzelman, P., Schoborg, J. A. & Jewett, M. C. pH responsive granulocyte colony-stimulating factor variants with implications for treating Alzheimer's disease and other central nervous system disorders. *Protein Eng. Des. Sel.* **28**, 481–489 (2015).
40. Yang, J., Kanter, G., Voloshin, A., Levy, R. & Swartz, J. R. Expression of active murine granulocyte-macrophage colony-stimulating factor in an *Escherichia coli* cell-free system. *Biotechnol. Prog.* **20**, 1689–1696 (2004).
41. Young, T. S., Ahmad, I., Yin, J. A. & Schultz, P. G. An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J. Mol. Biol.* **395**, 361–374 (2010).
42. Des Soye, B. J., Patel, J. R., Isaacs, F. J. & Jewett, M. C. Repurposing the translation apparatus for synthetic biology. *Curr. Opin. Chem. Biol.* **28**, 83–90 (2015).
43. Voloshin, A. M. & Swartz, J. R. Efficient and scalable method for scaling up cell free protein synthesis in batch mode. *Biotechnol. Bioeng.* **91**, 516–521 (2005).
44. Zawada, J. F. et al. Microscale to manufacturing scale-up of cell-free cytokine production-a aew approach for shortening protein production development timelines. *Biotechnol. Bioeng.* **108**, 1570–1578 (2011).
45. Caschera, F. & Noireaux, V. Synthesis of 2.3 mg/ml of protein with an all *Escherichia coli* cell-free transcription-translation system. *Biochimie* **99**, 162–168 (2014).
46. Jewett, M. C. & Swartz, J. R. Mimicking the *Escherichia coli* cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol. Bioeng.* **86**, 19–26 (2004).
47. Raucher, D. & Ryu, J. S. Cell-penetrating peptides: strategies for anticancer treatment. *Trends Mol. Med.* **21**, 560–570 (2015).
48. Despanie, J., Dhandhukia, J. P., Hamm-Alvarez, S. F. & MacKay, J. A. Elastin-like polypeptides: therapeutic applications for an emerging class of nanomedicines. *J. Control Release* **240**, 93–108 (2016).
49. Catherine, C. et al. Engineering thermal properties of elastin-like polypeptides by incorporation of unnatural amino acids in a cell-free protein synthesis system. *Biotechnol. Bioproc. E* **20**, 417–422 (2015).
50. Wu, I. L. et al. Multiple site-selective insertions of noncanonical amino acids into sequence-repetitive polypeptides. *Chembiochem* **14**, 968–978 (2013).
51. Johnson, D. B. et al. RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat. Chem. Biol.* **7**, 779–786 (2011).
52. Kuznetsov, G. et al. Optimizing complex phenotypes through model-guided multiplex genome engineering. *Genome Biol.* **18**, 100 (2017).
53. Ostrov, N. et al. Design, synthesis, and testing toward a 57-codon genome. *Science* **353**, 819–822 (2016).
54. Richardson, S. M. et al. Design of a synthetic yeast genome. *Science* **355**, 1040–1044 (2017).
55. Annaluru, N. et al. Total synthesis of a functional designer eukaryotic chromosome. *Science* **344**, 55–58 (2014).
56. Wang, K. et al. Defining synonymous codon compression schemes by genome recoding. *Nature* **539**, 59–64 (2016).
57. Zhang, Y. et al. A semi-synthetic organism that stores and retrieves increased genetic information. *Nature* **551**, 644–647 (2017).
58. Jewett M. C., Calhoun K.A., Voloshin A., Wuu J. J., & Swartz J. R. An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol. Syst. Biol.* **4**, 220 (2008).
59. Kwon, Y. C. & Jewett, M. C. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Sci. Rep.* **5**, 8663 (2015).
60. Gallagher, R. R., Patel, J. R., Interiano, A. L., Rovner, A. J. & Isaacs, F. J. Multilayered genetic safeguards limit growth of microorganisms to defined environments. *Nucleic Acids Res.* **43**, 1945–1954 (2015).
61. Henderson J. W., Brooks, A. Improved Amino Acid Mmethods using Agilent ZORBAX Eclipse Plus C18 Columns for a Variety of Agilent LC Instrumentation and Separation Goals. (Agilent Technologies, Santa Clara, CA, 2010).

Acknowledgements

This work was supported by the Office of Naval Research (N00014-11-1-0363), the Defense Advanced Research Projects Agency (N66001-12-C-4211), the David and Lucile Packard Foundation, the Chicago Biomedical Consortium with support from the Searle Funds at the Chicago Community Trust, the Air Force Research laboratory (FA8650-15-2-5518), the Army Research Office (W911NF-11-1-0445 and W911NF-16-1-0372), the National Science Foundation (MCB-1716766), the National Institutes of Health (NIH-MSTP-TG-T32GM07205), the Dreyfus Teacher-Scholar program, and the Arnold and Mabel Beckman Foundation. This research was carried out in collaboration with the National Resource for Translational and Developmental Proteomics under Grant P41 GM108569 from the National Institute of General Medical Sciences, National Institutes of Health. R.W.M. is a recipient of the Ford Foundation Fellowship and NSF Graduate Fellowship. B.J.D. is a recipient of the NSF Graduate Fellowship. J. K. is an NDSEG Fellow. We thank Professor James R. Swartz for providing pK7-catGM-CSF plasmid, Professor Bradley Bundy for providing pY71 plasmids, and Professor Peter G. Schultz for

providing the pEVOL plasmid. We thank Dr Rui Gan for constructing the pDAK-pAzFRS plasmid. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of Air Force Research Laboratory, the Department of Defense, or the U.S. Government.

Author contributions

R.W.M., S.H.H., B.J.D. and M.C.J. designed the experiments and study. R.W.M., R.D.M., M.G.W., A.S., B.J.D. and S.H.H. performed MAGE genome engineering. R.W.M. prepared cell extracts and reagents. R.W.M., N.I.M., M.G.W., A.K.C., and S.H.H. performed CFPS reactions and ncAA incorporation experiments. R.G.D., P.M.T. and N.L.K. carried out the quantitative mass spectrometry analysis. Y.C.K. performed mRNA and DNA stability analysis. J.E.K. performed HPLC analysis. S.H.H., M.A. and R.W.M. constructed ELP plasmids, and R.W.M. and C.X.C. performed expression and purifications. J.R.P. carried out genome sequencing and analysis. F.J.I. supervised the construction of the recoded strain, a subset of the ELP constructs, and genome sequence analysis. R.W.M., S.H.H., B.J.D. and M.C.J. wrote the manuscript. All authors discussed the experimental results and commented in writing the manuscript.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-03469-5.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.