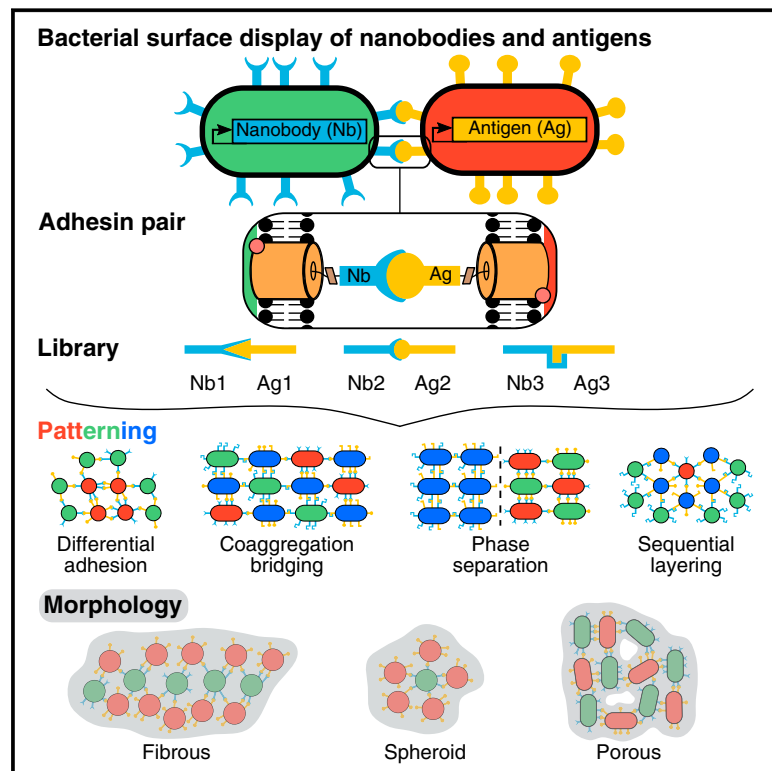# A Synthetic Bacterial Cell-Cell Adhesion Toolbox for Programming Multicellular Morphologies and Patterns

Graphical Abstract

Authors

David S. Glass, Ingmar H. Riedel-Kruse

Correspondence

ingmar@stanford.edu

In Brief

The development of a genetically encoded toolkit of surface-bound nanobodies and antigens in *E. coli* allows for precise manipulation of cell-cell adhesion and rational design of diverse self-assembled multicellular patterns and morphologies.

Highlights

- Orthogonal, composable adhesin library allows control over specificity and affinity

- Adhesion is maintained during cell growth and division

- Cultures form lattice-like, phase separation, and differential adhesion patterns

- Compatibility with synthetic biology standards allows complex multicellular designs

CellPress

Article

A Synthetic Bacterial Cell-Cell Adhesion Toolbox for Programming Multicellular Morphologies and Patterns

David S. Glass[1] and Ingmar H. Riedel-Kruse[1,2,*]
[1]Department of Bioengineering, Stanford University, 318 Campus Drive, Stanford, CA 94305, USA
[2]Lead Contact
*Correspondence: ingmar@stanford.edu
https://doi.org/10.1016/j.cell.2018.06.041

SUMMARY

Synthetic multicellular systems hold promise as models for understanding natural development of biofilms and higher organisms and as tools for engineering complex multi-component metabolic pathways and materials. However, such efforts require tools to adhere cells into defined morphologies and patterns, and these tools are currently lacking. Here, we report a 100% genetically encoded synthetic platform for modular cell-cell adhesion in *Escherichia coli*, which provides control over multicellular self-assembly. Adhesive selectivity is provided by a library of outer membrane-displayed nanobodies and antigens with orthogonal intra-library specificities, while affinity is controlled by intrinsic adhesin affinity, competitive inhibition, and inducible expression. We demonstrate the resulting capabilities for quantitative rational design of well-defined morphologies and patterns through homophilic and heterophilic interactions, lattice-like self-assembly, phase separation, differential adhesion, and sequential layering. Compatible with synthetic biology standards, this adhesion toolbox will enable construction of high-level multicellular designs and shed light on the evolutionary transition to multicellularity.

INTRODUCTION

Multicellular organisms display a variety of morphologies (three-dimensional structures) and patterns (spatial distributions of cell types) over multiple length scales using multiple cell types. There is growing interest in synthetic biology (Davies, 2008; Chen and Silver, 2012; Tabor et al., 2009; Basu et al., 2005; Tamsir et al., 2011; Danino et al., 2010; Liu et al., 2011) to engineer such multicellular arrangements in order to harness their unique abilities, such as separating intermediates in increasingly complex metabolic pathways (Chen and Silver, 2012; Avalos et al., 2013) or programing structured living materials (Jin and Riedel-Kruse, 2018; Nguyen et al., 2014; Chen et al., 2015) and tissues (Sia et al., 2007; Scholes and Isalan, 2017; Cachat et al., 2016). Syn-

thetic circuits have been engineered, for example, that direct cells on 2D substrates into self-organized ring patterns (Basu et al., 2005; Morsut et al., 2016; Payne et al., 2013). Such patterns were enabled by synthetic implementations of two tools: cell-cell signaling (Adams et al., 2014; Tamsir et al., 2011; Basu et al., 2005; Ortiz and Endy, 2012) and differentiation (Gardner et al., 2000; Bonnet et al., 2012). For natural multicellular organisms the key third tool for directing spatial organization is cell-cell adhesion (Rokas, 2008; Lyons and Kolter, 2015), but comparable synthetic tools are lacking (Davies, 2008; Teague et al., 2016).

Some synthetic cell-cell adhesion tools have, in fact, been developed to adhere various cell types (Cachat et al., 2016; Veiga et al., 2003; Piñero-Lambea et al., 2015; O'Brien et al., 2015; Todhunter et al., 2015; Koo et al., 2015), but have limited use for multicellular engineering due to (1) having limited control over specificity (Cachat et al., 2016; Veiga et al., 2003), (2) only mediating adhesion among very different cell types such as bacteria and mammalian cells (Piñero-Lambea et al., 2015), (3) being directly coupled to signaling events (Younger et al., 2017), or (4) having a non-genetic basis requiring chemical modifications that are diluted by growth (O'Brien et al., 2015; Todhunter et al., 2015; Koo et al., 2015). We propose that a synthetic cell-cell adhesion toolbox should have the following properties: it should be (1) genetically encoded, (2) decoupled from native signaling and adhesion, (3) easily extendable to an arbitrary library of adhesins, (4) tunable in binding strength and binding specificity, and (5) compatible with cell growth and division.

Here, we developed a synthetic cell-cell adhesion toolbox in *E. coli* that meets these criteria and enables controlled multicellular self-assembly (Figure 1). We quantified control over adhesive specificity (Figure 2) and strength (Figure 3). We further characterized the capability of synthetic adhesion to produce defined patterns and morphologies (Figure 4) even during cell growth and division (Figure 5). Finally, we demonstrated that these controls can be adjusted combinatorially to rationally design a variety of morphologies and patterns motivated by known natural processes (Figure 6).

RESULTS

Nanobody-Antigen Interactions Enable Design of a Synthetic Cell-Cell Adhesion Toolbox

We designed our adhesion toolbox from three elements: a transcriptional regulator, an outer membrane anchor, and an adhesin
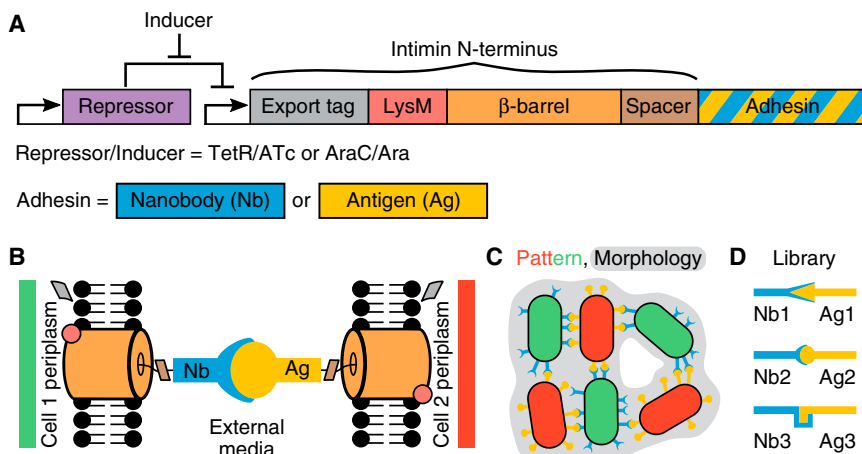
Figure 1. Design of a Synthetic Cell-Cell Adhesion Toolbox that Allows Aggregation of Multicellular Patterned Morphologies

(A) Adhesin constructs consist of a single coding sequence, with a nanobody (Nb) or antigen (Ag) serving as adhesin fused to the autotransporter intimin N terminus from enterohemorrhagic *E. coli*. A repressor (TetR or AraC), coupled to an inducer (ATc or Ara, respectively) that relieves the repression, can be used to control expression levels. TetR/ATc regulation was used except where noted otherwise.

(B) At the periplasm, intimin folds into the outer membrane, displaying the spacer and adhesin (Nb or Ag) outside the cell, mediating cell-cell adhesion.

(C) Nb-Ag interactions between cells can mediate production of microscopic patterns (spatial organization of cell types, denoted in color) and morphologies (overall spatial structure of all cells, denoted by gray background).

(D) A library of adhesin pairs can be used to expand adhesion capabilities.

See also Table S1.

library (Figure 1). We implemented the regulator using a standard TetR or AraC repressor, controlled by addition of small molecule inducers anhydrotetracycline (ATc) or arabinose (Ara), respectively (Figure 1A). We used TetR/ATc regulation except where noted otherwise. Outer membrane anchoring is achieved via the intimin N terminus of enterohemorrhagic *E. coli* (EHEC O157:H7), a reverse autotransporter and surface display system that includes a short N-terminal signal peptide (export tag) to direct its trafficking to the periplasm, a LysM domain for peptidoglycan binding, and a β-barrel for transmembrane insertion (Piñero-Lambea et al., 2015; Veiga et al., 2003; Salema et al., 2013) (Figures 1A and 1B). Fused to the C terminus of this truncated intimin is an adhesive domain, which we term the "adhesin." Expression of this entire fusion protein, which we term an "adhesin construct," is controlled by the repressor binding to a pTet (for TetR) or pBAD (for AraC) promoter upstream of intimin.

We implemented the adhesins using nanobodies and their corresponding antigens (Muyldermans, 2013; Salema et al., 2013) (Figures 1A and 1B). Nanobodies, the variable domains of camelid heavy-chain antibodies, can be expressed on bacterial surfaces due to their small size (~125 amino acids) and stability under a variety of conditions (Muyldermans, 2013). The combination of their single-domain structure and the intimin autotransporter allows the entirety of a highly specific, cell surface-bound adhesin to be encoded as a single fusion protein. We hypothesized that two *E. coli* strains displaying a nanobody (Nb) and a corresponding antigen (Ag), respectively, would specifically adhere to each other via the Nb-Ag interaction (Figures 1B and 1C) and that this would allow controlled morphology and patterning of multicellular assemblies (Figure 1C). We further hypothesized that an orthogonal library of such adhesin pairs could be established (Figure 1D) to expand patterning capabilities.

Specificity Control Demonstrates Orthogonality and Composability

To implement the design in Figure 1, we obtained a library of adhesin sequences from the VIB Nanobody Core (STAR Methods;

Table S1) consisting of 8 antigens of ≲125 amino acids each and 52 corresponding nanobodies that target these antigens (most of the antigens were paired with multiple nanobodies). For motivations behind the choice of these adhesins, see Figure S1. We cloned the 8 + 52 = 60 adhesins as fusions to intimin and transformed this adhesin construct library into MG1655 wild-type K-12 *E. coli* with a single adhesin construct per strain (see Figure S1B for all plasmid maps). To screen this library for adhesion, stationary-phase cultures were allowed to stand unshaken alone or in mixture with their corresponding Ag- or Nb-expressing partner strain. Cell-cell adhesion was detectable by macroscopic aggregation and settling (Figures 2A and 2B) within ~1 hr (Figure S1C). We quantified this aggregation by measuring optical density (OD_{600}) of cells remaining unaggregated in the upper half of the cultures after 24 hr (in order to approximate equilibrium and avoid false negatives). We identified three pairs of strains that aggregated and fell out of solution, comprising 6 strains termed Ag1–3 and Nb1–3 for the antigens and corresponding nanobodies, respectively (Figures 2C and S1D). Importantly, no such aggregation occurred in unmixed cultures, in mixtures without induction by ATc, or in a Null control containing the autotransporter but no adhesin (Figure 2C), indicating that the aggregation was driven by specific Nb-Ag binding interactions as designed. Roughly consistent with a naive estimate of 50% success rate (see Figure S1), 3/8 of the antigens mediated successful adhesion. Furthermore, these 3 antigens mediated adhesion with almost all of their corresponding nanobodies (Figure S1D), including 13 nanobodies targeting Ag3 (see below); conversely, none of the 33 nanobodies against the other 5 antigens appeared to mediate specific adhesion.

We next tested whether this nanobody-antigen library is both orthogonal (i.e., that adhesins only interact with designed partners) and composable (i.e., that arbitrary combinations of multiple adhesins function simultaneously within one cell). To demonstrate orthogonality, we assayed aggregation in all pairwise mixtures of the 7 strains used in Figure 2C (Ag1–3, Nb1–3,
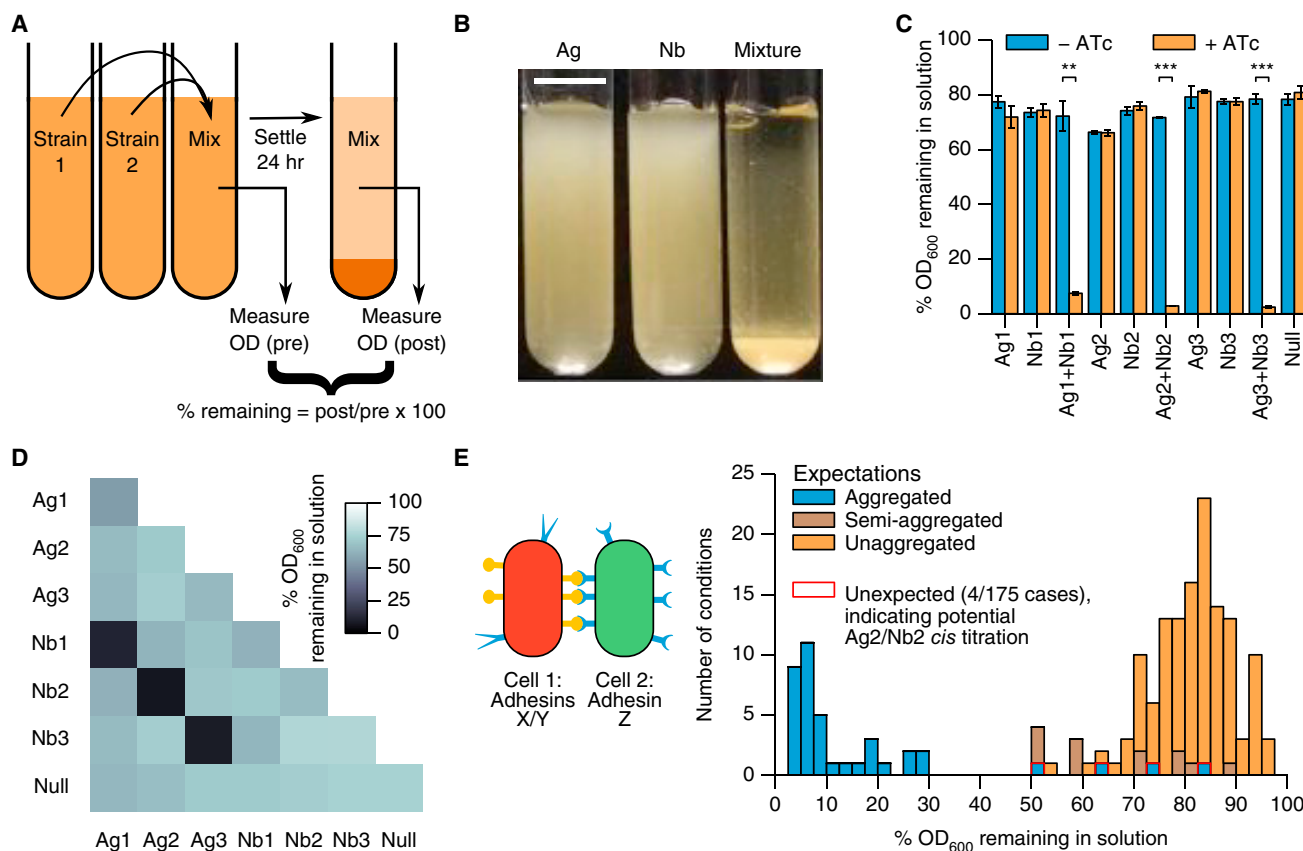
Figure 2. A Library of Adhesin Constructs Enables Multiple Ways of Tuning Adhesion Specificity

(A) An aggregation assay using optical density (OD_{600}) measurements allows the quantification of binding strength and specificity between cells.

(B) Binding in a 1:1 ratio of Ag2:Nb2 cell types leads to macroscopic aggregation and settling. Scale bar, 1 cm.

(C) Aggregating mixtures for three Ag-Nb pairs show significant settling compared to unmixed, uninduced, and no-adhesin (Null) conditions.

(D) Nb/Ag-based adhesion interactions are orthogonal, as strains only aggregate significantly with their designed partner strain ($p < 0.005$ for 2-tailed t test compared to median $\%OD_{600}$ remaining in solution).

(E) Multiple adhesin constructs can be used simultaneously (composability). Twenty-five strains containing all permutations of 5 adhesin constructs (intimin fusions to Ag2, Ag3, Nb2, Nb3, Null) on medium-copy (X) and low-copy (Y) plasmids were mixed with the original single-adhesin construct, medium-copy strains (Z). Aggregating cultures ($\%OD_{600} \lesssim 40$) behave as expected except for the strains containing both Ag2 and Nb2 in the same cell, indicating *cis* interactions for this pair (Figure S2). Semi-aggregated refers to expected mixtures of non-aggregating and (homophilically) aggregating cells. Displayed values are averages for n = 3 samples. Error bars, ±1 SD, **$p < 0.01$ or ***$p < 0.001$ according to a 2-tailed paired t test.

See also Figure S1.

Null), and indeed, significant reduction in supernatant density only occurred for mixtures of the designed Ag-Nb pairs (Figure 2D). To demonstrate composability of arbitrary pairs, we sought to systematically test triplet combinations of Ag-Nb interactions. Toward that end, we focused on Ag2, Ag3, Nb2, Nb3, and Null adhesin constructs (we will refer to Null as an adhesin construct for consistency despite its lack of an actual adhesin). We produced a set of 25 strains comprising all pairwise permutations of these 5 adhesin constructs, with one adhesin construct on a low-copy (pSC101 origin) plasmid and the other on a medium-copy (p15A origin) plasmid for convenience. We assayed aggregation of each of the 25 strains when mixed with the original 5 strains expressing just one of the 5 adhesin constructs, for a total of 25 × 5 = 125 conditions (Figure 2E, left). We also assayed aggregation in unmixed samples of the 25 strains with and without induction by ATc, for an

additional 25 × 2 = 50 conditions. Many of these combinations (e.g., Ag3/Nb3) have potential uses in patterning as we show later on; the future utility of other combinations (e.g., Ag2/Ag2) might be questionable at this point but are included so as to be systematic. Of the 125 + 50 = 175 total conditions, all but four (97.7%) behaved as expected (Figure 2E, right), aggregating if and only if a nanobody and its corresponding antigen were both present in the mixture (Figure S2 includes full dataset). Some combinations were expected to be "semi-aggregated" in the sense that a strain expressing both an antigen and its corresponding nanobody (e.g., Ag3/Nb3) would self-aggregate, leaving an orthogonal binder (e.g., Ag2) in solution. The remaining four represent unsuccessful homophilic adhesion (i.e., adhesion between like cells) of the cells producing both Ag2 and Nb2, which we speculate is due to *cis* titration of Nb2 by the much smaller (4 amino acid) Ag2 peptide when
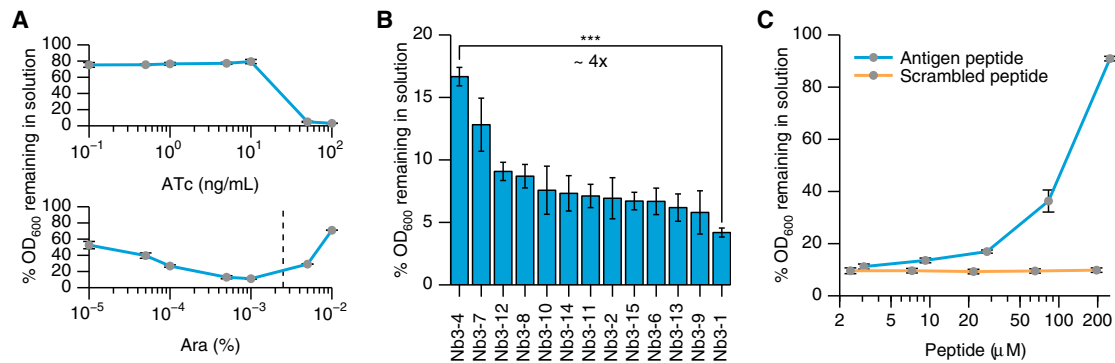
Figure 3. Adhesin Constructs Enable Multiple Ways of Tuning Adhesion Strength

(A) Aggregation of Ag2+Nb2 under different concentrations of ATc (top) and Ara (bottom) show induction control. Decrease of adhesion for Ara $\gtrsim 10^{-3}$% appears to be due to failure to fully process intimin with very high expression (Figures S3A–S3D) and thus Ara $> 10^{-3}$ should be avoided.

(B) Aggregation of the 13 different Nb3 variant strains that mediated adhesion (ranked by %OD result), when mixed with the Ag3 strain, demonstrates that adhesion can be controlled via individual Nb-Ag affinity. Expression levels of the various nanobodies are equivalent to within experimental error (Figure S3E), indicating that these differences are not simply due to differences in amount of adhesin as in (A).

(C) Aggregation of Ag2+Nb2 is diminished upon addition of soluble Ag2 peptide (sequence: EPEA) but not of a control peptide (sequence: PEAE). Displayed values are averages for n = 3 samples. Error bars, ±1 SD. Asterisks as in Figure 2.

both are expressed in the same cell (see Figure S2 for a further discussion).

Adhesion Strength between Cells Can Be Controlled Quantitatively by Multiple Strategies

We next sought to establish the ability to control adhesion strength quantitatively between cells through three independent methods. First, we controlled affinity on a per-cell basis by controlling adhesin construct expression level (Figure 3A). Varying inducer concentration over 3 orders of magnitude showed rather digital on/off control over aggregation by ATc and a more graded response but substantial leaky expression for Ara induction of an AraC-regulated construct, as would be expected for these induction systems (Figure S3). As an aside, we discovered that above 10^{-3}% Ara, the high level of expression may interfere with the processing of intimin, which may be the cause of decreased aggregation for such high Ara concentration (Figures S3A–S3D). Second, we controlled affinity on a per-molecule basis by individual Nb-Ag affinity. This we demonstrated using multiple different nanobodies against Ag3, which showed a range of binding strengths as indicated by their different levels of aggregation (Figure 3B). No detectable differences in intimin expression were observed among these strains (Figure S3E), indicating that these differences are not simply due to differences in amount of adhesin as with induction control. Third, we used a soluble peptide to competitively inhibit cell-cell adhesion. In particular, soluble Ag2 peptide blocked the formation of aggregates between Ag2 and Nb2 strains in a concentration-dependent manner, while a scrambled-sequence control peptide had no effect (Figure 3C). The potential *cis* titration of Nb2 by Ag2 mentioned above could provide a fourth method similar to competitive inhibition by soluble Ag2. Thus, we demonstrated three independent methods for quantitatively controlling adhesion strength, and other methods such as the speculated *cis* titration may be possible in the future as well.

Adhesion Mediates Self-Assembly of Multicellular Aggregates with Defined Morphologies and Precise Lattice-like Patterning

With this synthetic adhesion toolbox in hand, we explored what multicellular morphologies and spatial patterns could be achieved at a microscopic scale with a single adhesin pair expressed in two strains. We labeled Ag2- and Nb2-expressing strains with constitutive, cytoplasmic fluorescent proteins mRuby2 (red) and sfGFP (green), respectively, leading to extended aggregates with mesh-like patterns of alternating red and green cells (Figure 4A). The aggregates were observed to be essentially static structures over the timescale of microscopy acquisition, which allowed straightforward quantification of patterning by measuring cell centroid locations (see STAR Methods). The observed red-green binding specificity is statistically significant when quantifying the number of nearest neighbors (Figure 4B) around any given cell type, a metric that can be summarized concisely using a conditional probability table (Figure 4C; Data S1; STAR Methods). Taken together, these data show that synthetic heterophilic adhesion is able to mediate microscopic patterning of two cell types.

We then tested whether we could control overall morphology of the aggregates, in addition to their local patterning, by varying cell shape. Specifically, we expressed surface Ag2 and Nb2 in the spherical S1 strain (its shape due to an *mrdB* mutation) (Matsuzawa et al., 1973), shown in Figure 4D, and in a filamentous strain (its shape due to the stress of high-copy plasmid overexpression of the adhesin construct) (Wagner et al., 2007) (Figure S4; STAR Methods) shown in Figure 4E. The resulting aggregates differ significantly both in their microscopic porosity (fraction of space not occupied by cells), which we measured in aggregates that had settled for 24 hr to approximate equilibrium (Figures 4F and S4A–S4C), as well as in their macroscopic pellet size (Figures S4D and S4E). The data indicate that spheres pack more compactly than rods, which pack more compactly than filaments as would be expected by geometry. However,
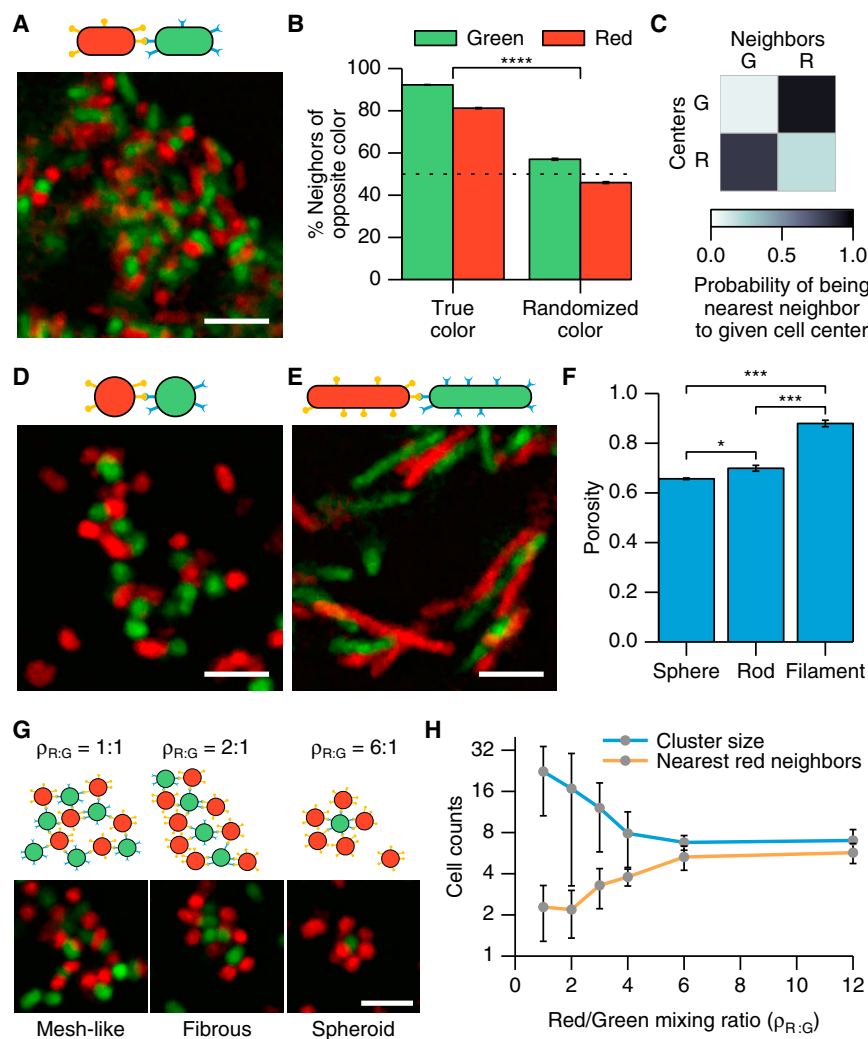
Figure 4. The Adhesion Toolbox Enables Self-Assembly of Multicellular Aggregates with Defined Morphologies and Precise Lattice-like Patterning

(A) A single confocal z slice from an Ag2+Nb2 aggregate. Red, Ag2 cells express cytoplasmic mRuby2. Green, Nb2 cells express cytoplasmic sfGFP.

(B) Quantification of 3D confocal stacks as in (A). Most cells neighboring any given cell are of the opposite cell type. In contrast, randomizing cell identities shows approximately uniform number of neighboring cell types ($p \approx 10^{-45}$ for true versus random coloring by χ^2-test with 2 df; see STAR Methods for details). Lower percentages for red cell centers is likely due to there being $24\% \pm 3\%$ more red cells than green cells in these samples (see G).

(C) Conditional probability table for (A), which reports the chance for a cell of a given color (column) being the nearest-neighbor to a cell of another color (row). See STAR Methods.

(D and E) Same as in (A) but with spherical (D) and filamentous cells (E), respectively.

(F) Porosity quantification of 24-hr aggregates from (A, D, and E) show significant differences in structure, with spheres packing more tightly than rods packing more tightly than filaments, as expected by geometry (24-hr images shown in Figure S4). Displayed values are averages for n = 3 confocal stacks with dimensions 212.5 μm × 212.5 μm × 6.4 μm (x × y × z).

(G) Increasing the density ratio $\rho_{R:G}$ of spherical cells predictably alters morphology (cluster size and shape) and patterning (nearest red neighbors per green cell) as available binding partners for green cells decreases (top: schematic, bottom: data). In particular, more red cells are bound to any given green cell and overall cluster size decreases as $\rho_{R:G}$ is increased.

(H) Quantification of (G), with displayed values averages over 9 clusters. Error bars, ±SEM in (B), otherwise ±1 SD. Asterisks as in Figure 2 with ****p < 0.0001. Scale bars, 5 μm.

See also Data S1.

there are significant differences in expression among these three strains as well, with spherical cells expressing much less of the adhesin constructs than the rod-shaped cells, which express at a much lower level than do the filamentous cells (Figures S4F and S4G). Thus, differences in aggregate morphology may be due to a combination of expression and cell shape effects. For example, spherical cells fall out of solution very slowly, with no macroscopic aggregates observable after 1 hr, which could be due to low expression. Moreover, other parameters may also affect aggregate morphology, such as total cell density or the amount of time allowed for settling. The important conclusion is that aggregate morphology can in fact be controlled by the combination of simple parameters such as cell shape and expression level.

Focusing further on the spherical cells, we next tested whether we could control aggregate size and morphology in microscopic aggregates by varying the density ratio $\rho_{R:G}$ of two adhering cell types (Figure 4G). We found a transition (Figures 4G and 4H) from large, mesh-like structures ($\rho_{R:G} \approx 1:1$) to more elongated

fibrous clusters $\rho_{R:G} \approx 2:1$) and eventually to small spheroids ($\rho_{R:G} \approx 6:1$). These different morphologies and cluster sizes are expected as the more abundant cell type makes maximal use of the available binding sites around the less abundant cell type by surrounding it. Thus, simple control of cell type ratios can be used to control aggregate size and morphology. Overall, the adhesion toolbox enables production of aggregates with lattice-like patterns as well as modulation of aggregate size and morphology. Ample opportunities exist for future research to explore the parameter space of such control, especially over aspects such as aggregate morphology and material properties.

Adhesion Is Compatible with Cell Growth and Division

We also wished to show that our toolbox is compatible with patterning throughout cell growth and division. To do so, we tracked small aggregates of exponential-phase cells in a microfluidic chamber (Lam et al., 2017) for several hours. We observed cells growing and dividing over multiple cell cycles while also adhering to other cells (Figure 5). Pairs of red and green cells
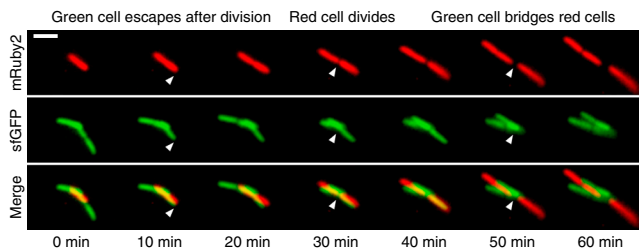
Figure 5. Self-Assembly of Multicellular Aggregates Is Compatible with Cell Growth and Division
Time lapse of adhesive co-cultures shows patterning even during cell growth and division (Ag2/sfGFP + Nb2/mRuby2). White arrows point to processes labeled on top. See also Video S1. Scale bar, 5 μm.

bound lengthwise gave rise to multiple generations of daughter cells similarly bound lengthwise, leading to small filaments of two to three cell widths. Absent an adjacent cell of the opposite color, daughter cells separated from the aggregate after division (Figure 5, left). Conversely, the presence of the opposite cell type maintained daughter cells as part of the aggregate by acting as an adhesive bridge (Figure 5, center, right). These results demonstrate that the bacteria are able to grow and divide while adherent; likewise, they continue to produce sufficient numbers of surface-displayed adhesins to support cell-cell binding during growth and division.

Complex Patterns Can Be Rationally Designed Using a Combinatorial Approach

Next, we sought to rationally design distinct patterns involving more than one adhesion pair in two cell types. We were motivated in these implementations by three key canonical patterning processes (Figures 6A–6C): differential adhesion (Steinberg, 1963), phase separation (Steinberg, 1963), and coaggregation bridging (Kolenbrander et al., 2006). First, the differential adhesion mechanism (Steinberg, 1963) enables patterning of cells that can all bind to each other through homophilic interactions but with different strengths, and where those with stronger binding localize to the center of an aggregate while those with weaker binding localize to the periphery. We achieved an analogous mechanism using only heterophilic Nb-Ag interactions (Figure 6A) with a mixture of four strains producing Ag2 or Nb2 at high or low levels. Here, both high-expressing cell types are labeled red (mRuby2) and both low-expressing cell types are labeled green (sfGFP) to match the canonical two cell-type differential adhesion (Steinberg, 1963). Note that this pattern is abolished when all cells produce adhesin at equal levels as expected (Figure S5). Second, phase separation, defined as the spontaneous separation of cell types into distinct aggregates, can occur in the limit of no adhesion interaction between two groups of cells (Steinberg, 1963). Note the contrast here between the no-binding condition of phase separation and the all-binding-all scenario of differential adhesion. We achieved this phase separation (Figure 6B) by generating a blue cell (Cerulean) that is homophilically adhesive through simultaneous expression of Ag3 and Nb3 and then co-mixing this cell with the heterophilically adhering Ag2 (green, Venus) and Nb2 (red, mCherry) cells demonstrated in Figure 4A. These three cell types then separated into homophilic (blue) and

heterophilic (green/red) phases as expected. Third, certain natural systems such as dental plaque biofilms exhibit a phenomenon known as coaggregation bridging (Kolenbrander et al., 2006), in which two otherwise non-interacting cell types adhere indirectly through an intermediate capable of binding both. We achieved this (Figure 6C) by generating a blue (Cerulean) cell type presenting both Ag2 and Ag3, which binds green (Venus) Nb3 and red (mCherry) Nb2 cells.

Finally, we wanted to demonstrate how varying other parameters such as cell shape, density ratio, and timing of culture mixing expands the available patterning space. We achieved this (Figure 6D) using the same adhesin set as in Figure 6C but using spherical cells, mixing the cells sequentially (red, blue, green), and in increasing densities of 1:6:36. This process led to bullseye patterns that differ qualitatively from Figure 6C. Note that this bullseye is accomplished solely through cell positioning (via adhesion), in contrast to methods using differentiation and signaling on pre-positioned cells (Basu et al., 2005; Morsut et al., 2016; Payne et al., 2013). We quantified each of the four patterns in Figure 6 using conditional probability tables as in Figure 4C (see bottom row of Figure 6), with each pattern differing quantitatively according to this metric (Figure S6). Altogether, Figure 6 demonstrates the adhesion toolbox's rich patterning capabilities, including patterning length scales of 2 cells (Figure 6B heterophilic phase), 3 cells (Figures 6C and 6D), and many cells (Figures 6A and 6B, homophilic phases).

Quantitative Predictions Enable Rational Design of Nearest Neighbor Interactions

To determine the extent to which patterns can be rationally designed at the level of nearest-neighbor interactions, we compared experimentally measured conditional probability tables for patterns in Figure 6 to theoretical estimates of these tables (Figure 6, bottom). We developed a simple heuristic that yields these estimates based on data from Figures 2, 3, and 4 (see STAR Methods for a full derivation). More precisely, we estimated the probabilities p_{ij} for cell type i (e.g., a green cell) to have a cell of type j (e.g., a red cell) as its nearest neighbor using the following heuristic:

$$p_{ij} = \frac{\rho_{j;i}(K_{ij} + \kappa_{ij})}{N_i}. \qquad \text{(Equation 1)}$$

Here, $\rho_{j;i}$ is the density ratio of cell type j to cell type i. K_{ij} is the binding strength between cell types i and j (normalized so that a direct Nb-Ag binding strength equals 1 at 100 ng/mL ATc induction). κ_{ij} is an apparent binding strength between cells i and j even when not directly bound ($K_{ij} = 0$). κ_{ij} arises when both i and j bind a third cell m, which may cause i and j to artifactually appear to be nearest neighbors in an image. $N_i = \sum_j \rho_{j;i}(K_{ij} + \kappa_{ij})$ is a normalization factor chosen to make the values of p_{ij} legitimate probabilities (i.e., $\sum_j p_{ij} = 1$).

Importantly, $\kappa_{ij} \approx 0.18$ for all i, j can be fit from the data in Figure 4C (see STAR Methods), binding strengths K_{ij} can be estimated from Figure 3A, and certain binding strength constants can be approximated as zero based on the orthogonality data of Figure 2 (Figure 6, schematics; STAR Methods). The mixing ratios $\rho_{j;i}$ are reported in the Figure 6 schematics. Taken together,
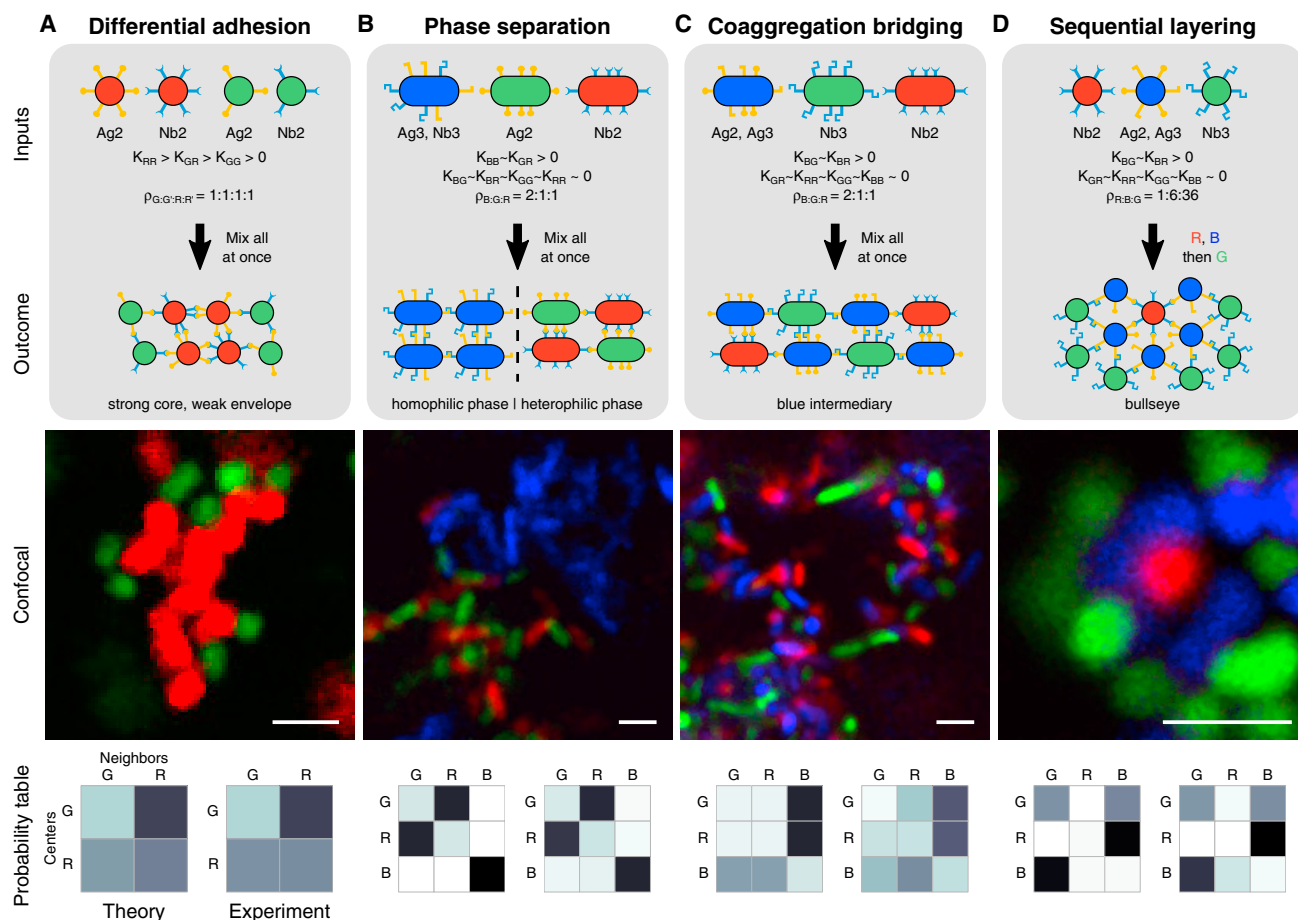
A Differential adhesion

Inputs

Ag2 Nb2 Ag2 Nb2

$K_{RR} > K_{GR} > K_{GG} > 0$

$\rho_{G:G':R:R'} = 1:1:1:1$

Mix all at once

Outcome

strong core, weak envelope

B Phase separation

Ag3, Nb3 Ag2 Nb2

$K_{BB} \sim K_{GR} > 0$
$K_{BG} \sim K_{BR} \sim K_{GG} \sim K_{RR} \sim 0$
$\rho_{B:G:R} = 2:1:1$

Mix all at once

homophilic phase | heterophilic phase

C Coaggregation bridging

Ag2, Ag3 Nb3 Nb2

$K_{BG} \sim K_{BR} > 0$
$K_{GR} \sim K_{RR} \sim K_{GG} \sim K_{BB} \sim 0$
$\rho_{B:G:R} = 2:1:1$

Mix all at once

blue intermediary

D Sequential layering

Nb2 Ag2, Ag3 Nb3

$K_{BG} \sim K_{BR} > 0$
$K_{GR} \sim K_{RR} \sim K_{GG} \sim K_{BB} \sim 0$
$\rho_{R:B:G} = 1:6:36$

R, B then G

bullseye

Confocal

Probability table

Neighbors
G R

Centers G R

Theory Experiment

Figure 6. Complex Multicellular Patterns Can Be Rationally Designed Using the Synthetic Adhesion Toolbox in a Combinatorial Fashion

(A) Difference in expression levels between highly adherent cells (Ag2/mRuby2 + Nb2/mRuby2, 100 ng/mL ATc) and weakly adherent cells (Ag2/sfGFP + Nb2/sfGFP, 0.0001% Ara) drives self-assembly through differential adhesion into clusters of red cells surrounded by green cells.

(B) Lack of adhesion between self-adherent homophilic (Ag3/Nb3/Cerulean) and heterophilic (Ag2/Venus + Nb2/mCherry) aggregates drives phase separation.

(C) The presence of a doubly adhesive strain (Ag2/Ag3/Cerulean) drives coaggregation bridging of non-interacting cells (Nb3/Venus and Nb2/mCherry).

(D) Sequential addition of excess binding cells can produce layered "bullseye" clusters (Nb2/mCherry + excess Ag2/Ag3/Cerulean, followed by excess Nb3/Venus). (A–D) Each panel includes, from top to bottom: (1) strains and adhesins used, including qualitative estimates of relative association constants (K) and density ratios (ρ) between green (G), red (R), and blue (B) cells; (2) mixing protocol; (3) expected patterning outcome and underlying mechanism; (4) typical confocal z slices (scale bars, 2.5 μm); and (5) conditional probability tables as in Figure 4C based on a heuristic of pairwise rules using data from Figures 2, 3, and 4 (Theory, left) and quantification of confocal images (Experimental, right), with color scale as in Figure 4C. These theoretical and experimental conditional probability tables agree quantitatively (Figure S6; STAR Methods). Rods are MG1655. Spheres are S1. Quantification is averaged over n = 3 confocal stacks (B and C) or 6 clusters of >15 cells (A) and 9 clusters (D) in confocal slices (see STAR Methods).
See also Figure S5 and Data S1.

these variables produce the theoretical conditional probability tables presented in Figure 6. Explicit calculations are available in the STAR Methods. Hierarchical clustering of all experimental and theoretical tables in Figure 6 show that predictions cluster with their corresponding experiments (Figure S6). Thus, we can rationally predict adhesion patterns at the level of nearest-neighbor interactions, as quantified by conditional probability tables, using Equation 1.

DISCUSSION

In summary, we established a synthetic cell-cell adhesion toolbox that, through quantitative control over key parameters,

enables rational programming of varied multicellular morphologies and patterns. The success of this toolbox relies on the strong, specific interactions of nanobody-antigen pairs (Muyldermans, 2013; Salema et al., 2013) and an outer membrane anchor from EHEC O157:H7 to display these proteins on the bacterial surface (Piñero-Lambea et al., 2015; Veiga et al., 2003; Salema et al., 2013). Quantitative characterization of pairwise (Figures 2A–2D and 3) and triplet (Figure 2E) interactions in macroscopic cultures, as well as microscopic quantification of spatial organization (Figure 4), enabled rational design of a variety of patterns and morphologies (Figure 6).

It should be noted that the conditional probability tables used here (e.g., bottom of Figure 6) are a local metric of patterning and

not a global metric of patterning or of morphology. For example, the tables show that red and green cells are arranged in an alternating pattern due to heterophilic binding (e.g., Figure 4C), but they make no statement about the overall size of the aggregate, such as that measured by the blue curve in Figure 4H. Likewise, these tables do not measure or predict the morphological arrangement of cells as would be necessary to describe, for example, the porosity measurements presented in Figure 4F. At the heart of this limited prediction is the fact that the aggregates are not perfect lattices, but rather exhibit a partially disordered packing of cells. Because the smallest length scale of the patterns is set by the dimension of the cells themselves, patterning on such disordered lattices is bound to be similarly stochastic. In the future, other metrics should be employed to more fully characterize the patterning and morphology on various length scales (e.g., order parameters previously used for other active multiparticle systems) (Ramaswamy, 2010).

The current work has considered primarily steady-state aggregates of highly adhesive, stationary cultures but not their underlying dynamics. Studying these dynamics of bacterial self-assembly and rearrangement could significantly increase our capabilities for rational design of multicellular patterning and morphology (Cademartiri and Bishop, 2015; Murugan et al., 2015; Huntley et al., 2016; Whitesides and Grzybowski, 2002). Biophysical characterization of these dynamics would in fact be required to understand and predict patterning behavior outside of this regime, where lower adhesive strength (higher off rate) on par with flagellar forces (Berry and Berg, 1997; Klamecka et al., 2015) should lead to dynamic patterning and morphologies that change over time. This will be especially pertinent in the context of growing cultures, and could open up new opportunities to develop active materials (Ramaswamy, 2010; Needleman and Dogic, 2017).

Rational design of patterning and morphology would also benefit from a "multicellular compiler," analogous to gene-circuit design tools for single-cell engineering (Nielsen et al., 2016; Salis et al., 2009). Using such a compiler, desired patterns and morphologies would be specified on a computer, and the appropriate cell types, mixing ratios, induction levels, and mixing order would be chosen algorithmically using predictors such as Equation 1. As noted earlier, only the local, stochastic nearest-neighbor patterning as given by conditional probability tables can currently be predicted for these disordered packings. Prediction of global metrics such as cluster size, porosity, and packing regularity would require further study, as would prediction of dynamic structure.

In order to increase the types of patterns and morphologies that can be generated, many extensions should be added to the adhesion toolbox. For example, a larger adhesin library could easily be constructed by screening through more nanobody-antigen pairs (Salema et al., 2013). The programmable patterns in Figure 6 are all based on adhesins distributed isotropically over the outer membrane, but sub-cellular localization, which has been documented for other autotransporters besides the intimin used here (Jain et al., 2006), would allow spatial symmetry breaking and the production of linear chains or sheets (Keller, 2006). Anecdotally, mechanical agitation reversibly disrupts the essentially static aggregates engineered in this work, but genet-

ically encoding an excreted competitive inhibitor (cf. Figure 3C) or intimin-specific protease could control dynamic, reversible adhesion on a microscopic scale. More broadly, this entire system is designed for *E. coli*, but porting to other cell types including eukaryotes could be accomplished through the use of suitable surface display anchors such as pDisplay (Morsut et al., 2016; Eiraku et al., 2002; Santiago et al., 2002; Forns et al., 1999).

Engineering of more complex synthetic multicellular systems will be enabled through the combination of adhesin-based control over morphology and patterning with cell-cell signaling (Adams et al., 2014; Tamsir et al., 2011; Basu et al., 2005; Ortiz and Endy, 2012; Toda et al., 2018), differentiation (Gardner et al., 2000; Bonnet et al., 2012), and gene regulatory logic (Tabor et al., 2009; Tamsir et al., 2011). With that goal in mind, all plasmid sequences used in this work were made compatible with the BioBricks standard (Shetty et al., 2008), one of several popular synthetic biology parts assembly standards (Casini et al., 2015). Such implementations should have broad utility for efficient pathway compartmentalization in metabolic consortia engineering (Chen and Silver, 2012; Avalos et al., 2013), implementation of cell-autonomous morphogenesis in engineered tissues (Sia et al., 2007; Scholes and Isalan, 2017; Cachat et al., 2016), and production of living materials (Jin and Riedel-Kruse, 2018; Nguyen et al., 2014; Chen et al., 2015). Compatibility with cell growth and division (Figure 5) will be a prerequisite for many of these designs.

Finally, it should be noted that synthetic biology has broadly enabled a build-to-understand methodology for studying the behavior of intracellular phenomena, such as protein production, gene network regulation, and genomic organization (Gardner et al., 2000; Chan et al., 2005; Temme et al., 2012; Hecht et al., 2017; Hutchison et al., 2016). Similarly, multicellular insights have been previously elucidated using synthetic analogs of two crucial multicellular processes, differentiation (Morsut et al., 2016) and cell-cell signaling (Basu et al., 2005). The contribution of a synthetic cell-cell adhesion toolbox provides the third pillar to complete a minimum set of tools required for multicellular organisms (Rokas, 2008; Lyons and Kolter, 2015), enabling controlled study of engineered multicellular interactions. Analogous to how minimal single-celled organisms can provide insights into the origin of life (Hutchison et al., 2016), we propose that minimal multicellular organisms using synthetic adhesion, differentiation, and signaling should provide bottom-up insights into natural development and the evolutionary transition to multicellularity (Rokas, 2008; Lyons and Kolter, 2015).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Strains and sequences
  - Culture conditions
- METHOD DETAILS
  - Aggregation assays
  - Peptides

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures, one table, one video, and one data file and can be found with this article online at https://doi.org/10.1016/j.cell.2018.06.041.

AUTHOR CONTRIBUTIONS

D.S.G. and I.H.R.-K. jointly conceived the project and wrote the paper. D.S.G. performed experiments and analysis.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Adams, B.L., Carter, K.K., Guo, M., Wu, H.C.C., Tsao, C.Y.Y., Sintim, H.O., Valdes, J.J., and Bentley, W.E. (2014). Evolved Quorum sensing regulator, LsrR, for altered switching functions. ACS Synth. Biol. *3*, 210–219.

Avalos, J.L., Fink, G.R., and Stephanopoulos, G. (2013). Compartmentalization of metabolic pathways in yeast mitochondria improves the production of branched-chain alcohols. Nat. Biotechnol. *31*, 335–341.

Basu, S., Gerchman, Y., Collins, C.H., Arnold, F.H., and Weiss, R. (2005). A synthetic multicellular system for programmed pattern formation. Nature *434*, 1130–1134.

Berry, R.M., and Berg, H.C. (1997). Absence of a barrier to backwards rotation of the bacterial flagellar motor demonstrated with optical tweezers. Proc. Natl. Acad. Sci. USA *94*, 14433–14437.

Bonnet, J., Subsoontorn, P., and Endy, D. (2012). Rewritable digital data storage in live cells via engineered control of recombination directionality. Proc. Natl. Acad. Sci. USA *109*, 8884–8889.

Cachat, E., Liu, W., Martin, K.C., Yuan, X., Yin, H., Hohenstein, P., and Davies, J.A. (2016). 2- and 3-dimensional synthetic large-scale de novo patterning by mammalian cells through phase separation. Sci. Rep. *6*, 20664.

Cademartiri, L., and Bishop, K.J.M. (2015). Programmable self-assembly. Nat. Mater. *14*, 2–9.

Casini, A., Storch, M., Baldwin, G.S., and Ellis, T. (2015). Bricks and blueprints: methods and standards for DNA assembly. Nat. Rev. Mol. Cell Biol. *16*, 568–576.

Chan, L.Y., Kosuri, S., and Endy, D. (2005). Refactoring bacteriophage T7. Mol. Syst. Biol. *1*, 2005.0018.

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

Chen, A.Y., Zhong, C., and Lu, T.K. (2015). Engineering living functional materials. ACS Synth. Biol. *4*, 8–11.

Danino, T., Mondragón-Palomino, O., Tsimring, L., and Hasty, J. (2010). A synchronized quorum of genetic clocks. Nature *463*, 326–330.

Davies, J.A. (2008). Synthetic morphology: prospects for engineered, self-constructing anatomies. J. Anat. *212*, 707–719.

De Genst, E.J., Guilliams, T., Wellens, J., O'Day, E.M., Waudby, C.A., Meehan, S., Dumoulin, M., Hsu, S.T., Cremades, N., Verschueren, K.H., et al. (2010). Structure and properties of a complex of α-synuclein and a single-domain camelid antibody. J. Mol. Biol. *402*, 326–343.

Eiraku, M., Hirata, Y., Takeshima, H., Hirano, T., and Kengaku, M. (2002). Delta/notch-like epidermal growth factor (EGF)-related receptor, a novel EGF-like repeat-containing protein targeted to dendrites of developing and adult central nervous system neurons. J. Biol. Chem. *277*, 25400–25407.

Forns, X., Emerson, S.U., Tobin, G.J., Mushahwar, I.K., Purcell, R.H., and Bukh, J. (1999). DNA immunization of mice and macaques with plasmids encoding hepatitis C virus envelope E2 protein expressed intracellularly and on the cell surface. Vaccine *17*, 1992–2002.

Gardner, T.S., Cantor, C.R., and Collins, J.J. (2000). Construction of a genetic toggle switch in *Escherichia coli*. Nature *403*, 339–342.

Hecht, A., Glasgow, J., Jaschke, P.R., Bawazer, L.A., Munson, M.S., Cochran, J.R., Endy, D., and Salit, M. (2017). Measurements of translation initiation from all 64 codons in E. coli. Nucleic Acids Res. *45*, 3615–3626.

Huntley, M.H., Murugan, A., and Brenner, M.P. (2016). Information capacity of specific interactions. Proc. Natl. Acad. Sci. USA *113*, 5841–5846.

Hutchison, C.A., 3rd, Chuang, R.Y., Noskov, V.N., Assad-Garcia, N., Deerinck, T.J., Ellisman, M.H., Gill, J., Kannan, K., Karas, B.J., Ma, L., et al. (2016). Design and synthesis of a minimal bacterial genome. Science *351*, aad6253.

Jain, S., van Ulsen, P., Benz, I., Schmidt, M.A., Fernandez, R., Tommassen, J., and Goldberg, M.B. (2006). Polar localization of the autotransporter family of large bacterial virulence proteins. J. Bacteriol. *188*, 4841–4850.

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

Jones, E., Oliphant, T., and Peterson, P. (2001). SciPy: Open source scientific tools for Python. http://www.scipy.org.

Keller, R. (2006). Mechanisms of elongation in embryogenesis. Development *133*, 2291–2302.

Klamecka, K., Severin, P.M., Milles, L.F., Gaub, H.E., and Leonhardt, H. (2015). Energy profile of nanobody-GFP complex under force. Phys. Biol. *12*, 056009.

Kolenbrander, P.E., Palmer, R.J., Jr., Rickard, A.H., Jakubovics, N.S., Chalmers, N.I., and Diaz, P.I. (2006). Bacterial interactions and successions during plaque development. Periodontol. 2000 *42*, 47–79.

Koo, H., Choi, M., Kim, E., Hahn, S.K., Weissleder, R., and Yun, S.H. (2015). Bioorthogonal click chemistry-based synthetic cell glue. Small *11*, 6458–6466.

Lam, A.T., Samuel-Gama, K.G., Griffin, J., Loeun, M., Gerber, L.C., Hossain, Z., Cira, N.J., Lee, S.A., and Riedel-Kruse, I.H. (2017). Device and programming abstractions for spatiotemporal control of active micro-particle swarms. Lab Chip *17*, 1442–1451.

Liu, C., Fu, X., Liu, L., Ren, X., Chau, C.K., Li, S., Xiang, L., Zeng, H., Chen, G., Tang, L.H., et al. (2011). Sequential establishment of stripe patterns in an expanding cell population. Science *334*, 238–241.

Lyons, N.A., and Kolter, R. (2015). On the evolution of bacterial multicellularity. Curr. Opin. Microbiol. *24*, 21–28.

Matsuzawa, H., Hayakawa, K., Sato, T., and Imahori, K. (1973). Characterization and genetic analysis of a mutant of Escherichia coli K-12 with rounded morphology. J. Bacteriol. *115*, 436–442.

Morsut, L., Roybal, K.T., Xiong, X., Gordley, R.M., Coyle, S.M., Thomson, M., and Lim, W.A. (2016). Engineering customized cell sensing and response behaviors using synthetic notch receptors. Cell *164*, 780–791.

Murugan, A., Zeravcic, Z., Brenner, M.P., and Leibler, S. (2015). Multifarious assembly mixtures: systems allowing retrieval of diverse stored structures. Proc. Natl. Acad. Sci. USA *112*, 54–59.

Muyldermans, S. (2013). Nanobodies: natural single-domain antibodies. Annu. Rev. Biochem. *82*, 775–797.

Needleman, D., and Dogic, Z. (2017). Active matter at the interface between materials science and cell biology. Nat. Rev. Mater. *2*, 17048.

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*, 5945.

Nielsen, A.A.K., Der, B.S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E.A., Ross, D., Densmore, D., and Voigt, C.A. (2016). Genetic circuit design automation. Science *352*, aac7341.

O'Brien, P.J., Luo, W., Rogozhnikov, D., Chen, J., and Yousaf, M.N. (2015). Spheroid and tissue assembly via click chemistry in microfluidic flow. Bioconjug. Chem. *26*, 1939–1949.

Ortiz, M.E., and Endy, D. (2012). Engineered cell-cell communication via DNA messaging. J. Biol. Eng. *6*, 16.

Payne, S., Li, B., Cao, Y., Schaeffer, D., Ryser, M.D., and You, L. (2013). Temporal control of self-organized pattern formation without morphogen gradients in bacteria. Mol. Syst. Biol. *9*, 697.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: machine learning in python. J. Mach. Learn. Res. *12*, 2825–2830.

Piñero-Lambea, C., Bodelón, G., Fernández-Periáñez, R., Cuesta, A.M., Álvarez-Vallina, L., and Fernández, L.A. (2015). Programming controlled adhesion of *E. coli* to target surfaces, cells, and tumors with synthetic adhesins. ACS Synth. Biol. *4*, 463–473.

Ramaswamy, S. (2010). The mechanics and statistics of active matter. Annu. Rev. Condensed Matter Phys. *1*, 323–345.

Rokas, A. (2008). The origins of multicellularity and the early history of the genetic toolkit for animal development. Annu. Rev. Genet. *42*, 235–251.

Salema, V., Marín, E., Martínez-Arteaga, R., Ruano-Gallego, D., Fraile, S., Margolles, Y., Teira, X., Gutierrez, C., Bodelón, G., and Fernández, L.Á. (2013). Selection of single domain antibodies from immune libraries displayed on the surface of *E. coli* cells with two β-domains of opposite topologies. PLoS ONE *8*, e75126.

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

Santiago, C., Björling, E., Stehle, T., and Casasnovas, J.M. (2002). Distinct kinetics for binding of the CD46 and SLAM receptors to overlapping sites in the measles virus hemagglutinin protein. J. Biol. Chem. *277*, 32294–32301.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Scholes, N.S., and Isalan, M. (2017). A three-step framework for programming pattern formation. Curr. Opin. Chem. Biol. *40*, 1–7.

Shetty, R.P., Endy, D., and Knight, T.F., Jr. (2008). Engineering BioBrick vectors from BioBrick parts. J. Biol. Eng. *2*, 5.

Sia, S.K., Gillette, B.M., and Yang, G.J. (2007). Synthetic tissue biology: tissue engineering meets synthetic biology. Birth Defects Res. C Embryo Today *81*, 354–361.

Steinberg, M.S. (1963). Reconstruction of tissues by dissociated cells. Some morphogenetic tissue movements and the sorting out of embryonic cells may have a common explanation. Science *141*, 401–408.

Tabor, J.J., Salis, H.M., Simpson, Z.B., Chevalier, A.A., Levskaya, A., Marcotte, E.M., Voigt, C.A., and Ellington, A.D. (2009). A synthetic genetic edge detection program. Cell *137*, 1272–1281.

Tamsir, A., Tabor, J.J., and Voigt, C.A. (2011). Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'. Nature *469*, 212–215.

Teague, B.P., Guye, P., and Weiss, R. (2016). Synthetic morphogenesis. Cold Spring Harb. Perspect. Biol. *8*, a023929.

Temme, K., Zhao, D., and Voigt, C.A. (2012). Refactoring the nitrogen fixation gene cluster from Klebsiella oxytoca. Proc. Natl. Acad. Sci. USA *109*, 7085–7090.

Thévenaz, P., Ruttimann, U.E., and Unser, M. (1998). A pyramid approach to subpixel registration based on intensity. IEEE Trans. Image Process. *7*, 27–41.

Toda, S., Blauch, L.R., Tang, S.K.Y., Morsut, L., and Lim, W.A. (2018). Programming self-organizing multicellular structures with synthetic cell-cell signaling. Science. https://doi.org/10.1126/science.aat0271.

Todhunter, M.E., Jee, N.Y., Hughes, A.J., Coyle, M.C., Cerchiari, A., Farlow, J., Garbe, J.C., LaBarge, M.A., Desai, T.A., and Gartner, Z.J. (2015). Programmed synthesis of three-dimensional tissues. Nat. Methods *12*, 975–981.

Tsai, J.C., Yen, M.R., Castillo, R., Leyton, D.L., Henderson, I.R., and Saier, M.H., Jr. (2010). The bacterial intimins and invasins: a large and novel family of secreted proteins. PLoS ONE *5*, e14403.

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

Wagner, S., Baars, L., Ytterberg, A.J., Klussmeier, A., Wagner, C.S., Nord, O., Nygren, P.A., van Wijk, K.J., and de Gier, J.W. (2007). Consequences of membrane protein overexpression in Escherichia coli. Mol. Cell. Proteomics *6*, 1527–1550.

Whitesides, G.M., and Grzybowski, B. (2002). Self-assembly at all scales. Science *295*, 2418–2421.

Younger, D., Berger, S., Baker, D., and Klavins, E. (2017). High-throughput characterization of protein-protein interactions by reprogramming yeast mating. Proc. Natl. Acad. Sci. USA *114*, 12166–12171.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Living Colors Full-Length GFP polyclonal Rabbit Antibody	Clontech	CAT #632592; RRID: AB_2336883
RhodamineRed-Xconjugated AffiniPure Goat AntiRabbit IgG H+L	Jackson ImmunoResearch	CAT #111-295-144; RRID: AB_2338028
Goat anti-GAPDH polyclonal antibody	Genscript	CAT #A00191-40; RRID:AB_914662
MonoRab Rabbit anti-Camelid VHH mAb	Genscript	CAT #A01860-200; RRID: AB_2734123
IRDye 680RD Donkey anti-Goat	Li-Cor	CAT #925-68074; RRID: AB_2650427
IRDye 800CW Donkey anti-Rabbit Ig(H+L)	Li-Cor	CAT #925-32213; RRID: AB_2715510
Bacterial and Virus Strains		
MG1655	Coli Stock Genome Center (CGSC)	CGSC #6300
S1	Coli Stock Genome Center (CGSC)	CGSC #6338
Chemicals, Peptides, and Recombinant Proteins		
Recombinant GFP protein	Clontech	CAT #632373
EPEA peptide (> 95% purity, sequence: EPEA)	Genscript	N/A
PEAE peptide (> 95% purity, sequence: PEAE)	Genscript	N/A
Critical Commercial Assays		
RNEasy Minikit with RNAprotect Bacteria Reagent	QIAGEN	CAT #74104, CAT #76506
SuperScript III Platinum SYBR Green One-Step qRT-PCR kit	Thermo-Fisher	CAT #11736059
Oligonucleotides		
NeaeD0_F (sequence: AAGCGGATAACGCCGATAC)	Jin and Riedel-Kruse, 2018	N/A
NeaeD0_R (sequence: GCACTGGCATTAAGTGCTGA)	Jin and Riedel-Kruse, 2018	N/A
16S_F280 (sequence: CGATCCCTAGCTGGTCTGAG)	This paper	N/A
16S_R511 (sequence: GTTAGCCGGTGCTTCTTCTG)	This paper	N/A
Recombinant DNA		
pDSG290 (pSB4A3_TetR_pTet_Neae2v1_antiP53TA)	This paper	GenBank: MH492375
pDSG291 (pSB4A3_TetR_pTet_Neae2v1_Null)	This paper	GenBank: MH492376
pDSG289 (pSB4A3_TetR_pTet_Neae2v1_antiEPEA)	This paper	GenBank: MH492377
pDSG288 (pSB4A3_TetR_pTet_Neae2v1_P53TA)	This paper	GenBank: MH492378
pDSG287 (pSB4A3_TetR_pTet_Neae2v1_EPEA)	This paper	GenBank: MH492379
pDSG312 (pSB3K3_AraC_pBAD_B0034_Neae2v1_antiEPEA)	This paper	GenBank: MH492380
pDSG310 (pSB3K3_AraC_pBAD_B0034_Neae2v1_A4_EPEA)	This paper	GenBank: MH492381
pDSG263 (pSB4A3_sfGFP)	This paper	GenBank: MH492382
pDSG262 (pSB3K3_mRuby2)	This paper	GenBank: MH492383
pDSG264 (pSB3K3_sfGFP)	This paper	GenBank: MH492384
pDSG6 (pSB1C3_Cerulean)	This paper; from Nate Cira	GenBank: MH492385
pDSG7 (pSB1C3_Venus)	This paper; from Nate Cira	GenBank: MH492386

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pDSG261 (pSB4A3_mRuby2)	This paper	GenBank: MH492387
pDSG8 (pSB1C3_mCherry)	This paper; from Nate Cira	GenBank: MH492388
pDSG323 (pSB3K3_TetR_pTet_Neae2v1_Null)	This paper	GenBank: MH492374
pDSG253 (pNeae2_N4-1_antiEPEA)	This paper	GenBank: MH492389
pDSG254 (pNeae2_A4_EPEA)	This paper	GenBank: MH492390
pDSG320 (pSB3K3_TetR_pTet_Neae2v1_antiEPEA)	This paper	GenBank: MH492391
pDSG318 (pSB3K3_TetR_pTet_Neae2v1_EPEA)	This paper	GenBank: MH492392
pDSG321 (pSB3K3_TetR_pTet_Neae2v1_antiP53TA)	This paper	GenBank: MH492393
pDSG319 (pSB3K3_TetR_pTet_Neae2v1_P53TA)	This paper	GenBank: MH492394
pDSG385 (pSB3K3_TetR_pTet_Neae2v1_N7-6_antiP53NLSRegHis6-R2PNR63)	This paper	GenBank: MH492395
pDSG403 (pSB3K3_TetR_pTet_Neae2v1_N8-6_antiP53TA-R4P8)	This paper	GenBank: MH492396
pDSG420 (pSB3K3_TetR_pTet_Neae2v1_A5_gelsolinFAF)	This paper	GenBank: MH492397
pDSG372 (pSB3K3_TetR_pTet_Neae2v1_N3-1_antiAkt3PH-3AKH13)	This paper	GenBank: MH492398
pDSG395 (pSB3K3_TetR_pTet_Neae2v1_N7-16_antiP53NLSRegHis6-R2PNR41)	This paper	GenBank: MH492399
pDSG364 (pSB3K3_TetR_pTet_Neae2v1_N2-1_antiAkt1PHMUT-R2AKM3)	This paper	GenBank: MH492400
pDSG361 (pSB3K3_TetR_pTet_Neae2v1_N1-1_antiAkt1PH-1AIP24)	This paper	GenBank: MH492401
pDSG400 (pSB3K3_TetR_pTet_Neae2v1_N8-3_antiP53TA-R4P25)	This paper	GenBank: MH492402
pDSG374 (pSB3K3_TetR_pTet_Neae2v1_N3-3_antiAkt3PH-4AKH8)	This paper	GenBank: MH492403
pDSG382 (pSB3K3_TetR_pTet_Neae2v1_N7-3_antiP53NLSRegHis6-R2PNR29)	This paper	GenBank: MH492404
pDSG373 (pSB3K3_TetR_pTet_Neae2v1_N3-2_antiAkt3PH-3AKH59)	This paper	GenBank: MH492405
pDSG411 (pSB3K3_TetR_pTet_Neae2v1_N8-14_antiP53TA-R3P36)	This paper	GenBank: MH492406
pDSG359 (pSB3K3_TetR_pTet_Neae2v1_A7_P53NLSRegHis)	This paper	GenBank: MH492407
pDSG386 (pSB3K3_TetR_pTet_Neae2v1_N7-7_antiP53NLSRegHis6-R2PNR84)	This paper	GenBank: MH492408
pDSG402 (pSB3K3_TetR_pTet_Neae2v1_N8-5_antiP53TA-R3P3)	This paper	GenBank: MH492409
pDSG404 (pSB3K3_TetR_pTet_Neae2v1_N8-7_antiP53TA-R3P28)	This paper	GenBank: MH492410
pDSG421 (pSB3K3_TetR_pTet_Neae2v1_A6_melanophilin)	This paper	GenBank: MH492411
pDSG408 (pSB3K3_TetR_pTet_Neae2v1_N8-11_antiP53TA-R3P90)	This paper	GenBank: MH492412
pDSG419 (pSB3K3_TetR_pTet_Neae2v1_A4_EPEA)	This paper	GenBank: MH492413
pDSG412 (pSB3K3_TetR_pTet_Neae2v1_N8-15_antiP53TA-R3P16)	This paper	GenBank: MH492414
pDSG401 (pSB3K3_TetR_pTet_Neae2v1_N8-4_antiP53TA-R3P9)	This paper	GenBank: MH492415
pDSG378 (pSB3K3_TetR_pTet_Neae2v1_N6-2_antimelanophilin-2MEL24)	This paper	GenBank: MH492416
pDSG365 (pSB3K3_TetR_pTet_Neae2v1_N2-2_antiAkt1PHMUT-R2AKM38)	This paper	GenBank: MH492417
pDSG362 (pSB3K3_TetR_pTet_Neae2v1_N1-2_antiAkt1PH-1AIP26)	This paper	GenBank: MH492418
pDSG363 (pSB3K3_TetR_pTet_Neae2v1_N1-3_antiAkt1PH-2AIP1)	This paper	GenBank: MH492419
pDSG383 (pSB3K3_TetR_pTet_Neae2v1_N7-4_antiP53NLSRegHis6-R2PNR77)	This paper	GenBank: MH492420
pDSG371 (pSB3K3_TetR_pTet_Neae2v1_N2-8_antiAkt1PHMUT-R2AKM29)	This paper	GenBank: MH492421
pDSG410 (pSB3K3_TetR_pTet_Neae2v1_N8-13_antiP53TA-R3P91)	This paper	GenBank: MH492422
pDSG346 (pSB3K3_TetR_pTet_Neae2v1_sfGFP)	This paper	GenBank: MH492423
pDSG358 (pSB3K3_TetR_pTet_Neae2v1_A3_Akt3PH)	This paper	GenBank: MH492424
pDSG376 (pSB3K3_TetR_pTet_Neae2v1_N5-1_antigelsolinFAF-FAF1)	This paper	GenBank: MH492425
pDSG381 (pSB3K3_TetR_pTet_Neae2v1_N7-2_antiP53NLSRegHis6-R2PNR34)	This paper	GenBank: MH492426
pDSG388 (pSB3K3_TetR_pTet_Neae2v1_N7-9_antiP53NLSRegHis6-R3PNR32)	This paper	GenBank: MH492427
pDSG406 (pSB3K3_TetR_pTet_Neae2v1_N8-9_antiP53TA-R3P6)	This paper	GenBank: MH492428
pDSG392 (pSB3K3_TetR_pTet_Neae2v1_N7-13_antiP53NLSRegHis6-R3PNR33)	This paper	GenBank: MH492429
pDSG339 (pSB3K3_TetR_pTet_Neae2v1_antiGFP)	This paper	GenBank: MH492430
pDSG397 (pSB3K3_TetR_pTet_Neae2v1_N7-18_antiP53NLSRegHis6-R3PNR70)	This paper	GenBank: MH492431
pDSG366 (pSB3K3_TetR_pTet_Neae2v1_N2-3_antiAkt1PHMUT-R2AKM5)	This paper	GenBank: MH492432

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pDSG377 (pSB3K3_TetR_pTet_Neae2v1_N6-1_antimelanophilin-2MEL35)	This paper	GenBank: MH492433
pDSG393 (pSB3K3_TetR_pTet_Neae2v1_N7-14_antiP53NLSRegHis6-R2PNR68)	This paper	GenBank: MH492434
pDSG356 (pSB3K3_TetR_pTet_Neae2v1_A1_Akt1PH)	This paper	GenBank: MH492435
pDSG407 (pSB3K3_TetR_pTet_Neae2v1_N8-10_antiP53TA-R3P1)	This paper	GenBank: MH492436
pDSG357 (pSB3K3_TetR_pTet_Neae2v1_A2_Akt1PHMUT)	This paper	GenBank: MH492437
pDSG379 (pSB3K3_TetR_pTet_Neae2v1_N6-3_antimelanophilin-2MEL5)	This paper	GenBank: MH492438
pDSG409 (pSB3K3_TetR_pTet_Neae2v1_N8-12_antiP53TA-R3P10)	This paper	GenBank: MH492439
pDSG375 (pSB3K3_TetR_pTet_Neae2v1_N4-1_antiEPEA-EPEA1)	This paper	GenBank: MH492440
pDSG380 (pSB3K3_TetR_pTet_Neae2v1_N7-1_antiP53NLSRegHis6-R2PNR16)	This paper	GenBank: MH492441
pDSG367 (pSB3K3_TetR_pTet_Neae2v1_N2-4_antiAkt1PHMUT-R2AKM2)	This paper	GenBank: MH492442
pDSG389 (pSB3K3_TetR_pTet_Neae2v1_N7-10_antiP53NLSRegHis6-R3PNR29)	This paper	GenBank: MH492443
pDSG396 (pSB3K3_TetR_pTet_Neae2v1_N7-17_antiP53NLSRegHis6-R3PNR83)	This paper	GenBank: MH492444
pDSG369 (pSB3K3_TetR_pTet_Neae2v1_N2-6_antiAkt1PHMUT-R2AKM18)	This paper	GenBank: MH492445
pDSG387 (pSB3K3_TetR_pTet_Neae2v1_N7-8_antiP53NLSRegHis6-R3PNR91)	This paper	GenBank: MH492446
pDSG405 (pSB3K3_TetR_pTet_Neae2v1_N8-8_antiP53TA-R3P18)	This paper	GenBank: MH492447
pDSG360 (pSB3K3_TetR_pTet_Neae2v1_A8_P53TA)	This paper	GenBank: MH492448
pDSG390 (pSB3K3_TetR_pTet_Neae2v1_N7-11_antiP53NLSRegHis6-R3PNR82)	This paper	GenBank: MH492449
pDSG370 (pSB3K3_TetR_pTet_Neae2v1_N2-7_antiAkt1PHMUT-R2AKM27)	This paper	GenBank: MH492450
pDSG368 (pSB3K3_TetR_pTet_Neae2v1_N2-5_antiAkt1PHMUT-R2AKM33)	This paper	GenBank: MH492451
pDSG394 (pSB3K3_TetR_pTet_Neae2v1_N7-15_antiP53NLSRegHis6-R3PNR26)	This paper	GenBank: MH492452
pDSG398 (pSB3K3_TetR_pTet_Neae2v1_N8-1_antiP53TA-R4P43)	This paper	GenBank: MH492453
pDSG399 (pSB3K3_TetR_pTet_Neae2v1_N8-2_antiP53TA-R3P12)	This paper	GenBank: MH492454
pDSG384 (pSB3K3_TetR_pTet_Neae2v1_N7-5_antiP53NLSRegHis6-R2PNR30)	This paper	GenBank: MH492455
pDSG391 (pSB3K3_TetR_pTet_Neae2v1_N7-12_antiP53NLSRegHis6-R3PNR15)	This paper	GenBank: MH492456
Software and Algorithms		
FIJI (Fiji Is Just ImageJ)	Schindelin et al., 2012	Version 1.51r
StackReg image plugin	Thévenaz et al., 1998	N/A
Imaris	Bitplane	Version 8.0.2
NearestNeighbors_patternquantification.py (Data S1)	This paper	N/A

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ingmar H. Riedel-Kruse (ingmar@stanford.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Strains and sequences

The two parent strains used in this study, and which were obtained from the Coli Genetic Stock Center (CGSC), are MG1655 (CGSC #6300) and S1 (CGSC #6338). The intimin display system was originally obtained as an IPTG-inducible high copy (pMB1 origin) plasmid termed pNeae2 (and anti-GFP derivative termed pNVgfp) from Luis Ángel Fernández (Piñero-Lambea et al., 2015). The Tet-expression and Ara-expression plasmids were constructed by synthesizing (using IDT's gBlock service) a restriction site-free version of the intimin N terminus and cloning via BioBrick suffix assembly into iGEM part numbers BBa_K145279 and BBa_I0500, respectively, or into pNeae2 for filamentous strains. Note that, in contrast to the high-copy expression of pNeae2, medium-copy expression of the adhesin constructs had limited effect on the shape of MG1655 cells (Figures 4H and 4I). Sequences for the nanobody/antigen library were obtained from the VIB Nanobody Core. These were then either synthesized as IDT gBlocks and incorporated into the Tet and Ara expression plasmids via BioBrick suffix assembly (adapted for fusions) or both synthesized and cloned into the Tet expression plasmid by Twist Bioscience. A full list of the nanobody/antigen sequences is available in Table S1. Fluorescent protein sequences were obtained from S. DePorter (mRuby2), P. Subsoontorn (sfGFP) or N. Cira (Venus, mCherry, Cerulean) and cloned into various constitutive expression plasmids (pSB1C3, pSB3K3, pSB4A3 with pλ or BBa_J23100, BBa_B0034 expression) by BioBrick assembly (see Table S1).

Culture conditions

Cultures for aggregation assays were grown at 37°C while shaking at 300 rpm in LB media + 100 ng/mL ATc (if induced and unless noted otherwise) for 24 hours to ensure stationary phase and consistent final density across samples. For the growth-phase aggregates in Figure 5, cultures were grown for 16 hours, backdiluted 1:1000, and grown for an additional 2 hours shaken at 37°C before mixing for aggregation.

METHOD DETAILS

Aggregation assays

Cultures were grown overnight at 37°C while shaking at 300 rpm in 7 mL LB + 100 ng/mL ATc (if induced and unless noted otherwise) for 24 hours to ensure stationary phase and consistent final density across samples. Filamentous morphology was accomplished by expressing the adhesin construct on the high-copy plasmid pNeae2 using 100 μM isopropyl β-D-1-thiogalactopyranoside (IPTG). Such high expression of membrane proteins has been shown to induce filamentous growth (Wagner et al., 2007). Cultures were then vortexed briefly and mixed 1:1 with other strains in deep 96-well plates at room temperature. Samples of 100 μL were taken from the mixtures immediately following mixture and 24 hours later, to ensure equilibrium, from the top ∼25% of the well ("supernatant"). Samples were transferred to 96-well assay plates and OD_{600} was measured on a Tecan infinite M1000 plate reader.

Peptides

EPEA and PEAE peptides were synthesized by Genscript at >95% purity. The lyophilized peptides were resuspended in water, and their concentration was quantified on a NanoDrop One using the A205/31 method.

Aggregation time lapses

Cultures were grown and mixed as above, and then transferred to 10 mL clear plastic test tubes, taped to a black felt background with an overhead fluorescent lamp for a dark field effect. Samples were photographed on a Nexus 5X smartphone using the TimeLapse Video Recorder app. Quantification was done in FIJI (Schindelin et al., 2012) by subtracting grayscale values of the upper one third of the test tubes minus neighboring test tubes in the same image containing only media.

Microscopy

Epifluorescence was performed on a Leica DMI6000B microscope using the GFP and TX2 filter sets, along with brightfield images, and a 40x 0.6 NA objective. Confocal microscopy was performed on a Leica DMRXE microscope using a 63x 1.2 NA water-immersion objective with excitation of 488 nm for sfGFP, 496 nm for Venus, 543 nm for mCherry or mRuby2, and 458 nm for Cerulean. Images for Figures S4 and S5 were obtained on a Zeiss LSM 880 confocal microscope at the Beckman Cell Sciences Imaging Facility using 40x 1.3 NA and 63x 1.4 NA oil-immersion objectives, respectively. For these, excitations wavelengths were 488 nm for sfGFP and 594 for mRuby2. Emission ranges for confocal were manually adjusted to maximize signal and avoid bleed-through. All confocal images were taken after allowing 600 – 1200 μL mixtures to settle in 1.5 mL microcentrifuge tubes for approximately 1 hour at room temperature. For each sample, approximately 20 μL of aggregate was extracted from the bottom of the tube using a wide-orifice pipette tip, transferred to a double-sided tape microscope slide chamber, covered with a coverslip, and sealed with Thomas Lubriseal stopcock grease. Sections varying from approximately 20 – 200 μm were imaged by confocal microscopy. Bleed-through from blue to green channels was corrected in 3-color images by subtracting the blue channel from the green channel. For display purposes, a 2-pixel median filter was applied to images in Figures 4, 5, and 6, and brightness/contrast were adjusted in FIJI for entire images to assure channels appear similar.

Microscopic time lapse

Overnight cultures were grown for 16 hours, backdiluted 1:1000, grown for 2 hours shaken at 37°C, mixed 1:1 (1.2 mL total), and allowed to settle within a 1.5 mL test tube at 37°C for an additional 2 hours. Using a sterile syringe, ∼100 μL were slowly transferred from the middle of the tube to a microfluidic device made from a layer of polydimethylsiloxane (PDMS) over a glass coverslip, containing an inlet, outlet, and 2 mm × 12 mm × 0.1 mm chamber (Lam et al., 2017). The chamber was connected using sterilized steel pins and tubing to two reservoirs of media (3 mL and 2.9 mL) and imaged using epifluorescence in a humidified, 37°C chamber every 5 minutes for 6 hours. Brightness and contrast were automatically adjusted in FIJI for each frame to help identify cells. StackReg ImageJ plugin (Thévenaz et al., 1998) was used to maintain orientation of cells in Figure 5.

Porosity quantification

To measure porosity (fraction of volume not occupied by cells), 3D confocal stacks of aggregates were processed in FIJI as follows. First, 8 z slices were selected from the center of the aggregate. A 2-pixel median filter was used, and then the images were thresholded using Phansalkar auto local threshold with default parameters. This was done separately for each color channel, and then the color channels were summed. "Analyze particles" was then used to exclude particles smaller than $1 \mu m^2$. The porosity is reported as one minus the average value of pixels in the thresholded, channel-summed, 3D image stack.

Immunostaining

Equal cell numbers (as calculated by volume × $OD_{600} \approx 220\mu L \cdot OD_{600}$) of each sample were spun down at 6000 rpm for 1 min and resuspended in 100 μL phosphate buffered saline (PBS) + 0.5% bovine serum albumin (BSA). Cells were then either left as is, mixed with either 1 μL Clontech Living Colors Full-Length GFP polyclonal Rabbit Antibody (cat #632592), or mixed with 1 mg/mL Clontech recombinant GFP protein (cat #632373) and incubated at room temperature for 30 min under aluminum foil. Samples were then washed 3 times by spinning down at 6000 rpm for 1 min and resuspending in 100 μL PBS + 0.5% BSA. For samples treated with GFP, the cells were resuspended the final time without BSA. For those treated with antiGFP, samples were subsequently mixed with 1 μL Jackson ImmunoResearch RhodamineRed-Xconjugated AffiniPure Goat AntiRabbit IgG H+L (cat #111-295-144) and incubated at room temperature for 30 min under aluminum foil. Samples were then washed 3 times again in PBS + 0.5% BSA, with the last resuspension leaving out the BSA. Samples were then transferred to microscope slides and imaged as above.

Protein extraction

Whole cell protein extracts were used for polyacrylamide gel electrophoresis (PAGE). Samples of 100 μL of 24 hr overnight culture were centrifuged and resuspended in 150 μL of a 95°C mixture of resuspension buffer, formulated as 750 μL 8x sodium dodecyl sulfate (SDS) in glycerol and bromothymol blue, 750 μL 850 mM 1,4-dithiothreitol (DTT), and 4.5 mL MECB buffer + 0.5% NP40 buffer. Samples were stored at −80°C.

Western blotting

Protein extract samples (5 μL) were loaded onto 4 − 20% acrylamide gradient gels and run for ∼ 20 min at 250V in SDS running buffer (14.4 mg/L glycine, 3 g/L Tris, 1 g/L SDS in water). Gels were then transferred to a blotting apparatus and run at 100V at 4°C in transfer buffer (formulated as 15.15 g/L tris, 72 g/L glycine, 20% methanol) for 1 hr. Blots were washed for 20 min in TBS (tris-buffered saline, formulated as 9.7 g/L tris 70.14 g/L NaCl pH 7.5) + 0.1% tween-20 + 5% nonfat dry milk (TBS-TM). Blots were then incubated on a rotor in 3 mL TBS-TM + 3 μL goat anti-GAPDH (Genscript #A00191-40) + 3 μL rabbit anticamelid VHH (Genscript #A01860-200). Blots were then washed twice in TBS + 0.1% tween-20 (TBS-T) for 5 min, then once in TBS-TM for 10 min. Blots were then incubated in TBS-TM + 1:15,000 donkey anti-goat 680 nm fluorescent secondary antibody (Li-Cor #925-68074) + 1:15,000 donkey anti-rabbit 800 nm fluorescent secondary antibody (Li-Cor #925-32213). Blots were finally washed three times in TBS-T, rinsed with 1x TBS without tween, and imaged on a LiCOR infrared scanner.

qRT-PCR

Total RNA was extracted from 40 μL of stationary, 24-hour cultures using QIAGEN RNEasy Minikit with RNAprotect Bacteria Reagent protocol 4 (using proteinase K) and on-column DNase treatment. Quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) was performed in 20 μL reactions on a 96-well plate using the SuperScript III Platinum SYBR Green One-Step qRT-PCR kit (Thermo-Fisher) spiked with 10 nM fluorescein for calibration. Thermocycling was carried out on a Bio-Rad iCycler (using primers NeaeD0_F, NeaeD0_R, 16S_F280, and 16S_R511 listed in the Key Resources Table), which reported the threshold cycle (C_T) values used here.

Nearest neighbor quantification

Confocal z-stacks were analyzed using the Bitplane Imaris software package after subtracting the blue channel from the green channel to remove bleed-through on 3-color images, and individual cell centroids were identified using the "surfaces" tool. No filtering or brightness/contrast adjustment was performed in advance. Centroids were analyzed using a custom python script and the scipy NearestNeighbors API (Jones et al., 2001) (see Data S1). For microscopic spherical cell aggregates, cell movement between z-stacks precluded automated centroid detection, so centroids were detected in 2D slices either by Imaris as above or by multi-tracker point tool in FIJI, following which the centroids were quantified as above. For Figure 4B, data from 3 confocal stacks were each sub-sampled for 5% of the centroids, and the 3 sets of data were pooled for both true and randomized colors. Counts of nearest neighbor colors were used as expected (randomized) and true (observed) values in a χ^2-test with 2 degrees of freedom. Distances in 3D aggregates were measured in voxels, rather than microns; analysis using microns does not substantially alter the results.

Hierarchical clustering of probability tables

Clustering was done using the scikit-learn hierarchy library (Pedregosa et al., 2011) with default parameters. Each probability table was converted into a 9×1 (A) or 4×1 (B) vector for the analysis. For differential adhesion and bullseye patterns, an arbitrary set of 3 replicates was chosen out of those available; the clustering holds true for all 6 (differential adhesion) and 9 (bullseye) replicates.

Definition of conditional probability tables

Conditional probability tables were presented in Figures 4C and 6A–6D, including theoretical tables in Figures 6 and S6. Here a conditional probability table refers to a matrix of probability values with each entry $p_{ij} = P(j$ is closest neighbor to $i)$ giving the probability for cell type i to have a nearest neighbor cell of type j. That is, each row references a single cell type i (green ($i = G$), red ($i = R$), or blue ($i = B$)), and the columns correspond to different conditions (that cell j, which is the closest neighbor to i, is green ($j = G$), red ($j = R$), or

blue (j = B)). In Figure 4C, for example, p_{ij} is high for $i \neq j$ and low for $i = j$, indicating that both green and red cells bind most closely to cells of the opposite color. Note that by these definitions each row must always sum to 1.

Heuristic expectation for probability tables
In Figure 6, expected probability tables were calculated using a heuristic estimate based on pairwise interaction data available in Figures 2, 3, and 4. This includes data on adhesion strength as a function of inducer concentration (Figure 3A), the two-cell heterophilic probability table (Figure 4C), and the density variation data (Figure 4H). We state the heuristic estimate first, and then discuss the reasoning for each term in the heuristic. For simplicity, cell positions are analyzed based on their centroids, and each cell has only one nearest neighbor.

Heuristic equation probability tables
The heuristic estimate for the probabilities is given by

$$p_{ij} = \frac{\rho_{j:i}(K_{ij} + \kappa_{ij})}{N_i} \tag{S1}$$

where the variables in this equation are defined as follows:

- p_{ij}, as defined above, is the probability for cell type i to have a cell of type j as its nearest neighbor.
- $\rho_{j:i}$, as used in Figures 4 and 6, is the density ratio of cell type j to cell type i.
- K_{ij}, as used in Figure 6, is the binding strength between cell types i and j (normalized so that a direct Nb-Ag binding strength equals 1 at maximum induction of 100 ng/mL ATc).
- κ_{ij} is an apparent binding strength between two cells i and j that do not bind directly via Nb-Ag interaction, but neighbor one another by both binding a third cell m. From a practical imaging standpoint, this causes i and j to neighbor one another in processed images, sometimes more closely than either neighbors m. Thus i and j appear to be bound despite the fact that they lack any physical adhesion K_{ij}.
- $N_i = \sum_j \rho_{j:i}(K_{ij} + \kappa_{ij})$ is a normalization factor chosen so that $\sum_j p_{ij} = 1$ to make the values of p_{ij} legitimate probabilities (i.e., that each row of the p_{ij} matrix sums to 1).

Reasoning behind the heuristic equation
There are three factors we need to explain in Equation 1: ρ, K, and κ. The factor N_i is required by the definition of p_{ij} as a probability.

The fact that p_{ij} is directly proportional to ρ comes from the data in Figure 4H, where for a ratio between $\rho_{R:G} = 1$ and $\rho_{R:G} = 6$ the number of nearest red neighbors scales approximately linearly. That is, for $\rho_{R:G} = 1$ the number of nearest neighbors is approximately 1, for $\rho_{R:G} = 2$ the number of nearest neighbors is approximately 2, and so on. We reasoned that the probability of having a cell as a nearest neighbor scales directly with the number of nearest neighbors, and thus we arrived at the proportionality to ρ in Equation 1. Note that the colon in the subscript is used to point out that $\rho_{j:i} \neq \rho_{i:j}$ (in fact, $\rho_{j:i} = 1/\rho_{i:j}$). Also, as an aside, note that the nearest neighbors curve in Figure 4H saturates at $\rho_{R:G} = 6$. Were we interested in mixing ratios greater than 6 we would therefore take $\rho = 6$ rather than the actual mixing ratio. This situation is not relevant to the cases discussed here.

The fact that p_{ij} is directly proportional to K comes from the data in Figure 3A, where a decreased binding strength left more cells unaggregated. We reasoned that with fewer cells bound, we should expect fewer nearest neighbors on average, and thus a proportionally lower probability, giving the proportionality in Equation 1. Although we did not take direct measurements of the physical binding strength (e.g., in nM), we used the data from Figure 3A to estimate relative binding strength. We took the data from Figure 3A and normalized the maximum expected binding strength (induction at 100 ng/mL ATc) to 1 for convenience since the equation is normalized anyway by N. This normalized, or relative, value is then taken as our definition for the binding strength K. Note that $K_{ij} = K_{ji}$ by definition. Note also that based on Figure 2C, there appears to be little noticeable difference in binding strength between different Nb-Ag pairs, and so we assumed that all such pairs have equivalent maximum binding strength for this analysis.

The κ term is required, because we know from Figure 4C that even with no direct binding interaction between cell types i and j, p_{ij} will not equal zero. This is in fact expected, since if i and j both bind a third cell m, then i and j may be positioned directly next to one another despite no direct Nb-Ag binding event between i and j. In analyzing the image data, i and j would then appear as neighbors, thus appearing to have an apparent binding strength that is non-zero. If i and j are in separate phases (such as green and blue cells in Figure 6B) or otherwise expected to always be >1 cell length apart (such as red and green cells in Figure 6D), then $\kappa = 0$ because the two cells can never be neighbors. Ultimately, the value for this parameter when non-zero is empirical, and we estimated it from the data in Figure 4C (see below). Note that $\kappa_{ij} = \kappa_{ji}$ by definition as with K_{ij}, and we retain the subscript to point out that for some interactions κ is zero and for others it is a non-zero parameter.

Estimating κ from Figure 4C
The experimental values in Figure 4C are:

	G	R
G	0.08	0.92
R	0.81	0.19

We followed the heuristic above by first stating the known values for ρ and K. All the densities are approximately equal, so $\rho_{G:G} = \rho_{R:R} = \rho_{R:G} = \rho_{G:R} = 1$. The two cell types are at maximum induction, so $K_{GR} = K_{RG} = 1$, and there is no direct binding between like cells, so $K_{GG} = K_{RR} = 0$. Meanwhile, κ is non-zero in all interactions, since all cells are within the same phase (e.g., a green cell can neighbor another green cell when both bind to a third red cell). Thus, we arrived at the following estimate:

$$p_{ij} = \begin{pmatrix} \kappa/N & (1+\kappa)/N \\ (1+\kappa)/N & \kappa/N \end{pmatrix}. \tag{S2}$$

Equating this matrix with the table of values from above, averaging the two κ cases, and remembering that the rows must sum to one, we solved for κ and N:

$$\frac{\kappa}{N} \approx (0.08 + 0.19)/2 = 0.13 \tag{S3}$$

$$\frac{2\kappa + 1}{N} = 1.$$

This yields $N \approx 1.35$ and, more importantly,

$$\kappa \approx 0.18. \tag{S4}$$

We used this value of κ in all further estimates. This value could theoretically be different for spherical and rod-shape cells, or for cells that swim more or less vigorously. We were only concerned with rough predictions here, however, so we did not take into account such potential variabilities.

Note that since we used the data from Figure 4C to determine κ, it is no longer appropriate to compare the raw probabilities from Figure 4C to the expected probabilities using the value $\kappa = 0.18$. Therefore, where we compared various conditional probability tables to one another via hierarchical clustering in Figure S6, we used another dataset to represent the two-cell type heterophilic pattern. That is, we used Ag2/Venus and Nb2/mCherry cells like those used in Figure 6B (these are slightly different than those in Figure 4C in using Venus and mCherry instead of sfGFP and mRuby2) and mixed them. The conditional probability tables from the resulting aggregates is termed "Figure 4C equivalent" and was used in Figure S6.

Calculation of expected probability tables
We used the heuristic given above to determine the expected conditional probability tables for Figure 6. The conditional probability tables derived below are displayed in Figure 6 under the "Expectation" column.
Figure 6A expectation
Here the cells are all present in equal ratios, so $\rho_{j:i} = 1$ always. No cells are expected to be in totally separate phases, so $\kappa = 0.18$ is always non-zero. Red cells are induced maximally (100 ng/mL ATc) and green cells at $10^{-4}\%$ Ara. Reading the data off of Figure 3A, we see that the green-green binding strength is then about 12% of the red-red binding strength. Thus, $K_{RR} = 1$ and $K_{GG} = 0.12$. We did not measure values directly for K_{GR}. However, we estimated this value from the composability data in Figure S2. In particular, there appears to be full titration of Nb2 within Ag2/Nb2 strains, which prevents any aggregation with Ag2 strains (Figure S2, first plot, row 3, column 1), but Ag2/Nb2 strains can still fully aggregate with Nb2 strains using what remains of the displayed Ag2 (Figure S2, third plot, row 3, column 1). Thus with the smaller amount of available displayed Ag2 these cells still mediate aggregation similar in magnitude to full, untitrated expression when mixed with fully expressing Nb2 cells. This suggests that in binding events between two strains with different numbers of adhesins, the adhesion tends to be similar to the more strongly adhering strain. Given that, we will assume as a rough estimate that $K_{GR} \approx 0.8$.

Plugging these values of ρ, K, and κ into Equation 1, we find that $N_G = 1.28$ and $N_R = 2.16$. Altogether, this yields the following conditional probability table:

	G	R
G	0.23	0.77
R	0.45	0.55

Figure 6B expectation
In this case, $\rho_{G:G} = \rho_{G:R} = \rho_{R:G} = \rho_{R:R} = \rho_{B:B} = 1$, $\rho_{B:G} = \rho_{B:R} = 2$, and $\rho_{R:B} = \rho_{G:B} = 1/2$. The binding constants are $K_{GR} = K_{BB} = 1$ and zero for all other cases (remembering that $K_{ij} = K_{ji}$). These values can be read off of the schematic in Figure 6B. Meanwhile, $\kappa_{BG} = \kappa_{BR} = 0$ and $\kappa = 0.18$ otherwise (again, remembering that $\kappa_{ij} = \kappa_{ji}$).

Plugging these values of ρ, K, and κ into Equation S1, we find that $N_G = N_R = 1.36$ and $N_B = 1.18$. Altogether, this yields the following conditional probability table:

	G	R	B
G	0.13	0.87	0
R	0.87	0.13	0
B	0	0	1.0

Figure 6C expectation

In this case, $\rho_{G:G} = \rho_{G:R} = \rho_{R:G} = \rho_{R:R} = \rho_{B:B} = 1$, $\rho_{B:G} = \rho_{B:R} = 2$, and $\rho_{R:B} = \rho_{G:B} = 1/2$ as in the previous case. The binding constants are $K_{GB} = K_{RB} = 1$ and zero for all other cases (remembering that $K_{ij} = K_{ji}$. These values can be read off of the schematic in Figure 6C. Meanwhile, $\kappa = 0.18$ for all cases because all cells can neighbor one another indirectly.

Plugging these values of ρ, K, and κ into Equation S1, we find that $N_G = N_R = 2.72$ and $N_B = 1.36$. Altogether, this yields the following conditional probability table:

	G	R	B
G	0.066	0.066	0.868
R	0.066	0.066	0.868
B	0.434	0.434	0.132

Figure 6D expectation

In this case, $\rho_{G:G} = \rho_{R:R} = \rho_{B:B} = 1$, $\rho_{G:B} = \rho_{B:R} = 6$, $\rho_{B:G} = \rho_{R:B} = 1/6$, $\rho_{G:R} = 36$, and $\rho_{R:G} = 1/36$. The binding constants are $K_{GB} = K_{RB} = 1$ and zero for all other cases as in the previous case (remembering that $K_{ij} = K_{ji}$). These values can be read off of the schematic in Figure 6D. Meanwhile, $\kappa = 0.18$ for all cases as in the previous case except for $\kappa_{RG} = \kappa_{GR} = 0$. This is because the sequential addition sequesters the red cells behind a shell of blue cells, and so green and red cells are always expected to be >1 cell length apart.

Plugging these values of ρ, K, and κ into Equation S1, we find that $N_G = 0.377$, $N_R = 7.26$, and $N_B = 7.457$. Altogether, this yields the following conditional probability table:

	G	R	B
G	0.478	0	0.522
R	0	0.025	0.975
B	0.950	0.026	0.024

QUANTIFICATION AND STATISTICAL ANALYSIS

Number of samples, definition of values (e.g., means) and error bars (e.g., standard deviations) are described in the figure captions. The t tests were performed in Excel and are similarly described in the figure captions. The normality assumption for t tests was checked by either Shapiro-Wilk test ($p > 0.05$ showing inability to reject normality) for each set of samples, or by visual inspection in a QQ-plot (for Figure S4I, where the large number of data points makes Shapiro-Wilk less effective). The χ^2-test was applicable to Figure 4B due to the large number of independent measurements on cell centroids, and a 5% sub-sampling of the centroids was performed to ensure that assayed centroids were spatially distant and thus likely uncorrelated.

DATA AND SOFTWARE AVAILABILITY

The nearest-neighbor quantification script that generated all experimental conditional probability tables (see Figures 4C, 6, S5, and S6) is provided in Data S1. All software code and all other data not present in the main text or the supplements is freely available upon request.
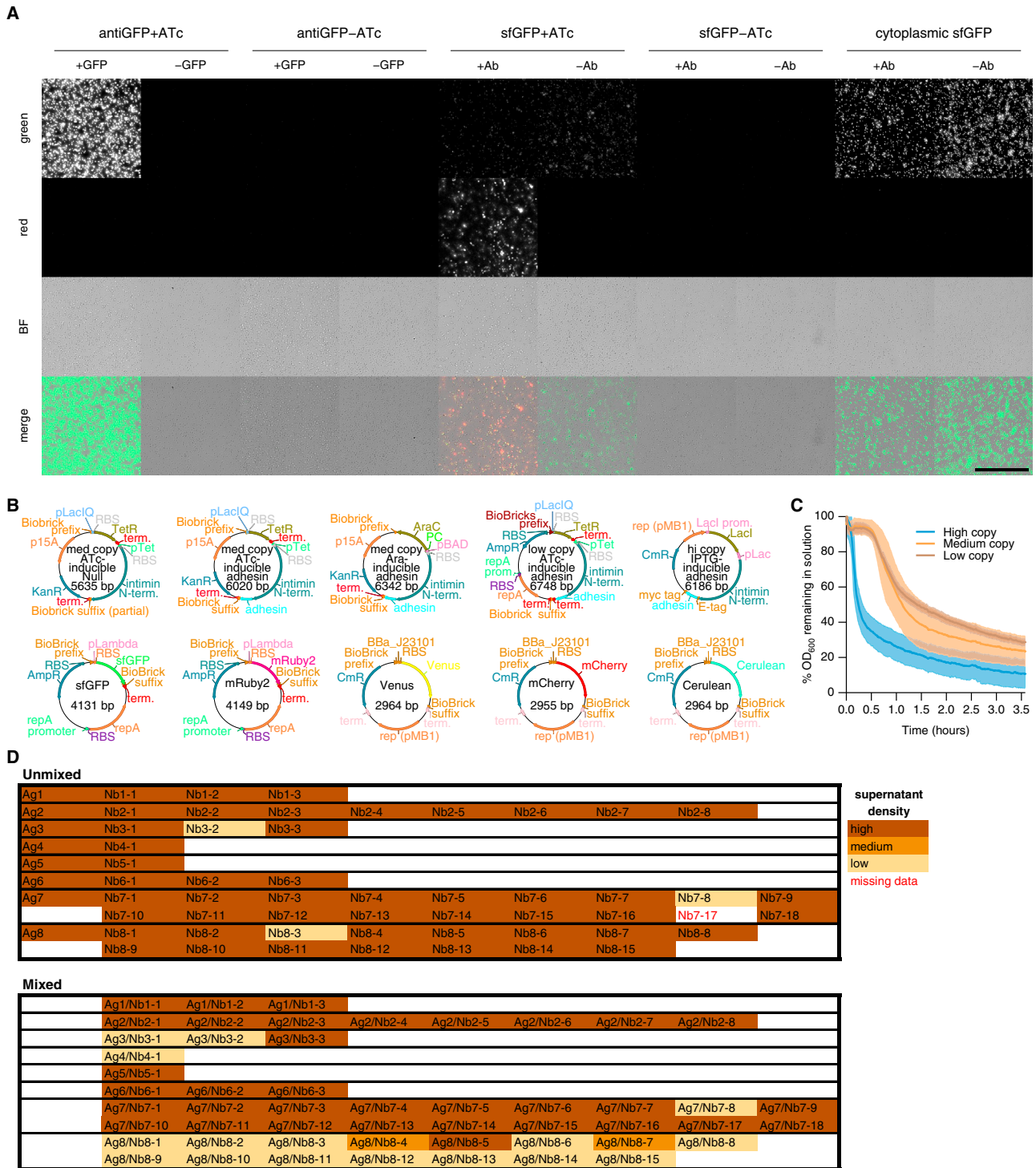
Supplemental Figures

Figure S1. A Library Screen of 8 Antigens and 52 Corresponding Nanobodies Resulted in the Three Nb-Ag Pairs Featured in the Main Text, Related to Figure 2

(A) The first Nb-Ag pair that we attempted to implement in the design from Figure 1 was a commonly available antiGFP-GFP pair (Piñero-Lambea et al., 2015). Immunostaining of cells producing intimin fusions to GFP or antiGFP demonstrated that each adhesin in this pair could bind its partner in solution, but showed poor antigen expression or folding and no adhesion. Vertical labels: color channel. Topmost horizontal labels: cell types (producing antiGFP, GFP or cytoplasmic GFP, and with or without induction by ATc for the first two cell types). Lower horizontal label: stain (either soluble GFP or anti-GFP antibody with red-fluorescent secondary antibody). Scale bar: 50 μm. Surface-displayed antiGFP and GFP were both stained by their respective binding partners, but cytoplasmically produced

(legend continued on next page)

GFP was not. This indicates successful surface display. These strains did not mediate adhesion, however (data not shown), and surface display of GFP gives very weak green signal. The native extracellular (C-terminal) portion of intimin discluded here is only slightly larger than GFP, and consists of mostly Ig-like repeats (Tsai et al., 2010), leading us to reason that GFP's large, complex structure might not have trafficked effectively or folded properly at the membrane. We thus focused on antigens less than the size of a nanobody ($\leq$125 amino acids). Note that a potential downside of small antigens is that fusion to intimin likely restricts the orientation of binding interactions, which we reasoned might interfere with adhesion. For example, one of the 8 antigens (termed Ag2 below) is only 4 amino acids long, and a known crystal structure (De Genst et al., 2010) shows the C-terminal carboxyl bound to an arginine within the nanobody binding pocket. Were this antigen fused to the N terminus of intimin, as would be the case with a non-inverse autotransporter (Salema et al., 2013), adhesion would likely be impossible. A small antigen also might have its nanobody-recognized epitope blocked close to the membrane to due steric effects. A naive estimate might therefore indicate that around 50% of antigens tested would have their epitopes hidden due to orientation and thus fail to mediate adhesion. This led to our decision to screen a library of small antigens and their corresponding nanobodies. (B) Maps for plasmids used in this work. Top row: adhesin construct plasmids. Bottom row: fluorescence plasmids. (C) Time courses of aggregation under low-, medium-, and high-copy plasmid expression show aggregation in $\leq$1 hour. Ag- and Nb-expressing cell types were mixed and allowed to settle, with OD quantified in FIJI by comparing the upper one-third of the tube to neighboring blank (LB media) tubes. Low and medium-copy expression yielded similar aggregation (mostly aggregated within ~1 hour). High copy expression caused cells to fall out of solution substantially faster (mostly aggregated within ~20 min). Averages are for n = 3 replicate tubes, error regions are $\pm$1 SD. Cells were mixed at time 0. (D) Initial qualitative aggregation assays for the full adhesin library. Nb3-2, Nb7-8, and Nb8-3 showed promiscuous binding, potentially to some bacterial surface protein, and were not examined further. Nb7-17 did not mediate adhesion in mixture, so little information is lost because of the missing data for the unmixed condition. Note that for antigens mediating successful adhesion, most or all corresponding nanobodies bind successfully, and likewise antigens that do not mediate adhesion do not mediate adhesion for any nanobody. This suggests that the limiting factor in finding working adhesion pairs has to do with the antigen (otherwise one would expect little correlation between a given antigen working with one nanobody versus another). For simplicity, elsewhere in the text the successfully aggregating strains Ag3, Nb3-1, Ag4, Nb4-1, Ag8, and Nb8-1 are referred to as Ag1, Nb1, Ag2, Nb2, Ag3, and Nb3, respectively. In comparing among the Nb8 variants, Nb8-1 through Nb8-15 are referred to as Nb3-1 through Nb3-15 (Figures 3B and S3E), with Nb3-1 equivalent to Nb3. Under this updated naming convention, Nb3-3 and Nb3-5 are not analyzed further as they failed to specifically mediate aggregation. All sequence data is presented in Table S1.
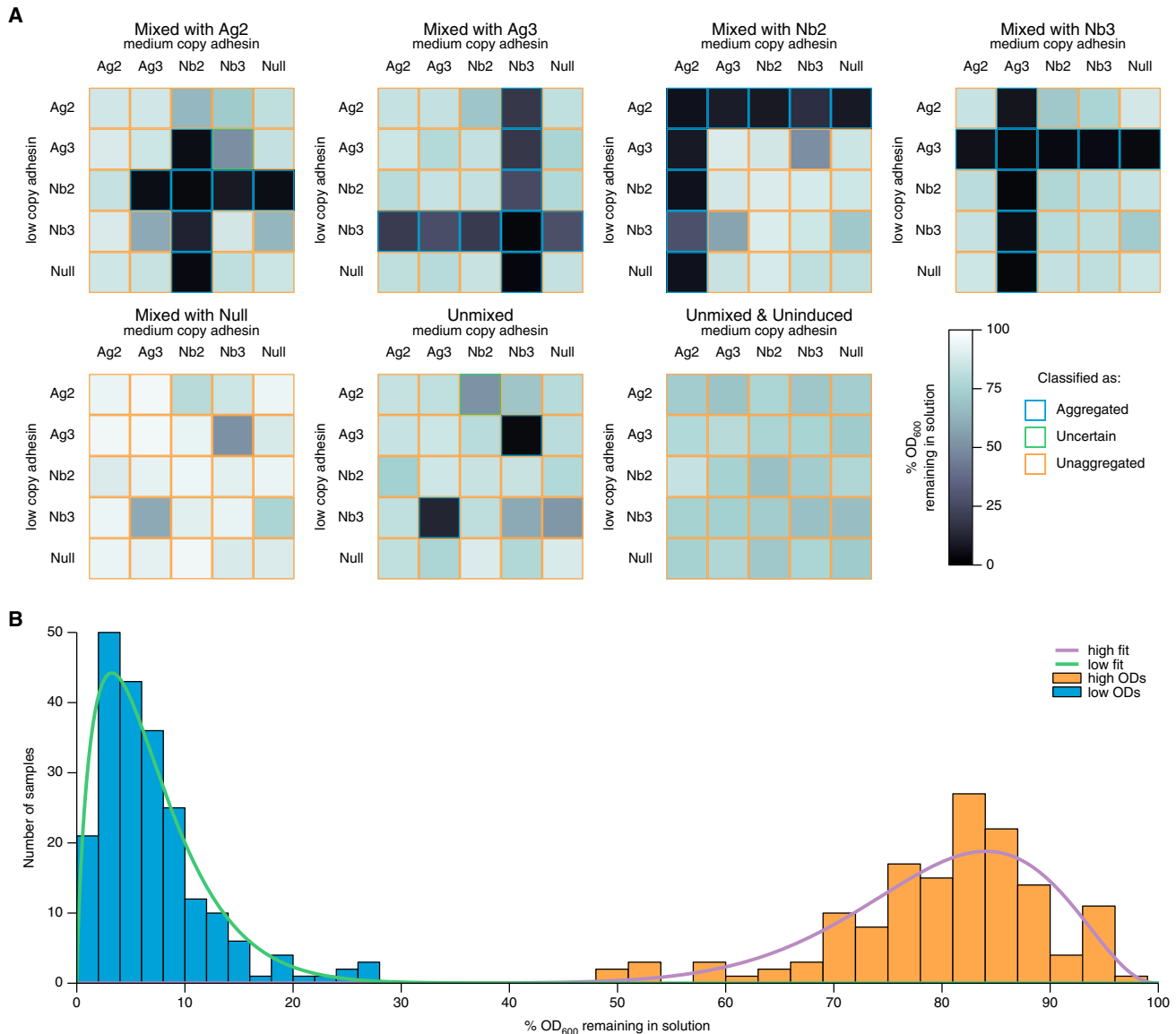
Figure S2. Full, Detailed Dataset Corresponding to Figure 2E in the Main Text, Related to Figure 2

(A) Twenty-five strains containing all permutations of 5 adhesins (Ag2, Ag3, Nb2, Nb3, Null) on medium-copy plasmid (top axis) and low-copy plasmid (left axis) were each measured by the aggregation assay (Figure 2A) under 7 conditions (panel title): in mixture with strains producing a single medium-copy adhesin construct (Ag2, Ag3, Nb2, Nb3, Null corresponding to the first 5 plots) and alone (unmixed) either with or without induction by ATc (final 2 plots). For example, in the upper left plot, row 2, column 3 shows that a strain producing Ag3 on a low copy plasmid and Nb2 on a medium copy plasmid aggregates when mixed with a strain producing Ag2 on a medium-copy plasmid. These data show that aggregation only occurs under conditions when both an Ag and its corresponding Nb are present. Note that, as expected, this includes the Ag3/Nb3 strains when unmixed (dark entries in second to last plot), demonstrating that a homophilically adhesive strain can be achieved with a single cell producing both Ag and Nb within a single cell. By contrast, the expectedly equivalent Ag2/Nb2 strains do not aggregate when unmixed (second to last plot, row 3, column 1 and row 1, column 3). In fact, they do not aggregate even when mixed with an Ag2 strain (upper left plot, row 3 column 1 and row 1, column 3). However, they do aggregate when mixed with Nb2 (third plot, row 3 column 1 and row 1, column 3), indicating that there is enough available Ag2 to bind Nb2, but no available Nb2. We term this a potential *cis* titration effect, since Ag2 seems to titrate away Nb2 from use for aggregation (which requires binding in *trans* with Ag2 on the surface of another cell). That is, unbound Ag2 on the Ag2/Nb2-producing cells is available for use in adhesion, whereas no or little Nb2 is available. This matches the idea that Ag2 and Nb2 bind to one another in *cis* on the surface of Ag2/Nb2. This assymetry is probably not due to differences in expression between Ag2 and Nb2 in Ag2/Nb2-producing cells, as can be seen from the fact that a medium-copy Ag2 and low-copy Nb2 (upper left panel, row 3 column 1) behaves the same as a low-copy Ag2 and medium-copy Nb2 (upper left panel, row 1 column 3). It is also unclear why this titration effect occurs with the Ag2/Nb2 pair and not with the Ag3/Nb3 pair. There is a strong asymmetry in the sizes of the antigen and nanobody in the Ag2/Nb2 pair (4 and 126 amino acids, respectively), which is not present in the Ag3/Nb3 pair (102 and 111 amino acids, respectively). We thus speculate that this size asymmetry somehow gives rise to the *cis* titration effect. The colors outlining each entry of each plot indicate whether the data for that condition was classified as aggregating or non-aggregating (see B). Heatmap for n = 3 averages. (B) Data points from (A) split into two well-defined clusters of "aggregated"

and "unaggregated," with a clear gap separating the two. Classification of these two-adhesin interaction data was done using an *ad hoc* predictor. All of the 175 conditions plotted in (A) were combined in a single histogram, shown here. There is a clear gap in the histogram between what conditions have aggregated (low %) and those that have not (high %). The data were split accordingly into datasets above and below 40%. Beta distribution fits to each of these two classes are plotted here as well. Each of the 175 conditions (with their triplicate measurements) were tested against the 99th percentile of the low distribution and the 1st percentile of the high distribution with a 1-tailed t test (using the triplicate measurements to determine the sample standard deviation). Conditions with $p < 0.05$ for one but not the other were classified as part of the corresponding group. Others are unclassified. The results of the classifier are contained in (A).
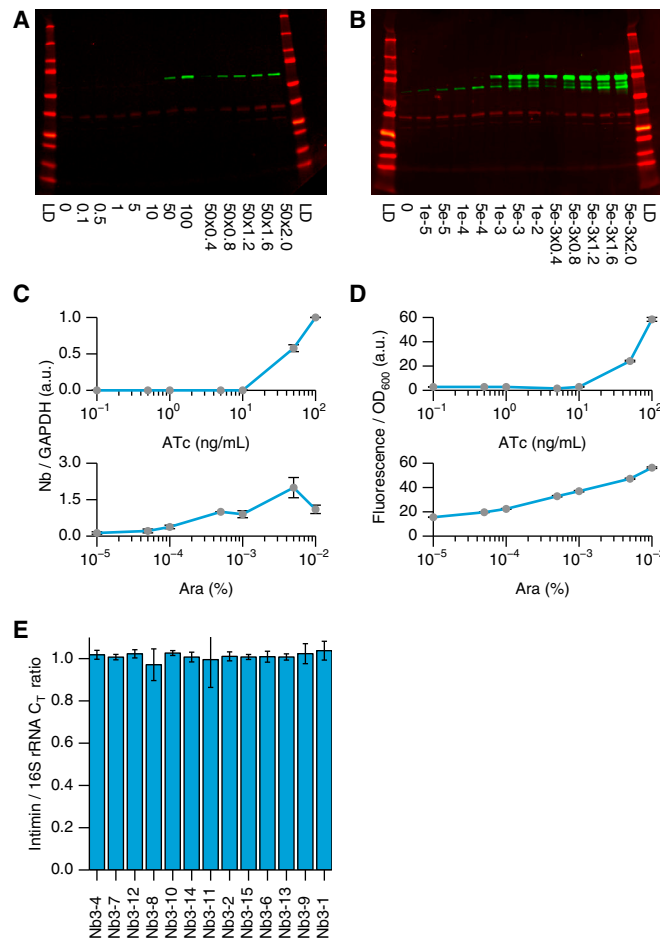
Figure S3. Quantification of Adhesion Construct Expression Demonstrates the Validity of Aggregation Assay Measurements, Related to Figure 3

(A) Western blot of ATc induction curve. Lanes are marked by the ATc concentration in ng/mL with an additional loading volume used for calibration standards (e.g., 50×0.4 = 50 ng/mL ATc with 0.4 × 5 μL = 2 μL loading volume; 1 × = 5 μL loading volume implicit where not reported). Lanes marked as "LD" are protein ladders (Bio-rad #161-0374). Staining is red for GAPDH (major band at expected 36 kDa) and green for Nb (expected band around 81 kDa). The major band runs fast, around 65 kDa, with partially and fully unfolded bands observed in high expression, as has been reported elsewhere (Salema et al., 2013). (B) Same as (A), but for Ara induction curve, with Ara concentration reported in % w/v. (C) Quantification of (A) and (B), using the lower Nb band, normalized for each replicate to 100 ng/mL ATc sample (top) or to 5×10^{-4} % Ara (bottom). This band correlates with adhesion measured by aggregation in Figure 3A. For Ara induction above 10^{-3}%, intense upper bands outside of the quantifiable range may represent overly high expression that interferes in some way with aggregation (although the cells do not become filamentous as with high-copy expression). Due to the shift in intensity from the lower band to the upper band at high Ara induction, we speculate that there may be some differences in intimin processing at very high expression. Practically, this implies that Ara concentration should not exceed 10^{-3} for induction. (D) Expression of cytoplasmically produced GFP under the same inducer systems as (C), showing similar dynamics except in the case of high induction by Ara. Note maximum expression matches at 0.01% Ara and 100 ng/mL ATc (samples are comparable, as they were measured in the same experiment). (E) Expression levels are consistent across the Nb3 variants tested in Figure 3B as measured by qRT-PCR cycle threshold of intimin normalized to that of 16S ribosomal RNA. This indicates that differences in aggregation are due to intrinsic affinity rather than expression differences. Samples are displayed in the same order as in Figure 3B. Displayed values are for n = 3 PCR runs of both intimin and 16S measurements, other than variants 8–15, which had only 2 replicates for the 16S measurements due to issues with the edge of the PCR plate. Error bars: ±1 SD.
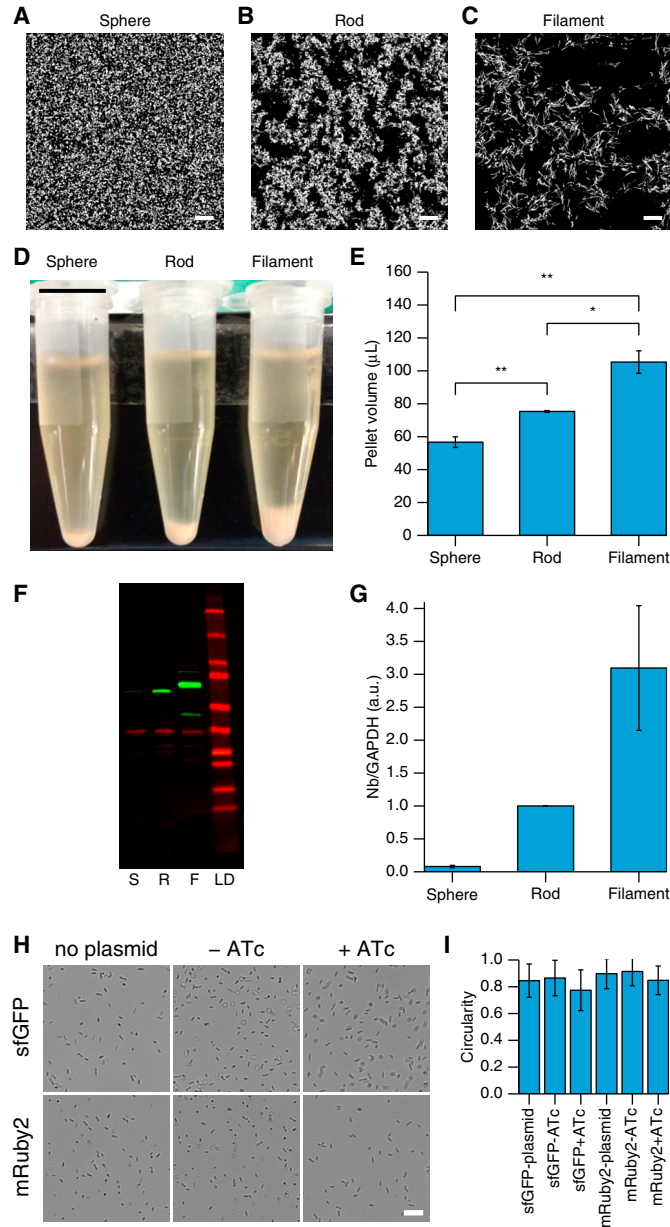
Figure S4. Quantification of Macroscopic Aggregates Demonstrates Differing Aggregate Structure, Matching Microscopic Estimates of Porosity in Figure 4F, Related to Figure 4

(A, B, C) Representative Z-projection of $212.5\mu m \times 212.5\mu m \times 6.4\mu m$ ($x \times y \times z$) confocal stacks of (A) spherical, (B) rod-shaped, and (C) filamentous bacteria used in quantifying microscopic porosity in Figure 4F. Scale bars: 20 μm. (D) Spherical, rod-shaped, and filamentous cell types aggregate into macroscopic pellets of different sizes. Scale bar: 1 cm. (E) Quantification of (D). These data match the microscopic porosity measurements of the main text, in that the majority of the volume in filament pellets does not contain any cells, rod pellets have a smaller volume fraction free of cells, and spherical cells pack most densely. *p-value*s are 0.0019 (sphere-filament), 0.0081 (sphere-rod), and 0.016 (filament-rod). (F) Western blot of expression in the three cell types (S: sphere, R: rod, F: filament, LD:ladder as in Figures S3A and S3B). Note that the filamentous cell's construct is 2.73 kDa longer due to the presence of myc and E tags. Additionally, a weak upper band corresponding to the fully unfolded protein and a lower band corresponding to just the extracellular portion (spacer and Nb) are visible for the filamentous sample, representing probable proteolytic cleavage. (G) Quantification of (F), including only the major band for the filamentous strain, showing very different expression levels in the three cell types. Values are normalized to the rod-shaped Nb/GAPDH value for each replicate. (H) Brightfield images of cells with constitutive production of sfGFP (top row) or mRuby2 (bottom row), and containing no adhesion plasmid (left column), uninduced adhesion plasmid (middle column), or induced adhesion plasmid (right column). (I) Quantification of (H) based on particle analysis in FIJI. Although induction significantly increases the mean cell length (decreases circularity) with $p \ll 0.01$ by t test averaged over $700 - 1500$ cells, the population is so broad that the length differences are not distinguishable for individual cells. Displayed bar-graph values are for N = 3 samples. Error bars: ± 1 SD.
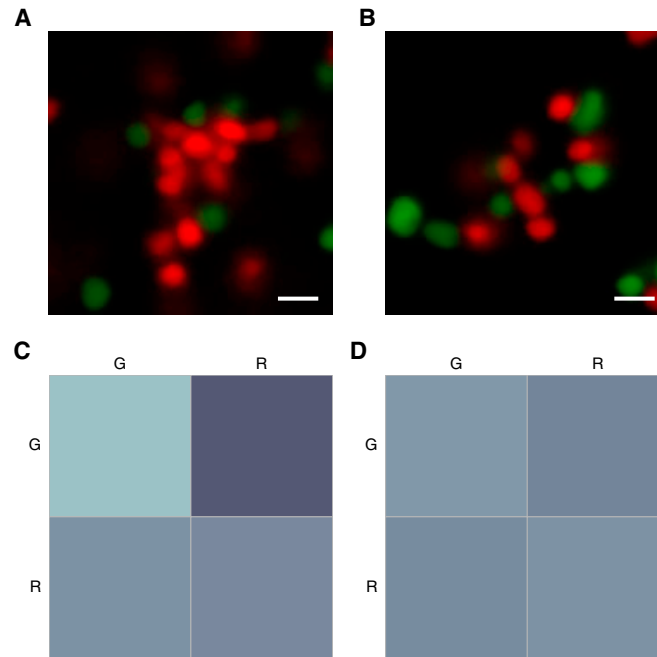
Figure S5. Full Induction of Green Cells Converts Differential Adhesion Patterning into Random Patterning, Related to Figure 6A
(A) Representative confocal image of clusters formed from mixtures of cells as in Figure 6A, with green cells induced at 10^{-5}% Ara. (B) Same as (A) except with maximal induction of Nb and Ag production in the green cells (10^{-3}% Ara), showing apparent random arrangement of red and green cells. (C,D) Conditional probability tables quantifying (A) and (B), respectively, with color scale as in Figure 4C. Scale bars: 2 μm.
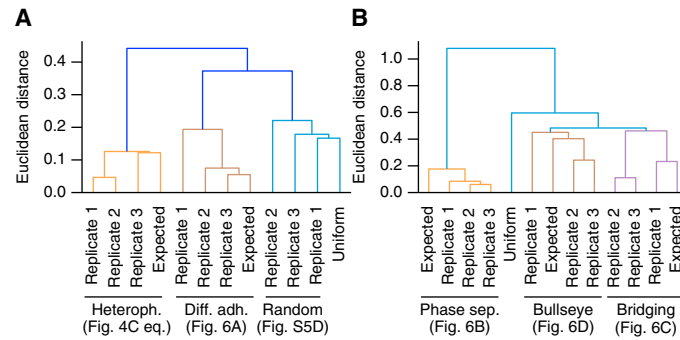
Figure S6. Hierarchical Clustering of Conditional Probability Tables Demonstrates that Experimental Probabilities Match Expectations and Are Non-random, Related to Figure 6

Hierarchical clustering for (A) 2-color and (B) 3-color patterns of conditional probability tables cluster meaningfully. Note that in both cases, a uniform conditional probability table clusters separately from the patterned samples, but clusters with the random patterns formed in Figures S5B and S5D. Most importantly, the theoretically expected probability tables (see STAR Methods) cluster with the experimentally determined probabilities. In (A), rather than using the data from Figure 4C directly (which is used to determine the expected probability table, see STAR Methods), we used replicates from a repeated set of experiments on similar strains (the red and green cells in Figure 6B without the blue cells present), which we termed "Figure 4C equivalent." For (B), bridging patterns (data from Figure 6C) and bullseye (data from Figure 6D), which use the same adhesin combinations, cluster nearby. Phase separation patterns (data from Figure 6B) clusters most distantly, presumably because of the large difference in probabilities of neighboring blue cells. Clustering was done using the scikit-learn hierarchy library (Pedregosa et al., 2011) with default parameters. Each probability table was converted into a 9x1 (A) or 4x1 (B) vector for the analysis. For differential adhesion and bullseye patterns, an arbitrary set of 3 replicates was chosen out of those available; the clustering holds true for all 6 (differential adhesion) and 9 (bullseye) replicates.