Electronic Supplementary Material (ESI) for Soft Matter.

Supplementary Information Appendix

This document includes:

Other supplementary materials for this manuscript include the following:

Movie S1

Supplementary information

Microscopy details:

Bright field images were collected in direct illumination (no phase). Fluorescence acquisitions were performed using either the green channel filters for GFP (Ex. 482/35, DM 506 Em. FF01-536/40). Excitation was performed using a LED box (CoolLed pE 4000). Image acquisition was performed using a Hamamatsu ORCA-R2 EMCCD camera for time-lapse acquisitions of 1344 × 1024 pixels images with 12 bits grey level depth (4096 grey levels) and captured an *xy* field of view of 330 μm × 430 μm. Bright field and fluorescence images were usually collected for 12 hours at frequencies ranging from 0.15 to 12 frames per hour. Excitation LEDs were set at a 50% power level and exposure time were 100 ms for fluorescence images acquisition.

Flow cytometry measurements:

Scattering and fluorescence dot plots are recorded for all the *E. coli* strains of interest using cell suspensions taken in their exponential growth phase. The forward scattering (FSC) threshold considered to sort bacteria from background is taken equal to 10 a.u. (valid for a given set of flow cytometer parameters defined and kept constant throughout the experiment). A fluorescence threshold equal to 6 a.u. is defined to isolate fluorescent GFP-expressing bacteria (the residual non-fluorescent objects, including debris and non-fluorescent, possibly dead cells, are removed from the analysis). The fluorescence intensity of single bacteria is expected to follow a lognormal distribution (1)S which is extracted from the whole population distribution by fitting the left part of the histogram (See Fig. S2 frames A.c and B.c) with the

function $y = \frac{A}{x\sigma\sqrt{2\pi}} e^{\frac{-(ln(x)-\mu)^2}{2\sigma^2}}$, where $A = N_{sc}.dx$ gives the number of single cells N_{sc}. σ and μ are the parameters of the lognormal distribution related to m, the mean and v, the variance of the single cells fluorescence distribution by $m = e^{\left(\mu + \frac{\sigma^2}{2}\right)}$ and $v = e^{(2\mu + \sigma^2)} e^{(\sigma^2 - 1)}$. The number of aggregates N_{ag}, is derived by subtracting N_{sc} from the total number of objects of the fluorescent population N_T. We then obtain the fraction of aggregates of the population, $[N_{ag}/(N_{sc} + N_{ag})]$. In order to take into account the aggregation number, expected to be, for each object, proportional to its mean fluorescence intensity, we defined the aggregation index as the fraction of aggregates multiplied by the mean fluorescence intensity of the aggregates, I_{ag}, which provides $i_{ag} = [N_{ag}/(N_{sc} + N_{ag})] \cdot I_{ag}$.

Diffusion coefficient determination:

The diffusion coefficient of single cells from each strain was measured by single-particle tracking with the TrackMate plugin in the Fiji software libraries. Movies were analyzed with the following parameters: estimated blob diameter, 3 µm, and threshold, 1 µm for each strain; linking max distance, 3 µm; gap-closing max distance, 3 µm; gap-closing max frame, 2; spot filtering: quality between 1.7 and 10 (allows to select single cells); track filtering process: duration of track above 40 s. For each selected trajectory, MatLab was used to calculate the Mean Square Displacement (MSD) versus time. Each MSD was fitted, on the first 20 points of the data, by a linear relation with a correlation coefficient >0.99. The diffusion coefficient *D* was calculated according to: $MSD(t) = 4Dt$ (Table S1). To determine if two means were significantly different, we performed Welsh two sample t-tests (Table S2).

Numerical simulations:

As outlined in the main text, we constructed a minimal and generic model to explain the formation of patterns in biofilms. Our model is based on the cell-cell interaction, the single-cell diffusion on the substrate, which is affected by the cell-substrate interaction, and on cell division, without any additional ingredient (see main text).

To simulate the evolution of this system, we perform agent-based simulations, in the framework of the Gillespie algorithm (2), which allows exact stochastic simulations without artificial discretization of time. We assume that detailed balance is satisfied.

Let us consider an elementary transition between an initial state i and a final state f, corresponding to a cell diffusing from one lattice site to an empty neighboring site, while all other particles remain at the same place. Detailed balance ensures that the rate r_{if} of this elementary transition is related to the rate r_{fi} of the inverse elementary transition through $\frac{r_{if}}{r_{fi}} = e^{E_i - E_f}$, where E_i denotes the total energy of the system in the initial state i, while E_f denotes the total energy of the system in the final state f, expressed in units of the thermal energy kT. More specifically, we choose a dynamics of traps, with $r_{if} = \frac{1}{\tau} e^{E_i}$, where $\frac{1}{\tau}$ denotes the rate at which an isolated cell moves to a nearest-neighbor site. Here, $\tau = \frac{r^2}{4D}$, where r represents the distance between two neighboring lattice sites, i.e. the characteristic size of a bacterium. We choose τ as our time unit.

To describe growth, we assume that cells that have no empty nearest neighbor site cannot divide, while all others divide with a division rate $\frac{1}{\tau_d}$. Upon division, the offspring cell fills one of the empty neighbor sites of the mother cell. This site is chosen at random among the available ones. Hence, an important parameter is τ_d, the division timescale, and more specifically its value in units of τ. In practice, $\tau_d >> \tau$.

Specifically, we implemented the following algorithm:

Initialization

- We draw the random initial position of each particle.

- We calculate all the elementary transition rates (i.e., all the rates for each particle to move to each of its empty nearest-neighbor sites, to divide and place progeny in an empty nearest-neighbor site) in the initial configuration of the system.

- We sum all these elementary transition rates in order to obtain the total transition rate R of the system.

First step

- We draw a random number in an exponential distribution with average 1/R: this is the time at which the first event occurs.

- We choose the event that occurs at this time, using the fact that each elementary transition with rate r occurs with a probability r/R.

- We update time, position data and transition rate data according to the event that has occurred.

This process is then iterated.

In our simulations, parameters were chosen to match experiments as closely as possible. In particular, the measured values of the diffusion coefficients *D* from Table S1 were used for each strain. Together with the characteristic size of bacteria, taken as r=3μm, which is an estimate of their diameter, this allows to estimate the diffusion timescale $\tau = \frac{r^2}{4D}$, which sets our time unit. Division time was further estimated as τ_d=156min from experiments. Because it is difficult to directly map the indexes of aggregation to the cell-cell interaction energy *E*, there is more uncertainty in our estimates of *E*. But in practice, the most important point is whether *E* is below or above the value corresponding to the phase transition, namely 1.25kT. In practice, patterns were well reproduced when *E*(F) was taken equal to 2kT, with robustness

around this value. To be consistent with aggregation indexes, we further took E(WT)=0, E(Curli)=kT and E(Ag43)=4kT for our simulations of these various strains.

Supplementary Figures

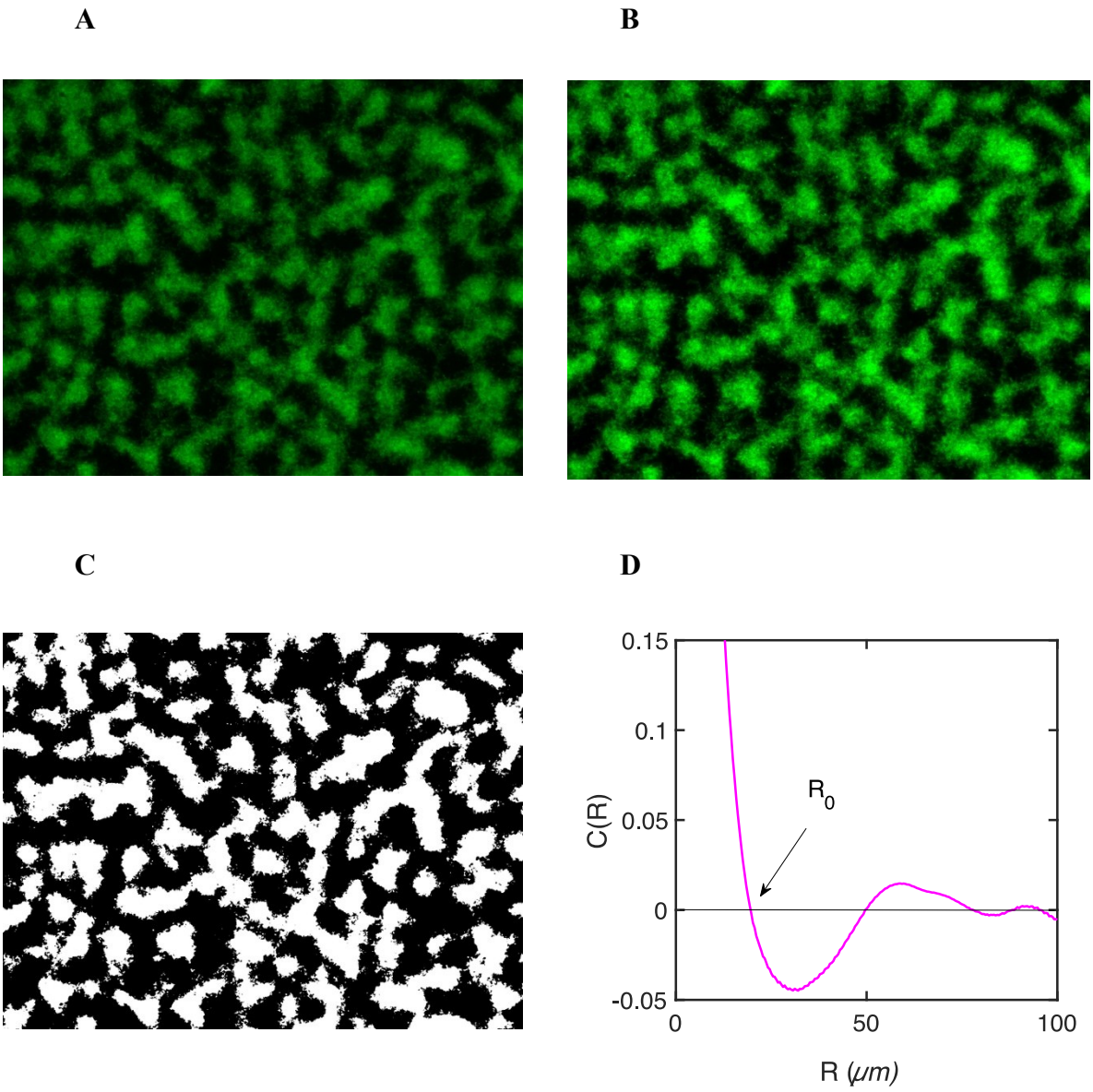

Fig. S1: Image analysis using imageJ

The initial 12-bit image is converted to 8-bit (A). Then the image auto-scaled (B). A binary image is obtained using automatic threshold menu (C). The spatial autocorrelation function of the binarized image is obtained from the 'Radially Averaged Autocorrelation' macro (D). The image size is 1344x1024 pixels or 433x330 μm^2.

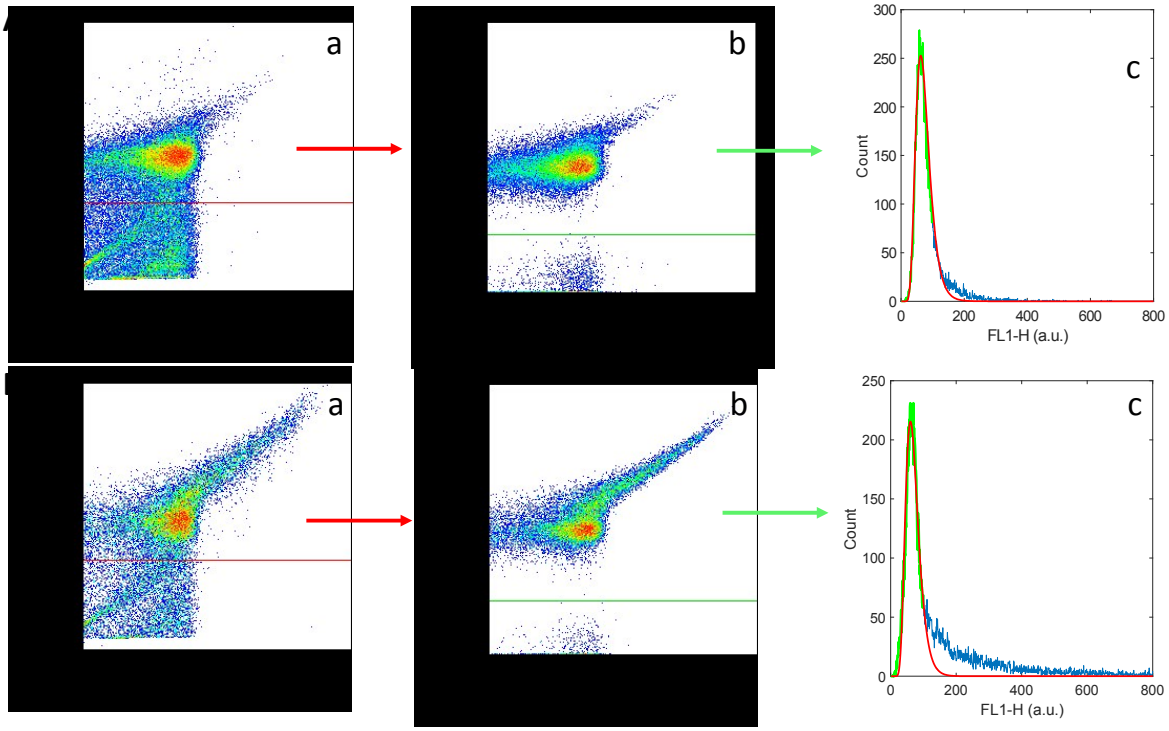

Fig. S2: Flow cytometry measurements enable aggregation index determination.

Typical scattering (a) and fluorescence (b) dot plots are shown together with fluorescence histograms (c) obtained with wild type (A) and Ag43-expressing (B) cell suspensions in the exponential growth phase. On the dot plots (a,b) the red and green lines show the side scattering (SSC) and fluorescence (FL1-H) limits of the background, 10 a.u. and 6 a.u., respectively. On the histograms (c) both experimental (in green and blue) and fitted curves (in red) are overlaid. The green curve is an extract of the blue one, marking the data set used to adjust the lognormal distribution of the single cells.

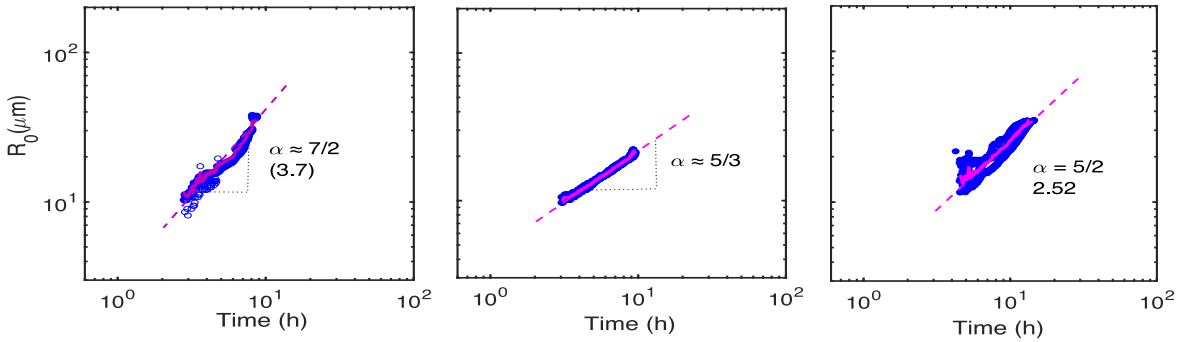

Fig. S3: Statistical dispersion of the experimental data. Coarsening data of F-expressing cell biofilms. (A), (B) and (C) show data from distinct experiments (performed on different days). The line in magenta represents the average over 3 distinct positions in the same channel (shown in blue). Although the overall behavior is well conserved from one replicate to the other, the power law exponent varies. Several experimental details may participate to these variations, such as the lag time before the cells transferred from planktonic culture to the channel resume their growth, or the manual injection into the channel, which could impact the exact number of cells that initially dwell in the channel.

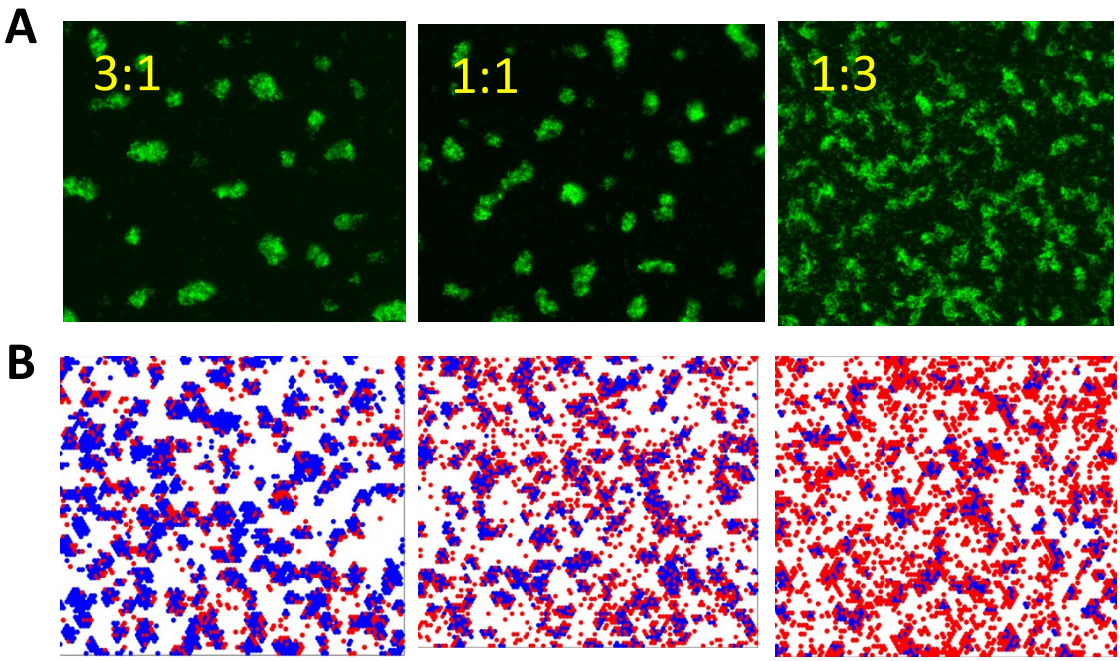

Fig. S4: Recruitment of WT cells in the pattern formed by F-expressing cells

This figure is a version of the main text Fig. 8 that shows F and WT cells in different colors (blue and red, respectively) in the simulation data (B). All other information is the same as in Fig. 8. It should be noted that in the experimental images (A) the non-labelled F cells appear in green due to their association with the fluorescent WT cells and the rather low magnification used in this experiment.

Supplementary tables:

STRAIN	Δt*	D (μm²/s)	SE**	#CELLS
F	10’	13e-3	0.2e-3	828
Ag43	15’	16.5e-3	0.9e-3	391
Wt	13’	12.1e-3	0.5e-3	366
Curli	21’	5.2e-3	0.3e-3	321

Table S1: Mean diffusion coefficients measured from single cell tracking in the early stages of biofilm formation.

* time elapsed from injection

** standard error of the mean

	F	WT	cu	Ag
F		< 2.2 e-16	**0.76**	<2.2e-16
W	<2.2e-16		<2.2e-16	1.8e-5
cu	**0.76**	<2.2e-16		<2.2e-16
Ag	<2.2e-16	1.8e-5	<2.2e-16	

Table S2: p-values calculated by Welch Two Sample t-tests. The red values indicate means that are not significantly different.

legend to supplementary movie

Movie S1: Biofilm initiates forming figures reminding spinodal decomposition of immiscible fluids. Bright field timelapse images of an F-expressing *E. coli* biofilm forming in a millifluidic channel 1mm side length continuously supplied with medium (1ml/h). A field of 353×338 μm^2 taken at the center of the channel is shown. The movie starts 2 hours after cells injection and lasts for about 6 hours (6.4 hours), 0.011 fps (90s between each image).

Supplementary information references

1. Robert L*, et al.* (2014) Division in Escherichia coli is triggered by a size-sensing rather than a timing mechanism. *BMC Biol* 12:17.
2. Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics* 22(4):403-434.