Data-Driven Models Reveal Mutant Cell Behaviors Important for Myxobacterial Aggregation

Zhaoyang Zhang,[a] Christopher R. Cotter,[b] ⓘ Zhe Lyu,[b] Lawrence J. Shimkets,[b] ⓘ Oleg A. Igoshin[a]

[a]Department of Bioengineering and Center for Theoretical Biological Physics, Rice University, Houston, Texas, USA
[b]Department of Microbiology, University of Georgia, Athens, Georgia, USA

ABSTRACT Single mutations frequently alter several aspects of cell behavior but rarely reveal whether a particular statistically significant change is biologically significant. To determine which behavioral changes are most important for multicellular self-organization, we devised a new methodology using *Myxococcus xanthus* as a model system. During development, myxobacteria coordinate their movement to aggregate into spore-filled fruiting bodies. We investigate how aggregation is restored in two mutants, *csgA* and *pilC*, that cannot aggregate unless mixed with wild-type (WT) cells. To this end, we use cell tracking to follow the movement of fluorescently labeled cells in combination with data-driven agent-based modeling. The results indicate that just like WT cells, both mutants bias their movement toward aggregates and reduce motility inside aggregates. However, several aspects of mutant behavior remain uncorrected by WT, demonstrating that perfect recreation of WT behavior is unnecessary. In fact, synergies between errant behaviors can make aggregation robust.

IMPORTANCE Self-organization into spatial patterns is evident in many multicellular phenomena. Even for the best-studied systems, our ability to dissect the mechanisms driving coordinated cell movement is limited. While genetic approaches can identify mutations perturbing multicellular patterns, the diverse nature of the signaling cues coupled to significant heterogeneity of individual cell behavior impedes our ability to mechanistically connect genes with phenotype. Small differences in the behaviors of mutant strains could be irrelevant or could sometimes lead to large differences in the emergent patterns. Here, we investigate rescue of multicellular aggregation in two mutant strains of *Myxococcus xanthus* mixed with wild-type cells. The results demonstrate how careful quantification of cell behavior coupled to data-driven modeling can identify specific motility features responsible for cell aggregation and thereby reveal important synergies and compensatory mechanisms. Notably, mutant cells do not need to precisely recreate wild-type behaviors to achieve complete aggregation.

KEYWORDS self-organization, development, modeling, *Myxococcus xanthus*, developmental biology, mathematical modeling

Development is one example of multiscale emergent behavior in which molecular interactions between cells allow self-organization into multicellular patterns. One of the most remarkable features of all types of development is how robust it is in the face of genetic and environmental perturbations, suggesting that backup systems are in place (1). While molecular genetics has identified mutations that impede multicellular development, even single mutations create downstream effects that influence multiple aspects of cell behavior and physiology. It is frequently difficult to ascertain which of the behavioral changes are deleterious to development and which can be tolerated. Here, we develop a new approach that leverages data-driven modeling to

determine whether a statistically significant trend in cell behavior results in a biologically significant alteration of the multicellular program. We demonstrate this approach by focusing on the full or partial rescue of the mutants during the multicellular development of *Myxococcus xanthus* biofilms.

Myxococcus xanthus is a rod-shaped member of the *Deltaproteobacteria* with a life cycle centered around surface motility of cells in a biofilm. *M. xanthus* has evolved multiple social mechanisms such as S-motility (2) and C-signaling (3–5) to achieve coordinated group behaviors such as predation (6), rippling (7–9), and development (7, 10, 43). Upon amino acid limitation, *M. xanthus* cells move into three-dimensional aggregates called fruiting bodies where they sporulate (11–13). Recent studies based on cell tracking have provided unprecedented detail of cell movement during development (14). In combination with mathematical modeling, these data sets unambiguously identified individual cell behaviors that are essential for aggregation (14, 15). These behaviors include reduced movement inside the aggregate and bias in the directed movement toward the aggregation centers, likely via chemotaxis (15). This methodology provides an unprecedented window into developmental behavior that is presently difficult to realize in larger organisms with thicker tissues or longer cell migration routes, such as the vertebrate neural crest, or in disease states such as tumor metastases.

In this work, we examined reciprocal interactions between wild-type (WT) cells mixed with nondeveloping mutants. More so than other bacteria, *M. xanthus* cell growth and development depend on neighboring cells, diffusing molecules, and the surrounding biotic and abiotic environment. To determine the factors that contribute to developmental robustness, we employed conditional mutants that were unable to develop on their own but will develop when mixed with WT cells. It is expected that the mutants respond to at least some of the conditions established by WT cells in the field of developing cells. The extent of the response is expected to reveal signaling and sensory transduction pathways that are essential for WT development and are defective in the mutants.

The extent of WT rescue of two mutants is examined in this work. The first of these contains a mutation in the *pilC* gene. PilC is an inner membrane protein located at the base of the pilus where it interacts with PilB and PilM to mediate pilus assembly (16, 17). This mutation interrupts pilus production (18) and consequently S-motility, one of the two motility systems in *M. xanthus* (19, 20). Aggregation can occur with the help of the A-motility system, which uses a novel molecular motor and focal adhesion complexes (21, 22). However, most S-system mutants fail to develop because they cannot produce an extracellular matrix (ECM) that is both essential for S-motility and vital for development. The ECM is required for some types of chemotaxis (23, 24) as well as for cell cohesion, which could inhibit motility inside the aggregate (25, 26). As shown in this work, *pilC* mutants cannot aggregate on their own but marginally improve when mixed with wild-type cells. The second mutation is a deletion of the *csgA* gene, which inhibits the production of one or more intercellular signals that are required for aggregation and sporulation (27). While CsgA signaling exerts control over most of development, the precise nature of the signals and their sensory pathways is only beginning to be revealed (28). *csgA* cells do not form fruiting bodies on their own (7) and respond much more completely to a WT cell developmental field than *pilC* (29). Although much is known about *M. xanthus* aggregation (7, 29–31), few quantitative data sets describe mutant cell movement during aggregation and the mechanism of their rescue (32).

To identify motility behaviors affecting mutant cell aggregation, we extended our previously developed approach that combines individual cell tracking with simulations driven by the accumulated cell behavior data (14). Directly applying experimental cell data to simulations allowed us to fully investigate the effect of each change in the mutant motility behavior on their aggregation. The results demonstrate that the WT developmental field is robust enough to nearly completely restore *csgA* development. By comparison, the *pilC* mutant has two striking sensory deficits that diminish its ability to accumulate inside the fruiting bodies. By exchanging particular aspects of cell
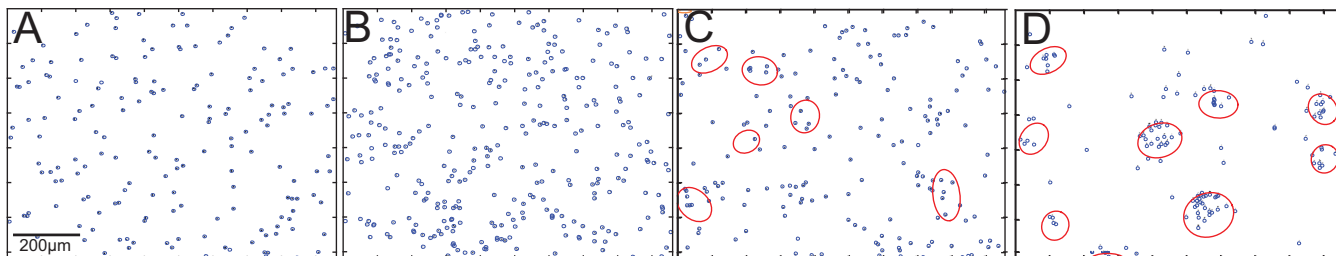
FIG 1 Cell distribution at the final frame of the experimental movies. Cells are segmented and shown as blue circles at the centroids of labeled cells. Red ellipsoids indicate the boundaries of aggregates segmented from the image after cells were filtered. (A) *pilC* cells alone. (B) *csgA* cells alone. (C) *pilC* mixed with WT cells. (D) *csgA* cells mixed with WT cells.

behavior between WT and mutant cells, our agent-based modeling was able to pinpoint specific differences in cell behavior that are most biologically significant.

RESULTS

Quantifying aggregation dynamics in mixtures of wild-type and mutant strains. Fluorescence microscopy was used to quantify the behavior of mutant cells at both single-cell and population levels. A small fraction of cells expressing the fluorescent protein tdTomato were mixed with cells expressing eYFP. Each cell expressing tdTomato is bright enough to be segmented and tracked, allowing quantification of their behaviors, whereas the weaker eYFP signal was used to quantify cell density during aggregate growth (14).

When either *pilC* or *csgA* cells are mixed with differentially labeled cells of their genotype, no aggregates are observed and the distribution of cells is nearly uniform at the final time point, i.e., at $T = 5$ h (Fig. 1A and B). Application of the 2-D Kolmogorov-Smirnov test (33) to cell positions shows that the null hypothesis of the uniform distribution of labeled cells cannot be rejected (P value > 0.95). Conversely, when tdTomato-labeled *csgA* cells are mixed with eYFP-labeled wild-type (WT) cells, *csgA* cells are overrepresented in the aggregates (Fig. 1D). The distribution of the cells is clearly nonuniform (P value < 0.05). For *pilC* cells mixed with WT cells, the rescue is less pronounced (Fig. 1C) and there is not sufficient evidence to reject the null hypothesis of uniform distribution of labeled cells (P value = 0.64). Below we describe a more sensitive metric to quantify aggregation rescue of mutant cells. As a comparison, Fig. S1 in the supplemental material shows the aggregation result of WT cells.

To quantify aggregate positions, densities, and sizes, we filtered out the tdTomato signal and then used the eYFP intensity to estimate cell density. These data were used to segment the aggregates and detect their boundaries and positions. For segmentation of the images in which aggregation was observed (mutant strains mixed with a majority of WT cells), we determined a threshold intensity that separates aggregates from the background using K-means clustering on the light intensity of each pixel in the final frame of the experimental movies. Dividing the light intensity of pixels into two clusters gives the threshold of light intensity for aggregates. Applying the same threshold throughout the sequence of time-lapse imaging, we can compare aggregate growth for different experiments. To compare the aggregation rate across different sets of experiments, we use the average aggregate size fraction, $F_{agg}(t)$, i.e., the total area of aggregates in each frame corresponding to time (t) divided by the field of view area. The results (Fig. 2A) indicate that aggregation of WT mixed with *pilC* cells is slightly slower than WT aggregation (data set from reference 14). On the other hand, WT cells mixed with *csgA* show faster aggregation. However, at the final time point, data sets lead to approximately the same area covered by aggregates, $F_{agg}(t_{final})$. Given that WT cells represent the overwhelming majority (>99.9%) of the cells, it is unlikely the observed differences are directly attributable to the presence of mutant cells. Instead, these differences are likely due to a slight variation of experimental conditions. Indeed, different biological repeats of the mixture experiments show differences in the aggre-
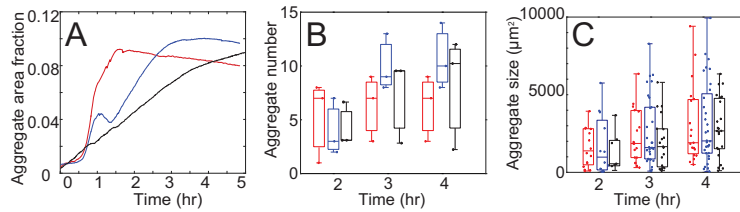
FIG 2 Aggregation variation between experiments: (A) WT cell aggregation rates vary between experiments. *y* axis is the aggregate area divided by the total area of the field of view. The red line is the average aggregation rate in experiments mixing *csgA* with WT cells. The blue line is the average aggregation rate of WT cells only (14). The black line is the average aggregation rate in experiments mixing *pilC* with WT cells. (B and C) Aggregate numbers (B) and each aggregate area (C) in experiments. Red is *csgA*, blue is WT, and black is *pilC*. Horizontal lines inside the boxes indicate the distribution median. Tops and bottoms of each box indicate 75th (q3) and 25th (q1) percentiles, respectively. Dots in panel B are aggregate number in each experiment, and dots in panel C are each aggregate area in experiments.

gation dynamics (Fig. 2B and C). Therefore, previously used metrics to characterize aggregation such as the fraction of cells within the current area of aggregates could be overly sensitive to this variability.

To quantify the distribution of the tracked cells relative to the aggregates in a way that is robust to the variability of aggregation rate, we decided to focus on the fraction of cells accumulated inside the final-frame boundaries of the aggregates. If the tracked cells were uniformly distributed, we would expect that fraction to be equal to the fraction of area covered by aggregates, i.e., $F_{agg}(t_{final})$. Therefore, to see if labeled cells are overrepresented, we focus on

$$P(t) = \frac{N_{in}(t)}{N_{tot}} - F_{agg}(t_{final}) \qquad (1)$$

Here, N_{in} is the number of tracked cells inside the final aggregate area and N_{tot} is the total number of tracked cells over the total field of view area. We do this calculation for each frame (at time *t*) and use it to quantify the aggregation rate of labeled cells.

The results for $P(t)$ quantification for aggregation of *csgA* mixed with WT (red) and *pilC* mixed with WT (black) cells are shown in Fig. 3A. To compare it with WT-only aggregation, we use a data set from reference 14 to compute the same quantity (Fig. 3A, blue line). The result shows that *csgA* has a similar aggregation rate to WT cells. In the final frame, cells inside aggregates are overrepresented by 50% of the total cell number, $P(t_{final}) \sim 0.5$. In contrast, *pilC* cells show much weaker aggregation, $P(t_{final}) \sim 0.1$. To test if overrepresentation of *pilC* mutants inside the aggregate is statistically significant, we performed a z-test. The null hypothesis is that the *pilC* cells are randomly distributed, and therefore the mean of $P(t_{final})$ is 0. The *P* value for accepting the null hypothesis is 0.002, indicating that the *pilC* mutant is partially rescued by WT cells.
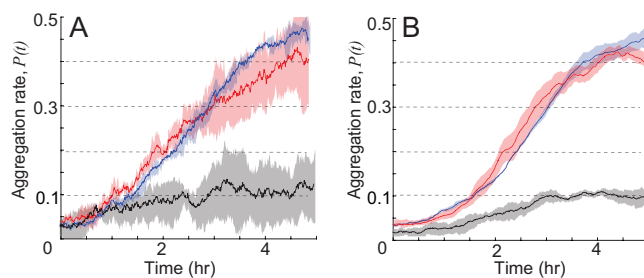

FIG 3 Comparison of aggregation rates (equation 1) between experiment (A) and simulation (B). Solid line is the average value, and shaded area is a standard deviation for each time point. Red, blue, and black colors correspond to WT, *csgA* mixed with WT, and *pilC* mixed with WT, respectively.
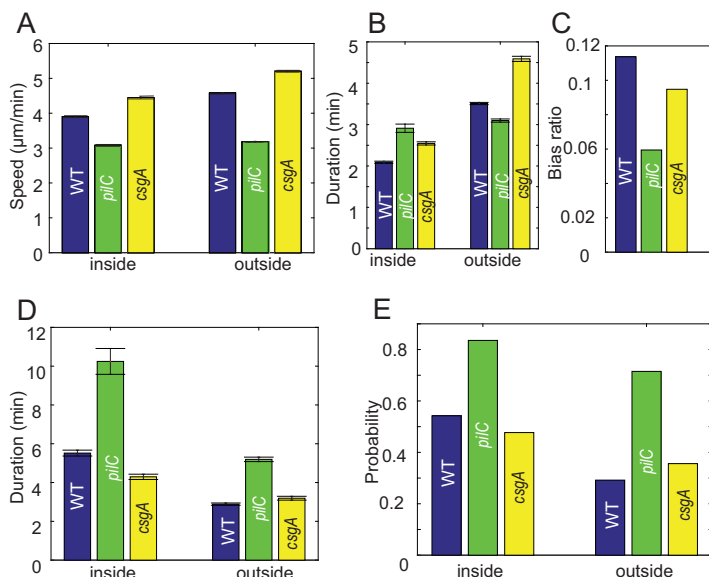
FIG 4 Experimental results of *pilC* and *csgA* cell behavior when mixed with WT cells compared with data from WT cells alone. Error bars represent bootstrapped 95% confidence interval of the means. (A) Persistent state speed of cells inside and outside aggregates. (B) Persistent state duration of cells inside and outside aggregates. (C) Bias ratio as defined in equation 2 shows the tendency of cells to extend their runs when approaching the aggregates. (D) Nonpersistent duration of cells inside and outside aggregates. (E) Probability of transitioning into a nonpersistent state after a persistent run for cells inside and outside aggregates.

Motility behaviors of rescued *pilC* and *csgA* cells differ from WT cells. To quantify single-cell behaviors, the cell trajectories were discretized into segments using the same method as in reference 14. The resulting segmented trajectories were then quantified as either persistent or nonpersistent run vectors. Persistent runs are interpreted as cells moving along their major axis using one or both motility systems whereas nonpersistent runs correspond to "stops" (or pauses) in progressive movements, during which cells can perhaps be pushed around by other cells. A run vector begins at a change of state (persistent to either nonpersistent or reversal) and ends at the next change of state. The properties of the resulting run vectors, such as duration (time between state changes) and speed (Euclidean distance over time), were used to quantify single-cell behavior during aggregation. The run vectors were also labeled with the distance to the nearest aggregate boundary and moving direction relative to the nearest aggregate center. Previous work has shown that WT cells have longer run durations when running toward an aggregate (bias effect), and cells decrease their motility inside aggregates ("traffic jam" effect) (14, 34). These effects have been shown to be important for aggregation (14, 15, 30). To quantify traffic jam and bias effects, we focus on the relationship between run vector properties and their distance and direction relative to aggregates.

To study the relationship between the run vector properties and the distance to aggregates, we divided the run vectors into 2 groups: those inside aggregates and those outside. Then we calculated the mean duration and speed for the persistent and nonpersistent state in each group (Fig. 4). We find that both WT and mutant cells mixed with WT cells display a traffic jam effect since they all have shorter persistent run durations and longer nonpersistent run durations inside aggregates (Fig. 4B and D). To quantify the bias in run duration, we divided the run vectors into 2 groups: those running toward aggregates and those running away. Then we define the bias ratio by

$$B = \frac{d_{to} - d_{away}}{d_{all}},\qquad(2)$$

where d_{to} is the average run duration of cells going toward aggregates, d_{away} is the

average run duration of cells going away from aggregates, and d_{all} is the average run duration of all cells. Figure 4C shows that each mutant mixed with WT cells has a bias ratio greater than 0, though both are less than WT.

To compare the traffic jam effect of *pilC* cells mixed with WT cells, we compared the speed and state durations of *pilC* and WT cells. In general, *pilC* cells exhibit longer stop durations, more frequent stops, and slower speeds, suggesting that loss of S-motility has compromised their overall mobility. Unlike WT cells, *pilC* cells show less than a 5% speed reduction inside aggregates during the persistent state (Fig. 4A) and show only 7% shorter persistent run durations (Fig. 4B). Furthermore, *pilC* cells show less bias in their run duration (Fig. 4C). Compared with WT and *csgA*, smaller differences between cell behaviors inside and outside aggregates may reduce the traffic jam effect, thereby impeding aggregation of *pilC* cells. On the other hand, *pilC* cells show a longer nonpersistent duration and a higher probability of transitioning to the nonpersistent state. However, the difference in the transitioning probability between inside and outside aggregates is smaller (Fig. 4D and E).

Similarly, we compared the traffic jam and bias effects of a *csgA*-WT mixture with WT cells. While *csgA* speed is ~20% faster than WT cells inside aggregates, *csgA* cells show proportional speed reduction inside aggregates (Fig. 4A). Similar to WT cells, *csgA* cells also have shorter persistent durations inside aggregates (Fig. 4B), and longer nonpersistent durations (Fig. 4D). Moreover, *csgA* cells increase their probability of transitioning to the nonpersistent state when inside the aggregates. However, the difference of this probability between inside and outside the aggregates of *csgA* cells is smaller than that of WT cells (Fig. 4E). All of the above behaviors reduce the motility of *csgA* cells inside the aggregates, likely creating a WT-like traffic jam effect.

In comparison with *pilC* cells, *csgA* cells likely have a stronger traffic jam effect due to a more pronounced reduction in speed (Fig. 4A) and persistent run duration (Fig. 4B) inside the aggregates. On the other hand, their traffic jam effect is expected to be weaker than WT due to reduced differences in nonpersistent duration (Fig. 4D) and probability between inside and outside (Fig. 4E). The *csgA* cells also have a weaker bias than WT cells (Fig. 4C). It remains to be seen why, despite a somewhat weaker bias and traffic jam effect, about the same proportion of *csgA* cells accumulate in the aggregate as WT cells (Fig. 3).

Data-driven models can match the aggregation dynamics of *pilC* and *csgA* cells based on the quantified motility parameters and their correlations. To more stringently test the effect of cell behaviors on aggregation, we extended the data-driven model approach used in our previous work (14) to model experiments with mixtures of two strains. To this end, we introduce a population of two agents corresponding to WT and mutant (either *pilC* or *csgA*) cells. Agent behaviors are chosen from the experimental data using K-nearest neighbor (KNN) sampling based on simulation time and the agents' distance and moving direction relative to the nearest aggregate. Given that the overwhelming majority of cells in the experiments are WT, we use only WT agent density to detect aggregates. This way, WT agents affect the behavior of mutant agents but not vice versa. At each time step, the WT density profile is estimated from the WT agent positions by kernel density estimation (KDE) (35), and the aggregates are then detected from the density profile. Thereafter, we pick agent behaviors and move agents accordingly. Each simulation was run for 5 h, after which we calculated the aggregation rate $P(t)$ as we did for the experiment. Simulations containing *csgA* agents mixed with WT agents display an aggregation rate similar to that of WT agents, whereas simulations with *pilC* agents exhibit much weaker aggregation (Fig. 3B). Comparing the results of these simulations to the experimental measurements (Fig. 3A), we concluded that the model can reproduce the aggregation dynamics for WT and each mutant cell mixture with WT. In other words, dependencies (correlations) included in the sampling of agent behavior contain sufficient information to recapture observed aggregation dynamics.

As a control for the previous simulations, we performed simulations where we removed all dependencies such that agent behavior was randomly chosen from the
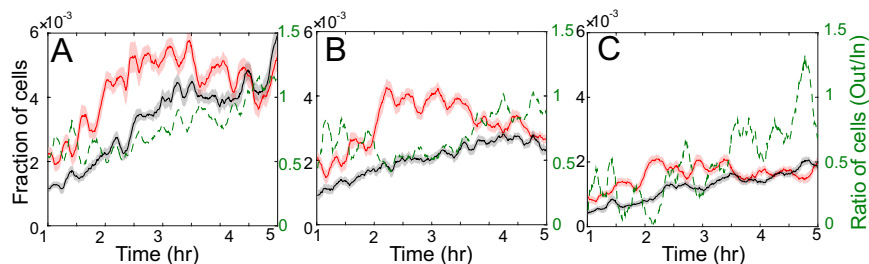
FIG 5 Analysis of agents moving in and out of aggregates at each time step for *csgA* (A), WT (B), and *pilC* (C). Red lines are the 10-min moving average of the fraction of agents moving into all aggregates at a given time. The fraction is calculated as a ratio of agents moving into the aggregate and the total number of agents simulated. Similarly, black lines are the 10-min moving average of the fraction of agents moving out of the aggregates. The shaded areas are 95% confidence interval of the moving mean. The right vertical axis and dashed green line show the ratio of number of agents moving out of aggregates divided by the number of agents moving in the aggregates.

whole data set. As expected, we did not see any aggregation for mutant mixtures or WT agents (Fig. S2A to C). This result shows that some combination of cell behavior dependence on time, distance, and direction to nearest aggregate is essential for aggregation. Since there are many cell behavior dependencies in this model, our next step is to find which dependencies are more important for aggregation.

Previous work on WT aggregation has shown that cell behaviors are different at different times during development, and this time dependence of cell behaviors affects aggregation dynamics (14). To determine whether time dependence is important for mutant cell aggregation, we performed simulations where agent behavior does not depend on time. Removing time dependence for WT aggregation causes $P(t_{final})$ to drop from ~0.45 to ~0.35 (Fig. S2F), which confirms our previous result (14) that time dependence helps WT aggregation. However, removing time dependence for mutant agents (while keeping it for WT agents) does not affect aggregation dynamics for either *pilC* (Fig. S2D) or *csgA* (Fig. S2E). This shows that behavior dependence on time is not important for mutant cell aggregation.

To further illuminate differences between WT, *cgsA*, and *pilC* strains, we used our simulations to quantify the fraction of cells that enter and exit aggregates as a function of time (Fig. 5). The results indicate notable differences. Comparing WT and *cgsA*, we can see that, although more *cgsA* cells reach the aggregates (red line in Fig. 5A versus Fig. 5B), a larger fraction of these cells leave (green line in Fig. 5A versus Fig. 5B). Therefore, we can hypothesize that reduced traffic jamming of *cgsA* cells is compensated by increase in motility. On the other hand, for *pilC* cells, the motility defects make them less likely to reach aggregates and slightly less likely to stay, leading to only weak aggregation. In what follows, we aim to test these hypotheses and relate these observations with the trends in cell behaviors quantified in Fig. 4.

Transitioning to and staying in the nonpersistent state in aggregates does not help *pilC* and *csgA* aggregation. Given that the increase of nonpersistent state duration increases the time that cells spend inside aggregates, we hypothesized that this effect is an essential component of the traffic jam effect and aids the aggregation of mutant cells. To test this hypothesis, we performed simulations where the nonpersistent state duration for agents is not conditional on their position relative to the aggregate. Surprisingly, removing this dependence does not have an obvious effect on *pilC* (Fig. 6A) or *csgA* (Fig. 6B) and leads to only a modest decrease in WT aggregation [~0.05 or ~10% drop in $P(t_{final})$; Fig. 6C]. This result shows that longer "stops" inside aggregates are not the main reason for successful aggregation.

To assess the effects of a higher probability of "stops" (i.e., nonpersistent runs) inside the aggregates, we performed simulations where the probability of transitioning to a nonpersistent state is independent of the agents' position (i.e., sampled from the same distribution inside and outside an aggregate). We discovered that removing this dependence does not affect aggregation for *pilC* (Fig. 6D) or *csgA* (Fig. 6E) and leads to
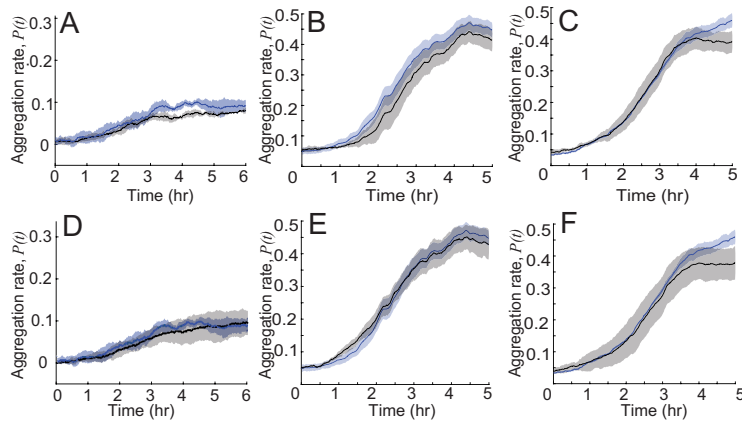
FIG 6 Comparison of the aggregation rates [quantified as $P(t)$, equation 1] (on y axis) from simulations of *pilC* (A and D) and *csgA* (B and E) cells mixed with WT cells and WT (C and F) cells alone. Blue line and shaded area are the simulation result under conditions where agent behavior is chosen from experimental data. Black lines represent simulations where nonpersistent behavior does not depend on distance to aggregates (A to C) or in which probability to nonpersistent state does not depend on distance to aggregates (D to F). Shaded areas show standard deviations.

only an ~0.07 (~15%) drop in $P(t_{final})$ for WT (Fig. 6F). It appears that longer nonpersistent state durations and a higher probability of transitioning to the nonpersistent state are not the main reasons for cell accumulation in aggregates. In summary, the difference in stopping probability and duration between inside and outside the aggregates is not critical for the traffic jam effect or can be compensated by other mechanisms.

Behaviors in the persistent state are critical for the aggregation. To test which persistent state behaviors are important for aggregation, we first removed the bias toward aggregates, which is the dependence of run duration on the angle between the moving cell and the closest aggregate. This leads to an ~0.03 drop in $P(t_{final})$ for *pilC* (Fig. 7A). For *csgA* (Fig. 7B) and WT (Fig. 7C), $P(t_{final})$ drops to the 0.15 to 0.2 range. This result shows that bias in run duration is essential, more so for *csgA* and WT aggregation than *pilC* aggregation. This also agrees with Fig. 4C where we showed that *csgA* and WT cells have larger bias ratios than *pilC* cells. However, given the overall poor aggregation of *pilC*, the decrease associated with lack of bias is still important and in relative terms
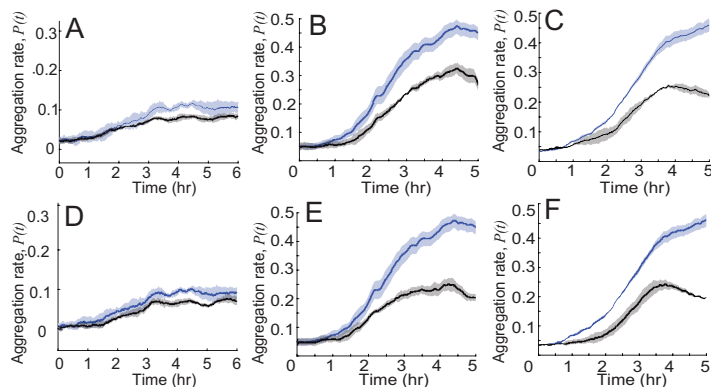

FIG 7 Comparison of the aggregation rates [quantified as $P(t)$, equation 1] (on y axis) of simulation of *pilC* (A and D) and *csgA* (B and E) cells mixed with WT cells and WT (C and F) cells alone. The blue lines and shaded area are simulation results under normal conditions. The black lines represent simulations where persistent behavior does not depend on run direction, i.e., without bias (A to C), or in which run duration does not depend on the distance to aggregates, i.e., without traffic jam (D to F). Shaded areas show standard deviations.
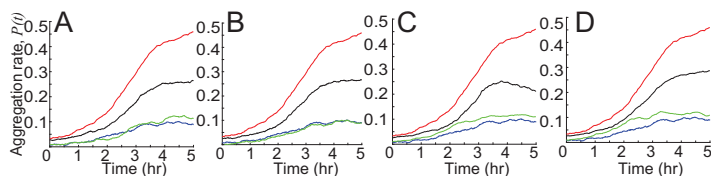
FIG 8 Simulations swapping WT data and *pilC* cell data demonstrate which mutant cell behaviors are sufficiently different from wild-type to affect the aggregation rate [quantified as *P*(*t*), equation 1] (on the *y* axis). Blue lines are simulation results of agents using *pilC* cell data. Red lines are simulation results of agents using WT cell data. Green lines are simulations of agents using *pilC* data with partial WT cell data. Black lines are simulations of agents using WT data with partial *pilC* cell data. (A) Green is agents using WT cell probability to nonpersistent state and other *pilC* data. Black is agents using *pilC* probability to nonpersistent state and other WT data. (B) Green show agents using WT cell nonpersistent state duration and other *pilC* data. Black is agents using *pilC* cell nonpersistent state duration and other WT data. (C) Green is agents using WT cell persistent state duration and other *pilC* data. Black is agents using *pilC* cell persistent state duration and other WT data. (D) Green is agents using WT cell persistent state speed and other *pilC* data. Black is agents using *pilC* cell persistent state speed and other WT data. Only mean values are plotted for clarity.

is just slightly weaker than that of the other strains [30% reduction of final $P(t_{final})$ for *pilC* versus 40% for *csgA* and 45% for WT].

Next, we attempt to make the cells behave the same way inside and outside the aggregate to remove the traffic jam effect but maintain the bias. First, we removed persistent state speed and duration dependence on the agents' distance to the nearest aggregate while keeping the dependence of run duration on the angle between the moving cell and the closest aggregate. The results show that removing the distance dependence decreases aggregation for all types of cells: $P(t_{final})$ drops ~0.03 for *pilC* (Fig. 7D) and drops ~0.25 for *csgA* (Fig. 7E) and WT (Fig. 7F) cells. Therefore, the reduction of speed and duration inside aggregates is important for aggregation. Interestingly, the reduction of speed and duration can also be considered a traffic jam effect. Comparing the traffic jam effects in the nonpersistent state, i.e., longer duration and higher probability of nonpersistent state inside aggregates, traffic jam effects in the persistent state appear to be more important. Notably, removing persistent speed and duration dependence on distance decreases aggregation more in *csgA* and WT cells than in *pilC* cells. Even considering the poor aggregation of *pilC*, the relative decrease in aggregation is still weaker for *pilC* [30% reduction of final $P(t_{final})$ for *pilC* versus 55% for *csgA* and 55% for WT]. This shows that *csgA* and WT cells have a stronger traffic jam effect than *pilC*, in agreement with Fig. 4A and B.

Different motility behaviors of *pilC* and *csgA* cells explain the partial rescue of *pilC* and full rescue of *csgA*. The results thus far match the observed behaviors of mutant cells with their observed aggregation dynamics. Next, we try to determine which mutant cell behaviors are responsible for the different aggregation rates compared with WT. To this end, we introduce a new "hybrid" simulation technique in which certain aspects of mutant and WT agent behaviors are swapped with one another or scaled to match the mean of another. For example, the experimental data show that *pilC* mutants switch to the nonpersistent state more frequently and stay in the nonpersistent state longer (Fig. 4D and E). To determine whether these behaviors contribute to weaker aggregation, we performed simulations where we swap some of the *pilC* motility behaviors with WT behaviors (Fig. 8). When agents use the *pilC* probability of transitioning to the nonpersistent state and WT data for other behaviors, aggregation drops [$P(t_{final})$ drops ~0.2] (Fig. 8A). On the other hand, agents' use of WT probability of transitioning to the nonpersistent state with *pilC* data for other behaviors does not improve *pilC* aggregation (Fig. 8B). To further confirm that the decrease in aggregation is due to longer stops or a higher stopping frequency rather than some other feature of the *pilC* data, we performed simulations of WT cells where we increased only the nonpersistent duration or nonpersistent probability to match the average data of *pilC* cells (Fig. S3). The aggregation rate is decreased compared to WT aggregation, suggesting that frequent stops are one of the major impediments to *pilC* aggregation.
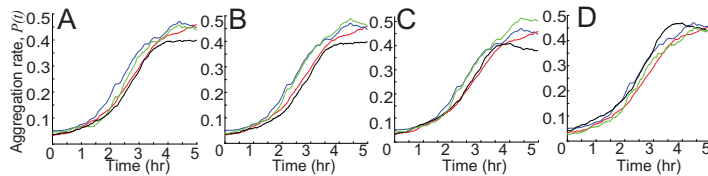
FIG 9 Simulations swapping WT data and *csgA* cell data demonstrate which mutant cell behaviors are sufficiently different from wild type to affect the aggregation rate [quantified as $P(t)$, equation 1] (on y axis). Blue lines are simulation results of agents using *csgA* cell data. Red lines are simulation results of agents using WT cell data. Green lines are simulations of agents using *csgA* data with partial WT cell data. Black lines are simulations of agents using WT data with partial *csgA* cell data. (A) Green is agents using WT cell probability to nonpersistent state and other *csgA* data. Black is agents using *csgA* probability to nonpersistent state and other WT data. (B) Green is agents using WT cell nonpersistent state duration and other *csgA* data. Black is agents using *csgA* cell nonpersistent state duration and other WT data. (C) Green is agents using WT cell persistent state duration and other *csgA* data. Black is agents using *csgA* cell persistent state duration and other WT data. (D) Green is agents using WT cell persistent state speed and other *csgA* data. Black is agents using *csgA* cell persistent state speed and other WT data. Only mean values are plotted for clarity.

To learn how *pilC* persistent behaviors affect aggregation, we performed simulations where agents use the WT persistent duration data combined with other *pilC* cell data and vice versa (Fig. 8C). Agents using *pilC* persistent duration combined with other WT data have reduced aggregation compared with WT [$P(t_{final})$ drops ~0.25]. Agents using WT persistent duration combined with other *pilC* data show improved aggregation over *pilC* [$P(t_{final})$ increases ~0.02]. This is not surprising since WT cells have a much stronger persistent duration bias, and stronger bias leads to more complete aggregation. Finally, agents using *pilC* persistent speed combined with other WT data have reduced aggregation compared with WT [$P(t_{final})$ drops ~0.2] whereas WT persistent speed combined with other *pilC* data improves *pilC* aggregation [$P(t_{final})$ increases ~0.02] (Fig. 8D). This is because *pilC* cells have similar speeds inside and outside aggregates whereas WT cells have slower speeds inside aggregates and this slowdown improves aggregation. Overall, our results show that weak aggregation of *pilC* is due to slow speed, longer nonpersistent durations, and a higher probability of transitioning to the nonpersistent state.

For *csgA* mutants, Fig. 4E shows that the difference for stopping probabilities and durations inside and outside aggregates is less pronounced than WT. To test whether these behaviors decrease aggregation, we performed a simulation where agents use WT data for probability of transitioning into the nonpersistent state and *csgA* data for other behaviors. This simulation does not improve *csgA* aggregation (Fig. 9A). But agents using *csgA* probability to transition to the nonpersistent state and WT data for other behaviors cause $P(t_{final})$ to drop ~0.05 compared with WT aggregation. Moreover, in Fig. 9B, agents using *csgA* nonpersistent duration and WT data for other behaviors show a slight decrease in aggregation compared with WT aggregation [$P(t_{final})$ drops ~0.05]. On the other hand, agents using WT nonpersistent duration and *csgA* data for other behaviors show a slight increase in aggregation compared with *csgA* aggregation [$P(t_{final})$ increases ~0.02]. These results show that the differences in nonpersistent state switching and duration between WT and *csgA* do not affect aggregation much.

To learn how *csgA* persistent behaviors affect aggregation, we performed a simulation where agents use persistent duration of WT cells and other behaviors of *csgA* cells (Fig. 9C). This leads to a slightly better aggregation compared with *csgA* cells [$P(t_{final})$ increases ~0.05]. Agents using *csgA* persistent duration and other WT cell behavior show a slightly lower aggregation compared with WT cells [$P(t_{final})$ drops ~0.07]. Note that WT persistent duration has a bigger bias but shorter duration. To learn whether a shorter duration will decrease *csgA* aggregation, we performed a simulation where agents use *csgA* data but scale the persistent duration to match the average duration of WT cells. This led to a lower aggregation (Fig. S4). These results show that the *csgA* weaker bias is partially compensated by the longer persistent duration.

Finally, to find whether the faster speed of *csgA* cells in the persistent state helps aggregation, we performed simulations where agents use *csgA* persistent speed and other WT behaviors (Fig. 9D). This leads to faster aggregation, but $P(t_{final})$ remains the same compared with WT. Moreover, agents using WT persistent speed and other *csgA* behaviors have a slightly lower aggregation rate compared with *csgA*. This result shows that *csgA* cells' faster speed compensates for the weaker (compared with WT cells) bias in persistent duration and explains why *csgA* and WT cells show similar aggregation rates. Therefore, the rescue of collective behaviors can occur even without the complete rescue of the underlying single-cell behaviors.

DISCUSSION

In this work, we developed a novel methodology to assess which aspects of individual cell behavior are responsible for observed trends in collective self-organization. We show that this approach is both robust and versatile by examining how aggregation is restored when *csgA* and *pilC* mutants are mixed with wild-type (WT) cells. The *csgA* rescue is quantitatively complete based on the percentage of cells in the aggregate whereas *pilC* rescue is marginal. As with our previous analysis of wild-type aggregation dynamics, we conclude that three features of cell behavior contribute to efficient accumulation in aggregates. First, the cells follow aligned paths that precede the appearance of aggregates such that their orientations are correlated with one another and with the direction to the nearest aggregate likely to appear along their path. That essentially reduces the search for aggregates to one dimension (1D). Second, due to the bias in persistent run durations, cells move longer when approaching an aggregate than when moving in the opposite direction. This biased random walk results in an increase in cell flux toward the aggregates and accelerates the aggregation dynamics. The biased walk appears to be due to chemotaxis (15), and several lipids are known to be chemoattractants (24, 44). Finally, once in aggregates, the cells are less likely to leave. Notably, all three strains displayed these three features, but their contributions are roughly proportional to the extent of aggregation, with *pilC* being the worst.

Our approach enables a detailed examination of motility parameters that mediate a particular aggregation feature. For example, we extended the data-driven modeling approach for hybrid populations of agents that correspond to wild-type and mutant behaviors by using swapped-data set sampling (Table 1) to pinpoint cell movement features that are responsible for trapping cells in an aggregate. When cells enter an aggregate, their speed decreases, the time spent in a persistent state decreases, their probability to transition to a nonpersistent state increases, and the time spent in the nonpersistent state increases. By swapping each of these features between WT and mutant in agent-based models, we discovered that longer nonpersistent state duration and a higher probability of transitioning to the nonpersistent state are not the main reasons cells accumulate in aggregates. Rather, reduction in speed and run duration are the most critical features that trap cells in aggregates. It is possible that this result will inform those now searching for the enigmatic chemical signal that traps cells in an aggregate.

Our approach also revealed that *csgA* achieved complete rescue despite differing in most WT motility parameters. Most notably, *csgA* cells show longer persistent run distances resulting from faster speeds and longer durations in the persistent state and compensate for a reduced biased random walk to produce similar aggregation to WT cells. Genetic studies have shown that *csgA* cells possess both motility systems. The mutant fails to develop specifically because it fails to produce one or more essential developmental signals that mediate aggregation in addition to a signal that induces sporulation. While the mutant clearly responds to the wild-type signal(s), it would appear that the *csgA* mutation causes a slight downstream effect in perception or motility regulation that reduces the bias. The chemical nature of the aggregation signal(s) remains unknown, but it seems to be a development-specific lipid-like molecule (5). The signal is unlikely to be the exopolysaccharide (EPS) used in S-motility as

TABLE 1 Summary of performed simulation and effect of mutant behaviors on aggregation

Observation	Simulation performed	Effect on aggregation
pilC mutants switch to nonpersistent state more frequently.	*pilC* agents use WT data for the probability of transitioning to the nonpersistent state and vice versa.	Frequent stops slow down aggregation (Fig. 8A), but the final aggregation result is similar after a longer simulation time (Fig. S3).
pilC mutants stay in nonpersistent state longer.	*pilC* agents use WT data for the duration and speed of the nonpersistent state and vice versa.	Longer stops slow down aggregation (Fig. 8B), but the final aggregation result is similar after a longer simulation time (Fig. S3).
Run duration for *pilC* mutants shows lesser dependence on cell density and smaller bias.	*pilC* agents use WT data for persistent duration and vice versa.	Smaller bias and difference between inside and outside aggregates for run durations impede aggregation (Fig. 8C).
pilC mutants do not show speed reduction inside the aggregates.	*pilC* agents use persistent speed scaled to match WT data and vice versa.	Lack of speed reduction inside the aggregates impedes aggregation (Fig. 8D).
Stopping probability inside and outside aggregates is less pronounced for *csgA* mutants.	*csgA* agents use WT data for the probability of transitioning to the nonpersistent state and vice versa.	The density dependence of stopping probability does not have a major effect on aggregation (Fig. 9A).
Stop duration inside and outside aggregates is less pronounced for *csgA* mutants.	*csgA* agents use WT data for the duration and speed of the nonpersistent state and vice versa.	Density dependence of stopping slightly impedes aggregation (Fig. 9B).
csgA mutants have a weaker bias but longer duration in the persistent state than WT.	*csgA* agents use WT data for the duration of the persistent state. Scaled the persistent duration of *csgA* agents to match mean values of WT.	While longer persistent duration helps *csgA* aggregation (Fig. S4), the weaker bias has an opposite and stronger effect. Overall, compared with other behaviors, *csgA* persistent duration impedes aggregation more (Fig. 9C).
csgA mutants have faster speed in persistent state.	*csgA* agents use WT data for the persistent speed and vice versa.	Faster speed speeds up aggregation (Fig. 9D).

csgA produces normal levels of EPS, shows normal agglutination, and possesses S-motility.

Finally, we attempted to learn why *pilC* rescue is so poor. *pilC* cells are significantly attenuated for all behaviors that are important for wild-type aggregation. The swapped data set sampling approach revealed that the failure of *pilC* cannot be attributed to one or two specific changes. *pilC* cells lack pili and EPS. S-motile cells use the pilus to attach to EPS on adjacent cells, and retraction of a motor at the base of the pilus pulls the cell forward. *pilC* cells have significantly decreased bias which our results suggest is due to reduced speed, reduced run durations, and increased frequency of transiting to the nonpersistent state. As the WT cells would be expected to provide normal levels of EPS and other required signals, *pilC* cells clearly lack the appropriate response. This is a striking finding in view of the observation that neither pili nor S-motility is required for aggregation. Some S mutants, like *pilA* (which encodes the pilus structural protein) and *pilT* (which encodes the pilus retraction motor), can aggregate using only the A-motility system (36, 37). The results point to a downstream defect, perhaps related to the perception of lipid chemoattractants, which has been noted in certain S mutants but not examined specifically in *pilC*. The *pilC* mutant also has a diminished traffic jam effect. *pilC* cells frequently leave aggregates and ultimately are about 4-fold less abundant than *csgA* or WT cells. Again, the as-yet-unknown signal(s) used to hold cells in aggregates should be in sufficient concentration, leading one to suspect that the problem is more specifically due to *pilC* response to the signal.

There is also a striking difference in the rescue of the two mutants for sporulation by WT cells. As a benchmark, WT cells form 0.14 spores per input cell with the remaining cells undergoing alternate developmental fates such as programmed cell death and formation of peripheral rods. WT cells efficiently rescue the sporulation of *csgA* mutants. *csgA* cells alone form 2×10^{-5} spores per input cell, which increases to 0.15 spores per input cell in the presence of WT cells. In contrast, WT cells do not rescue *pilC* sporulation. *pilC* cells alone form 3×10^{-3} spores per input cell, which increases minimally to 4×10^{-3} spores per input cell in the presence of WT cells. Sporulation is thought to have little bearing on aggregation since it occurs after aggregation is complete. Nevertheless, these results clearly support the ideas developed in this work for aggregation that *csgA* cells are proficient in responding to signals, unlike *pilC*.

Data-driven modeling approaches can identify important correlations that drive self-organization as a first step to understand the mechanisms. For example, in our previous work we have combined data-driven and mechanistic agent-based modeling to show that steric alignment and slime-trail following are sufficient to explain how cells align during aggregation (15). Moreover, we also showed that chemotaxis rather than contact-based signaling is responsible for bias in reversal times for cells moving toward versus away from the aggregates (15). However, the exact molecular pathway(s) responsible for chemotaxis has yet to be determined. Furthermore, molecular mechanisms for the interactions that decrease the residence time of cells inside the aggregates (traffic jam effects) are not clear. Another interesting observation is that both mutant strains considered here displayed differences from wild-type behavior in all the behaviors driving aggregation. That fact can indicate complex or possibly pleiotropic interactions between the pathways controlling cell behavior. Perhaps application of the developed methodology to mutant strains with more subtle phenotypes can shed light into these challenges in future work.

Multicellular self-organization is prevalent in biological systems but has proven challenging to study. There are complex feedback and compensatory mechanisms at the population level as well as pleiotropic effects of single mutations. Given significant heterogeneity of individual cell behaviors, small trends in behaviors between mutant strains could dissipate over time or in contrast could accumulate, leading to differences in the emergent patterns. Our results demonstrate how careful quantification of cell behavior coupled to data-driven modeling approaches can predict these effects and pinpoint important synergies and compensatory mechanisms.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. All *M. xanthus* strains were grown in CYE broth (1% Bacto Casitone [Difco], 0.5% yeast extract [Difco], 10 mM 4-morpholinepropanesulfonic acid [MOPS] [pH 7.6], and 0.1% MgSO$_4$), and development was induced on thin (10 ml in 100-mm petri dish) TPM agar [10 mM Tris-HCl (pH 7.6), 1 mM KH(H$_2$)PO$_4$ (pH 7.6), 10 mM MgSO$_4$, 1.5% agar (Difco)] plates containing 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) and 100 μM vanillate as described in reference 14. Strain LS3910 was constructed by electroporation (38) of pLJS145 (14) into LS2442 (39) Transformants were selected using CYE 1.5% agar plates containing 15 μg ml^{-1} oxytetracycline. *pilC* mutant LS3011 was constructed by Magellan mutagenesis of DK1622 as described in reference 40. Strain LS4223 was constructed by electroporating the tdTomato plasmid pLJS145 into LS3011 with selection on CYE agar containing 15 μg ml^{-1} oxytetracycline.

Fluorescence time-lapse microscopy. Time-lapse image capture was performed as described in reference 14. As in reference 14, the beginning of aggregation varied between replicates by up to 1 h. To avoid possible bias in movie alignment caused by differences in the aggregation rate of WT and mutant cells, the approach of using the fraction of tdTomato cells within the aggregates used in reference 14 was replaced with a technique that relied on YFP fluorescence. To quantify aggregation progress using YFP fluorescence, the 2-D Fourier transform coefficient magnitudes for wavelengths between 50 and 100 μm were summed for each frame. Aggregation start was then detected as the point at which the summed magnitude in the movie frames crossed 20% of the maximum value reached in that movie. Movies were then cropped to align the detected beginning of aggregation and equalize their lengths as described in reference 14.

Developmental assays. The developmental assays were performed by mixing the tdTomato fluorescent strain LS4223 or LS3909 with the YFP fluorescent wild-type strain LS3630 in a 1:10,000 ratio. tdTomato fluorescence indicated positions of the individual cells for strains LS4223 and LS3909, while the YFP fluorescence revealed the territories of the LS3630 cell aggregates. Specifically, 100 liters of the tdTomato-expressing strains and 1 ml of the YFP or nonfluorescent strains in the exponential phase were collected by centrifugation at 17,000 $\times$ g for 1 to 2 min and washed with 100 liters of ddH$_2$O, respectively. The tdTomato strains were further diluted to 5 $\times$ 10^6 cells ml^{-1}, while the YFP or nonfluorescent strains were concentrated to 5 $\times$ 10^8 cells ml^{-1} in ddH$_2$O. The diluted tdTomato strains were then mixed with the YFP or nonfluorescent strains in a 1:100 ratio, resulting in a final ratio of 1:10,000 between the tdTomato and the YFP or nonfluorescent cells. Thirty-five liters of the cell mixtures in 4 to 6 replicates was spotted onto a TPM plate and dried out in a 32°C incubator for 30 to 45 min. The plate was sealed with Parafilm and incubated in a 28°C dark room. With strains or mixtures that developed, development usually started between 7 and 10 h postincubation and produced stable aggregates in another 5 to 8 h. For time-lapse movies, images were captured at 30-s intervals beginning about 1 h before the initiation of aggregation and lasting until the formation of stable aggregates.

Viable spore data were obtained as noted previously (41).

Cell tracking, cell state detection, run vector extraction, and aggregate tracking. Cell tracking, cell state detection, and run vector extraction were performed as described in reference (14). Time-lapse images were band-pass filtered to filter out the background fluorescent signal. Thereafter, the MATLAB

function 'regionprops' was used to identify the centroid and orientation of every fluorescently labeled cell. Image-to-image linking of detected cell positions into trajectories is achieved using the method introduced in reference 42. To detect movement characteristics of the cell, an extended Kalman filter (EKF) was developed to estimate the most likely movement state of the cell at a certain time step. Here, we assume cells will be in one of three different states: persistent forward, persistent backward, and nonpersistent state. The EKF uses the anticipated cell position and detected cell position to calculate the probability of every state. The state with the maximum likelihood was then assigned as the movement state between the two images. Cell trajectories were then divided into run vectors, which start at the start of one contiguous movement state and end with the following change of state. The average speed, duration, distance, angle to the closest aggregate centroid, and distance to the closest aggregate boundary were calculated for every run vector. The detection of aggregate is based on the light intensity of pixels. The threshold of the aggregate light intensity is calculated using K-means clustering on the pixels in the final frames of experiments. Areas with light intensity higher than the threshold are considered aggregates. For simulations in Fig. S6 in the supplemental material, this threshold is increased by 20%.

Data-driven agent-based model. The agent-based model used here is adapted from our previous work (14). Given that simulations with the experimental mutant-to-WT ratio will lead to an unfeasible number of agents to simulate, we instead chose to implement the wide excess of WT cells via asymmetry in their interactions. We sample behaviors of both WT and mutant cells conditional only on the WT population distributions (see below). Each simulation consists of 10,000 WT agents and 8,000 mutant agents on a rectangular domain of 986 μm × 740 μm, equal to the microscope field of view, with periodic boundary conditions along each side. Each agent represents a single cell sampled from a biofilm of the same average density as in experiments (1.1 cells/μm^2), similar to sampling cell behaviors in the biofilm using a small number of fluorescently labeled cells. Similar to our previous model (14), each agent's behaviors such as run speed, run duration, and run angle are drawn from the experiment data based on the time since the beginning of the experiment, the angle between the cell orientation and the average bearing angle of neighboring runs, and distance and angle to nearest aggregate. Note that unlike our previous model (14), here we did not use local cell density extracted from the time-lapse microscopy to choose our agent behaviors since the light intensity in the experiment varies too much for reliable density estimates outside the aggregates. We use the same method as in reference 14 to select run behavior for agents. Varying the number of mutant agents simulated does not significantly alter the results (e.g., Fig. S5).

Since in the experiments the ratio of WT to mutant cells is over 10,000:1, it is fair to assume that WT cell behavior is not affected by the mutant cells. Therefore, in the simulation, we usually chose the agent behavior based solely on the population distribution of WT agents. Moreover, the density estimation in simulation uses only WT agents so that mutant agents will not affect WT agent behavior. However, in simulations where we swap some WT data with mutant data, e.g., in Fig. 8 and 9, we do this by replacing some mutant data with WT data or vice versa and feed the combined data to mutant agents. WT agents will always use WT data to provide background information such as neighbor cell alignment and density profile, etc.

For simulations where agents use both mutant data and WT data (Fig. 8 and 9), the simulation process is slightly different from the original (14). In particular, simulations where agents use WT data for nonpersistent probability or nonpersistent state behavior and mutant data for other behaviors, agents will choose their behaviors from WT data or mutant data accordingly using nearest-neighbor methods. A similar procedure is applied for simulations where agents use mutant data for nonpersistent probability or nonpersistent state behavior and WT data for other behaviors. For simulations where agents use WT data for persistent state speed or duration and mutant data for other behaviors, agents will choose their behaviors from mutant data only and then scale the persistent state speed or duration to match the mean of the WT data. This way, we can keep the correlation between the speed and duration in the mutant data. Moreover, to keep the bias in WT data, we split WT data into 2 branches: data of cells moving toward the aggregate and data of cells moving away from the aggregates. To keep the traffic jam effect in WT data, we further split the 2 data branches into smaller branches based on the distance to aggregate: each branch now contains data of cells with distance to aggregate within a 1-μm window and moving in the same direction. Then we calculate the mean speed or duration of each branch of data. We perform a similar calculation for mutant data and use the means to scale the agent behavior to match the WT data using the following equation:

$$B(dir, dis) = \frac{B_1(dir, dis)}{B_{mu}(dir, dis)} B_{WT}(dir, dis) \tag{3}$$

where B is the final scaled behavior (speed or duration) for the agent, dir is the moving direction (moving toward or away from aggregate) of the agent and dis is the distance to aggregate, B_1 is the selected behavior from mutant data, B_{mu} is the mean of the mutant data calculated as above, and B_{WT} is the mean of WT data calculated as above. For simulations where agents use mutant data for persistent state speed or duration and WT data for other behaviors, we apply a similar procedure, and the equation to scale the agent behavior becomes:

$$B(dir, dis) = \frac{B_1(dir, dis)}{B_{WT}(dir, dis)} B_{mu}(dir, dis) \tag{4}$$

where B_1 here is the selected behavior from WT data.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, PDF file, 0.5 MB.

FIG S2, PDF file, 0.3 MB.

FIG S3, PDF file, 0.7 MB.

FIG S4, PDF file, 0.5 MB.

FIG S5, PDF file, 0.5 MB.

FIG S6, PDF file, 0.9 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Shahbazi MN, Siggia ED, Zernicka-Goetz M. 2019. Self-organization of stem cells into embryos: a window on early mammalian development. Science 364:948–951. https://doi.org/10.1126/science.aax0164.

2. Hodgkin J, Kaiser D. 1979. Genetics of gliding motility in *Myxococcus xanthus* (*Myxobacterales*): two gene systems control movement. Mol Gen Genet 171:177–191. https://doi.org/10.1007/BF00270004.

3. Kim SK, Kaiser D. 1990. C-factor: a cell-cell signaling protein required for fruiting body morphogenesis in *M. xanthus*. Cell 61:19–26. https://doi.org/10.1016/0092-8674(90)90211-V.

4. Sager B, Kaiser D. 1994. Intercellular C-signaling and the traveling waves of *Myxococcus*. Genes Dev 8:2793–2804. https://doi.org/10.1101/gad.8.23.2793.

5. Boynton TO, Shimkets LJ. 2015. *Myxococcus csgA*, Drosophila Sniffer, and human HSD10 are cardiolipin phospholipases. Genes Dev 29:1903–1914. https://doi.org/10.1101/gad.268482.115.

6. Pérez J, Jiménez-Zurdo JI, Martínez-Abarca F, Millán V, Shimkets LJ, Muñoz-Dorado J. 2014. Rhizobial galactoglucan determines the predatory pattern of *Myxococcus xanthus* and protects *Sinorhizobium meliloti* from predation. Environ Microbiol 16:2341–2350. https://doi.org/10.1111/1462-2920.12477.

7. Shimkets LJ, Gill RE, Kaiser D. 1983. Developmental cell interactions in *Myxococcus xanthus* and the *spoC* locus. Proc Natl Acad Sci U S A 80:1406–1410. https://doi.org/10.1073/pnas.80.5.1406.

8. Börner U, Deutsch A, Reichenbach H, Bär M. 2002. Rippling patterns in aggregates of myxobacteria arise from cell-cell collisions. Phys Rev Lett 89:078101. https://doi.org/10.1103/PhysRevLett.89.078101.

9. Igoshin OA, Neu J, Oster G. 2004. Developmental waves in myxobacteria: a distinctive pattern formation mechanism. Phys Rev E Stat Nonlin Soft Matter Phys 70:041911. https://doi.org/10.1103/PhysRevE.70.041911.

10. Zusman DR. 1984. Cell-cell interactions and development in *Myxococcus xanthus*. Q Rev Biol 59:119–138. https://doi.org/10.1086/413780.

11. Kaiser D, Welch R. 2004. Dynamics of fruiting body morphogenesis. J Bacteriol 186:919–927. https://doi.org/10.1128/jb.186.4.919-927.2004.

12. Kuner J, Kaiser D. 1982. Fruiting body morphogenesis in submerged cultures of *Myxococcus xanthus*. J Bacteriol 151:458–461. https://doi.org/10.1128/JB.151.1.458-461.1982.

13. O'Connor KA, Zusman DR. 1989. Patterns of cellular interactions during fruiting-body formation in *Myxococcus xanthus*. J Bacteriol 171:6013–6024. https://doi.org/10.1128/jb.171.11.6013-6024.1989.

14. Cotter CR, Schüttler HB, Igoshin OA, Shimkets LJ. 2017. Data-driven modeling reveals cell behaviors controlling self-organization during *Myxococcus xanthus* development. Proc Natl Acad Sci U S A 114:E4592–E4601. https://doi.org/10.1073/pnas.1620981114.

15. Zhang Z, Igoshin OA, Cotter CR, Shimkets LJ. 2018. Agent-based mod-

eling reveals possible mechanisms for observed aggregation cell behaviors. Biophys J 115:2499–2511. https://doi.org/10.1016/j.bpj.2018.11.005.

16. Chang YW, Rettberg LA, Treuner-Lange A, Iwasa J, Søgaard-Andersen L, Jensen GJ. 2016. Architecture of the type IVa pilus machine. Science 351:aad2001. https://doi.org/10.1126/science.aad2001.

17. Bischof LF, Friedrich C, Harms A, Søgaard-Andersen L, van der Does C. 2016. The type IV pilus assembly ATPase PilB of *Myxococcus xanthus* interacts with the inner membrane platform protein PilC and the nucleotide-binding protein PilM. J Biol Chem 291:6946–6957. https://doi.org/10.1074/jbc.M115.701284.

18. Wu SS, Kaiser D. 1997. Regulation of expression of the *pilA* gene in *Myxococcus xanthus*. J Bacteriol 179:7748–7758. https://doi.org/10.1128/JB.179.24.7748-7758.1997.

19. Nudleman E, Kaiser D. 2004. Pulling together with type IV pili. J Mol Microbiol Biotechnol 7:52–62. https://doi.org/10.1159/000077869.

20. Li Y, Lux R, Pelling AE, Gimzewski JK, Shi W. 2005. Analysis of type IV pilus and its associated motility in *Myxococcus xanthus* using an antibody reactive with native pilin and pili. Microbiology 151:353–360. https://doi.org/10.1099/mic.0.27614-0.

21. Mignot T, Nöllmann M. 2017. New insights into the function of a versatile class of membrane molecular motors from studies of *Myxococcus xanthus* surface (gliding) motility. Microb Cell 4:98–100. https://doi.org/10.15698/mic2017.03.563.

22. Guzzo M, Murray SM, Martineau E, Lhospice S, Baronian G, My L, Zhang Y, Espinosa L, Vincentelli R, Bratton BP, Shaevitz JW, Molle V, Howard M, Mignot T. 2018. A gated relaxation oscillator mediated by *FrzX* controls morphogenetic movements in *Myxococcus xanthus*. Nat Microbiol 3:948–959. https://doi.org/10.1038/s41564-018-0203-x.

23. Kearns DB, Campbell BD, Shimkets LJ. 2000. *Myxococcus xanthus* fibril appendages are essential for excitation by a phospholipid attractant. Proc Natl Acad Sci U S A 97:11505–11510. https://doi.org/10.1073/pnas.210448597.

24. Kearns DB, Venot A, Bonner PJ, Stevens B, Boons GJ, Shimkets LJ. 2001. Identification of a developmental chemoattractant in *Myxococcus xanthus* through metabolic engineering. Proc Natl Acad Sci U S A 98:13990–13994. https://doi.org/10.1073/pnas.251484598.

25. Arnold JW, Shimkets LJ. 1988. Cell surface properties correlated with cohesion in *Myxococcus xanthus*. J Bacteriol 170:5771–5777. https://doi.org/10.1128/jb.170.12.5771-5777.1988.

26. Behmlander RM, Dworkin M. 1991. Extracellular fibrils and contact-mediated cell interactions in *Myxococcus xanthus*. J Bacteriol 173:7810–7820. https://doi.org/10.1128/jb.173.24.7810-7820.1991.

27. Hagen DC, Bretscher AP, Kaiser D. 1978. Synergism between morpho-

genetic mutants of *Myxococcus xanthus*. Dev Biol 64:284–296. https://doi.org/10.1016/0012-1606(78)90079-9.

28. Konovalova A, Wegener-Feldbrügge S, Søgaard-Andersen L. 2012. Two intercellular signals required for fruiting body formation in *Myxococcus xanthus* act sequentially but non-hierarchically. Mol Microbiol 86:65–81. https://doi.org/10.1111/j.1365-2958.2012.08173.x.

29. Lee K, Shimkets LJ. 1996. Suppression of a signaling defect during *Myxococcus xanthus* development. J Bacteriol 178:977–984. https://doi.org/10.1128/jb.178.4.977-984.1996.

30. Sliusarenko O, Zusman DR, Oster G. 2007. Aggregation during fruiting body formation in *Myxococcus xanthus* is driven by reducing cell movement. J Bacteriol 189:611–619. https://doi.org/10.1128/JB.01206-06.

31. Zhang H, Angus S, Tran M, Xie C, Igoshin OA, Welch RD. 2011. Quantifying aggregation dynamics during *Myxococcus xanthus* development. J Bacteriol 193:5164–5170. https://doi.org/10.1128/JB.05188-11.

32. Jelsbak L, Søgaard-Andersen L. 1999. The cell surface-associated intercellular C-signal induces behavioral changes in individual *Myxococcus xanthus* cells during fruiting body morphogenesis. Proc Natl Acad Sci U S A 96:5031–5036. https://doi.org/10.1073/pnas.96.9.5031.

33. Peacock J. 1983. Two-dimensional goodness-of-fit testing in astronomy. Mon Not R Astron Soc 202:615–627. https://doi.org/10.1093/mnras/202.3.615.

34. Igoshin OA, Welch R, Kaiser D, Oster G. 2004. Waves and aggregation patterns in myxobacteria. Proc Natl Acad Sci U S A 101:4256–4261. https://doi.org/10.1073/pnas.0400704101.

35. Botev ZI, Grotowski JF, Kroese DP. 2010. Kernel density estimation via diffusion. Ann Stat 38:2916–2957. https://doi.org/10.1214/10-AOS799.

36. Yang Z, Lux R, Hu W, Hu C, Shi W. 2010. PilA localization affects extracellular polysaccharide production and fruiting body formation in *Myxococcus xanthus*. Mol Microbiol 76:1500–1513. https://doi.org/10.1111/j.1365-2958.2010.07180.x.

37. Curtis PD, Atwood J, Orlando R, Shimkets LJ. 2007. Proteins associated with the *Myxococcus xanthus* extracellular matrix. J Bacteriol 189:7634–7642. https://doi.org/10.1128/JB.01007-07.

38. Youderian P, Burke N, White DJ, Hartzell PL. 2003. Identification of genes required for adventurous gliding motility in *Myxococcus xanthus* with the transposable element mariner. Mol Microbiol 49:555–570. https://doi.org/10.1046/j.1365-2958.2003.03582.x.

39. Curtis PD, Taylor RG, Welch RD, Shimkets LJ. 2007. Spatial organization of *Myxococcus xanthus* during fruiting body formation. J Bacteriol 189:9126–9130. https://doi.org/10.1128/JB.01008-07.

40. Youderian P, Hartzell PL. 2006. Transposon insertions of magellan-4 that impair social gliding motility in *Myxococcus xanthus*. Genet 172:1397–1410. https://doi.org/10.1534/genetics.105.050542.

41. Bullock HA, Shen H, Boynton TO, Shimkets LJ. 2018. Fatty acid oxidation is required for *Myxococcus xanthus* development. J Bacteriol 200:e00572-17. https://doi.org/10.1128/JB.00572-17.

42. Jaqaman K, Loerke D, Mettlen M, Kuwata H, Grinstein S, Schmid SL, Danuser G. 2008. Robust single-particle tracking in live-cell time-lapse sequences. Nat Methods 5:695–702. https://doi.org/10.1038/nmeth.1237.

43. Xie C, Zhang H, Shimkets LJ, Igoshin OA. 2011. Statistical image analysis reveals features affecting fates of *Myxococcus xanthus* developmental aggregates. Proc Natl Acad Sci U S A 108:5915–5920. https://doi.org/10.1073/pnas.1018383108.

44. Kearns DB, Shimkets LJ. 1998. Chemotaxis in a gliding bacterium. Proc Natl Acad Sci U S A 95:11957–11962. https://doi.org/10.1073/pnas.95.20.11957.