



Connecting the Dots between Mechanosensitive Channel Abundance, Osmotic Shock, and Survival at Single-Cell Resolution

 Griffin Chure,^a Heun Jin Lee,^d Akiko Rasmussen,^c Rob Phillips^{a,b,d}

^aDepartment of Biology and Biological Engineering, California Institute of Technology, Pasadena, California, USA

^bDepartment of Physics, California Institute of Technology, Pasadena, California, USA

^cInstitute of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen, United Kingdom

^dDepartment of Applied Physics, California Institute of Technology, Pasadena, California, USA

ABSTRACT Rapid changes in extracellular osmolarity are one of many insults microbial cells face on a daily basis. To protect against such shocks, *Escherichia coli* and other microbes express several types of transmembrane channels that open and close in response to changes in membrane tension. In *E. coli*, one of the most abundant channels is the mechanosensitive channel of large conductance (MscL). While this channel has been heavily characterized through structural methods, electrophysiology, and theoretical modeling, our understanding of its physiological role in preventing cell death by alleviating high membrane tension remains tenuous. In this work, we examine the contribution of MscL alone to cell survival after osmotic shock at single-cell resolution using quantitative fluorescence microscopy. We conducted these experiments in an *E. coli* strain which is lacking all mechanosensitive channel genes save for MscL, whose expression was tuned across 3 orders of magnitude through modifications of the Shine-Dalgarno sequence. While theoretical models suggest that only a few MscL channels would be needed to alleviate even large changes in osmotic pressure, we find that between 500 and 700 channels per cell are needed to convey upwards of 80% survival. This number agrees with the average MscL copy number measured in wild-type *E. coli* cells through proteomic studies and quantitative Western blotting. Furthermore, we observed zero survival events in cells with fewer than ~100 channels per cell. This work opens new questions concerning the contribution of other mechanosensitive channels to survival, as well as regulation of their activity.

IMPORTANCE Mechanosensitive (MS) channels are transmembrane protein complexes which open and close in response to changes in membrane tension as a result of osmotic shock. Despite extensive biophysical characterization, the contribution of these channels to cell survival remains largely unknown. In this work, we used quantitative video microscopy to measure the abundance of a single species of MS channel in single cells, followed by their survival after a large osmotic shock. We observed total death of the population with fewer than ~100 channels per cell and determined that approximately 500 to 700 channels were needed for 80% survival. The number of channels we found to confer nearly full survival is consistent with the counts of the numbers of channels in wild-type cells in several earlier studies. These results prompt further studies to dissect the contribution of other channel species to survival.

KEYWORDS biophysics, mechanosensation, osmoregulation, quantitative methods, single cell

Received 1 August 2018 **Accepted** 5 August 2018

Accepted manuscript posted online 10 September 2018

Citation Chure G, Lee HJ, Rasmussen A, Phillips R. 2018. Connecting the dots between mechanosensitive channel abundance, osmotic shock, and survival at single-cell resolution. *J Bacteriol* 200:e00460-18. <https://doi.org/10.1128/JB.00460-18>.

Editor Conrad W. Mullineaux, Queen Mary University of London

Copyright © 2018 Chure et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Rob Phillips, phillips@pboc.caltech.edu.

G.C. and H.J.L. contributed equally.

Changes in the extracellular osmolarity can be a fatal event for the bacterial cell. Upon a hypo-osmotic shock, water rushes into the cell across the membrane, leaving the cell with no choice but to equalize the pressure. This equalization occurs either through damage to the cell membrane (resulting in death) or through the regulated flux of water molecules through transmembrane protein channels (Fig. 1A). Such proteinaceous pressure release valves have been found across all domains of life, with the first bacterial channel being described in 1987 (1). Over the past 30 years, several more channels have been discovered, described, and (in many cases) biophysically characterized. *Escherichia coli*, for example, has seven of these channels (one MscL and six MscS homologs), which have various degrees of conductance, gating mechanisms, and expression levels. While they have been the subject of much experimental and theoretical dissection, much remains a mystery with regard to the roles their abundance and interaction with other cellular processes play in the greater context of physiology (2–8).

Of the seven channels in *E. coli*, the mechanosensitive channel of large conductance (MscL) is one of the most abundant and the best characterized. This channel has a large conductance (3 nS) and mediates the flux of water molecules across the membrane via an ~3-nm-wide pore in the open state (9, 10). Molecular dynamics simulations indicate that a single open MscL channel permits the flux of 4×10^9 water molecules per second, which is an order of magnitude larger than a single aquaporin channel (BNID 100479) (11, 12). This suggests that having only a few channels per cell could be sufficient to relieve even large changes in membrane tension. Electrophysiological experiments have suggested a small number of channels per cell (13, 14); however, more recent approaches using quantitative Western blotting, fluorescence microscopy, and proteomics have measured several hundred MscL per cell (3, 15, 16). To further complicate matters, the expression profile of MscL appears to depend on the growth phase, available carbon source, and other environmental challenges (3, 16, 17). While there are likely more than just a few channels per cell, why cells seem to need so many and the biological rationale behind their condition-dependent expression both remain a mystery.

While their biochemical and biophysical characteristics have received much attention, their connection to cell survival is understudied. Drawing such a direct connection between channel copy number and survival requires quantitative *in vivo* experiments. To our knowledge, the work presented in van den Berg et al. (8) is the first attempt to simultaneously measure channel abundance and survivability for a single species of mechanosensitive channel. While the measurement of channel copy number was performed at the level of single cells using superresolution microscopy, survivability after a hypo-osmotic shock was assessed in bulk plating assays, which rely on serial dilutions of a shocked culture followed by counting the number of resulting colonies after incubation. Such bulk assays have long been the standard for querying cell viability after an osmotic challenge. While they have been highly informative, they reflect only the mean survival rate of the population, obfuscating the variability in survival of members of the population. The stochastic nature of gene expression results in a noisy distribution of MscL channels rather than a single value, meaning those cells found in the long tails of the distribution have quite different survival rates than the mean but are lost in the final calculation of survival probability.

In this work, we present an experimental system to quantitatively probe the interplay between MscL copy number and survival at single-cell resolution, as shown in Fig. 1B. We generated an *E. coli* strain in which all seven mechanosensitive channels had been deleted from the chromosome, followed by the chromosomal integration of a single gene encoding an MscL-superfolder green fluorescent protein (sfGFP) fusion protein. To explore copy number regimes beyond those of the wild-type expression level, we modified the Shine-Dalgarno sequence of this integrated construct, allowing us to cover nearly 3 decades of MscL copy number. To probe survivability, we exposed cells to a large hypo-osmotic shock at controlled rates in a flow cell under a microscope, allowing the observation of the single-cell channel copy number and the resulting fate

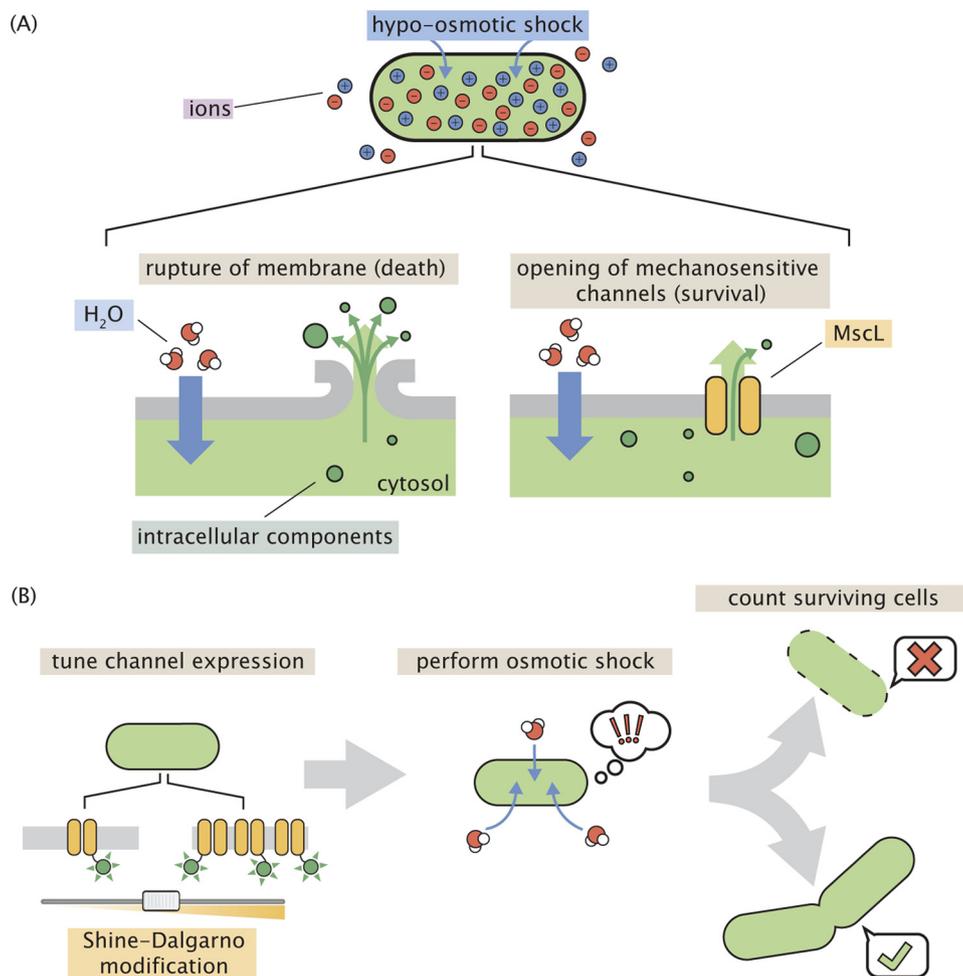


FIG 1 Role of mechanosensitive channels during hypo-osmotic shock. (A) A hypo-osmotic shock results in a large difference in the osmotic strength between the intracellular and extracellular spaces. As a result, water rushes into the cell to equalize this gradient, increasing the turgor pressure and tension in the cell membrane. If no mechanosensitive channels are present and membrane tension is high (left), the membrane ruptures, releasing intracellular content into the environment and resulting in cell death. If mechanosensitive channels are present (right) and membrane tension is beyond the gating tension, the mechanosensitive channel MscL opens, releasing water and small intracellular molecules into the environment, thus relieving pressure and membrane tension. (B) The experimental approach undertaken in this work. The number of mechanosensitive channels tagged with a fluorescent reporter is tuned through modification of the Shine-Dalgarno sequence of the *mscL* gene. The cells are then subjected to a hypo-osmotic shock and the number of surviving cells are counted, allowing the calculation of a survival probability.

of single cells. With this large set of single-cell measurements, we approach the calculation of survival probability in a manner that is free of binning bias, which allows the reasonable extrapolation of survival probability to copy numbers outside the observed range. In addition, we show that several hundred channels are needed to convey high rates of survival and observe a minimum number of channels needed to permit any degree of survival.

RESULTS

Quantifying the single-cell MscL copy number. The principal goal of this work is to examine the contribution of a single mechanosensitive channel species to cell survival under a hypo-osmotic shock. While this procedure could be performed for any species of channel, we chose MscL as it is the most well characterized and one of the most abundant species in *E. coli*. To probe the contribution of MscL alone, we integrated an *mscL* gene encoding an MscL-superfolder GFP (sfGFP) fusion into a strain in which all seven known mechanosensitive channel genes were deleted from the

Downloaded from <http://jlb.asm.org/> on November 17, 2018 by guest

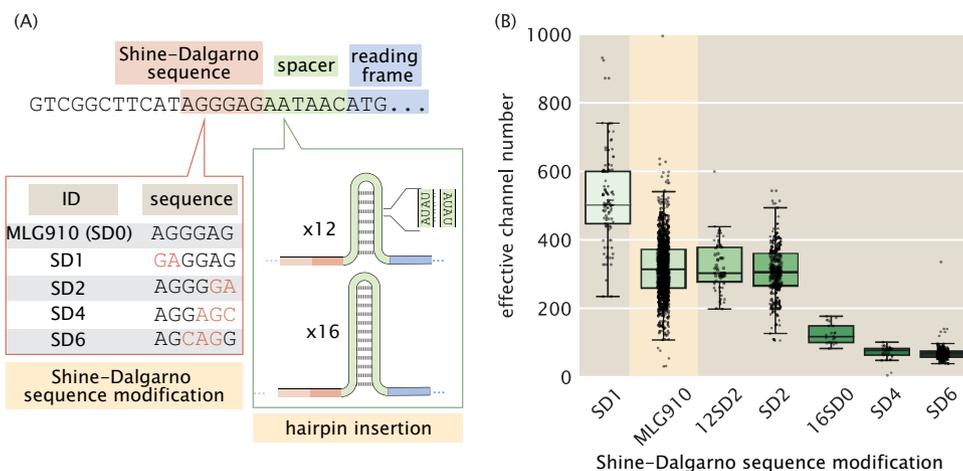


FIG 2 Control of MscL expression and calculation of channel copy number. (A) Schematic view of the expression modifications performed in this work. The beginning portion of the native *mscL* sequence is shown with the Shine-Dalgarno sequence, spacer region, and start codon shaded in red, green, and blue, respectively. The Shine-Dalgarno sequence was modified through the Salis laboratory's RBS (Ribosome Binding Site) Calculator (18, 19). The wild-type sequence (MLG910) is shown in black, with mutations for the four Shine-Dalgarno mutants highlighted in red. Expression was further modified by the insertion of repetitive AT bases into the spacer region, generating hairpins of various length which acted as a thermodynamic barrier for translation initiation. (B) Variability in effective channel copy number is computed using the standard candle. The boxes represent the interquartile region of the distribution, the center line displays the median, and the whiskers represent 1.5 times the maximum and minimum of the interquartile region. Individual measurements are denoted as black points. Data for the strain used for calibration of channel copy number (MLG910) are highlighted in yellow.

chromosome (5). Chromosomal integration imposes strict control on the gene copy number compared to expression from plasmid-borne systems, which is important to minimize variation in channel expression across the population and provide conditions more representative of native cell physiology. Abrogation of activity, mislocalization, or cytotoxicity are all inherent risks associated with creating chimeric reporter constructs. In Supplement SA in the supplemental material, we carefully dissect the functionality of this protein through electrophysiology (Fig. S1), measure the rate of fluorophore maturation (Fig. S2), and quantify potential aggregates (Fig. S3 and S4). To the best of our knowledge, the MscL-sfGFP fusion protein functions identically to the wild type, allowing us to confidently draw conclusions about the physiological role this channel plays in wild-type cells.

To modulate the number of MscL channels per cell, we developed a series of mutants which were designed to decrease the expression relative to that of the wild type. These changes involved direct alterations of the Shine-Dalgarno sequence, as well as the inclusion of AT hairpins of various lengths directly upstream from the start codon, which influences the translation rate and, hence, the number of MscL proteins produced (Fig. 2A). The six Shine-Dalgarno sequences used in this work were chosen using the RBS (ribosome binding site) Calculator from the Salis laboratory at the Pennsylvania State University (18, 19). While the designed Shine-Dalgarno sequence mutations decreased the expression relative to that of the wild type as intended, the distribution of expression levels was remarkably wide, spanning an order of magnitude.

To measure the number of MscL channels per cell, we determined a fluorescence calibration factor to translate arbitrary fluorescence units per cell to protein copy numbers. While there have been numerous techniques developed over the past decade to directly measure this calibration factor, such as quantifying single-molecule photobleaching constants or measuring the binomial partitioning of fluorescent proteins upon cell division (3, 20), we used *a priori* knowledge of the mean MscL-sfGFP expression level of a particular *E. coli* strain to estimate the average fluorescence of a single channel. In the work of Bialecka-Fornal et al. (3), the authors used single-molecule photobleaching and quantitative Western blotting to probe the expression of

MscL-sfGFP under a wide range of growth conditions. To compute a calibration factor, we used strain MLG910 [*E. coli* K-12 MG1655 $\phi(\text{mscL-sfGFP})$] as a “standard candle,” highlighted in yellow in Fig. 2B. This standard-candle strain was grown and imaged under conditions identical to those under which the MscL count was determined through fluorescence microscopy. The calibration factor was computed by dividing the mean total cell fluorescence by the known MscL copy number, resulting in a measure of arbitrary fluorescence units per MscL channel. Details regarding this calculation and appropriate propagation of error, as well as its sensitivity to various growth media, can be found in Materials and Methods, as well as in Supplement SB (Fig. S5 to S8).

While it is seemingly straightforward to use this calibration factor to determine the total number of channels per cell for wild-type or highly expressing strains, the calculation for the lowest-expressing strains is complicated by distorted cell morphology. We observed that, as the channel copy number decreased, cellular morphology became increasingly aberrant, with filamentous, bulging, and branched cells becoming more abundant (Fig. S7A). This morphological defect has been observed when altering the abundance of several species of mechanosensitive channels, suggesting that they play an important role in general architectural stability (3, 4). As these aberrant morphologies can vary widely in size and shape, calculating the number of channels per cell becomes a more nuanced endeavor. For example, taking the total MscL copy number for these cells could skew the final calculation of survival probability, as a large but severely distorted cell would be interpreted as having more channels than a smaller, wild-type-shaped cell (Fig. S7B). To correct for this pathology, we computed the average expression level per unit area for each cell and multiplied this by the average cellular area of our standard-candle strain, which is morphologically indistinguishable from wild-type *E. coli*, allowing the calculation of an effective channel copy number. The effect of this correction can be seen in Fig. S7C and D, which illustrate that there is no other correlation between cell area and channel expression.

Our calculation of the effective channel copy number for our suite of Shine-Dalgarno mutants is shown in Fig. 2B. The expression levels of these strains cover nearly 3 orders of magnitude, with the extremes ranging from approximately 4 channels per cell to nearly 1,000. While the mean values of each strain are somewhat distinct, the distributions show a large degree of overlap, making one strain nearly indistinguishable from another. This variance is a quantity that is lost in the context of bulk scale experiments but can be accounted for via single-cell methods.

Performing a single-cell hypo-osmotic challenge assay. To measure the channel copy number of a single cell and query its survival after a hypo-osmotic shock, we used a custom-made flow cell in which osmotic shock and growth can be monitored in real time using video microscopy (Fig. 3A). The design and characterization of this device have been described in depth previously (4) and are briefly described in Materials and Methods. Using this device, cells were exposed to a large hypo-osmotic shock by switching between LB Lennox medium supplemented with 500 mM NaCl and LB Lennox medium alone. Cells containing the six Shine-Dalgarno modifications shown in Fig. 2B (excluding MLG910) were subjected to hypo-osmotic shocks at controlled rates while under observation. After the application of the osmotic shock, the cells were imaged every 60 s for 4 to 6 h. Each cell was monitored over the outgrowth period and was manually scored as a survivor, a fatality, or inconclusive. The criteria used for scoring death were the same as those previously described by Bialecka-Fornal et al. (4). Survivors were defined as cells that underwent multiple divisions postshock. To qualify as survivors, cells must undergo at least two divisions, although more typically, four to eight divisions are observed without any signs of slowing down. Imaging was stopped when the survivor cells began to go out of focus or overlap each other. Survivors do not show any sign of ceasing division. More information regarding this classification can be found in Materials and Methods, as well as in Supplement SC (Fig. S9 to S10 and Tables S1 and S2). The experimental protocol can be seen in brief in Fig. 3B.

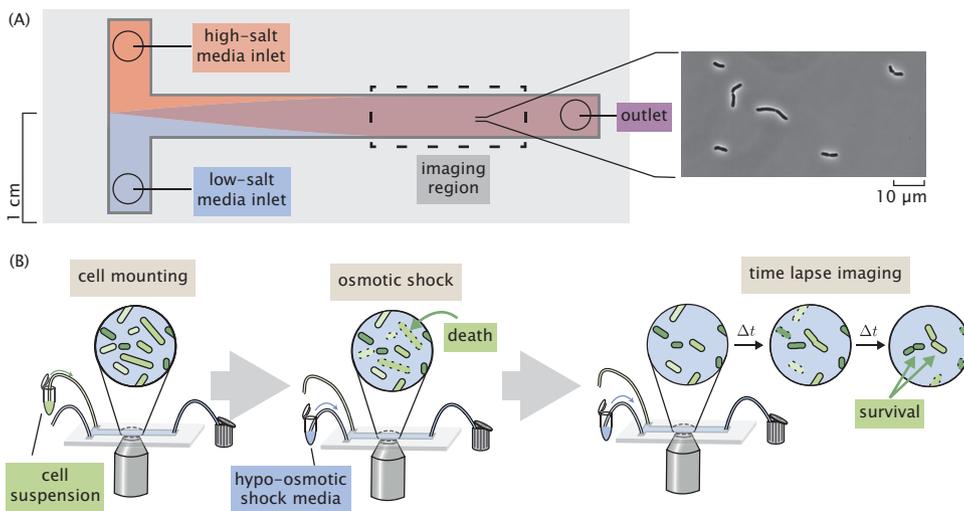


FIG 3 Experimental approach to measuring survival probability. (A) Layout of a home-made flow cell for subjecting cells to osmotic shock. Cells are attached to a polyethylenimine-functionalized surface of a glass coverslip within the flow chamber by loading a dilute cell suspension through one of the inlets. (B) The typical experimental procedure. Cells are loaded into a flow chamber as shown in panel A and mounted to the glass coverslip surface. Cells are subjected to a hypo-osmotic shock by flowing hypotonic medium into the flow cell. After shock, the cells are monitored for several hours and surviving cells are identified.

Due to the extensive overlap in expression levels between the different Shine-Dalgarno mutants (Fig. 2B), computing the survival probability by treating each mutant as an individual bin obfuscates the relationship between channel abundance and survival. To more thoroughly examine this relationship, all measurements were pooled, with each cell being treated as an individual experiment. The hypo-osmotic shock applied in these experiments was varied across a range of 0.02 Hz (complete exchange in 50 s) to 2.2 Hz (complete exchange in 0.45 s). Rather than pooling this wide range of shock rates into a single data set, we chose to separate the data into “slow shock” (<1.0 Hz) and “fast shock” (≥ 1.0 Hz) classes. Other groupings of shock rates were explored and are discussed in Supplement SD (Fig. S11 and S12). The cumulative distributions of channel copy numbers separated by survival rates are shown in Fig. 4. In these experiments, survival was never observed for a cell containing fewer than approximately 100 channels per cell, indicated by the red stripes in Fig. 4. This suggests that there is a minimum number of channels needed for survival that is on the order of 100 per cell. We also observe a slight shift in the surviving fraction of the cells toward

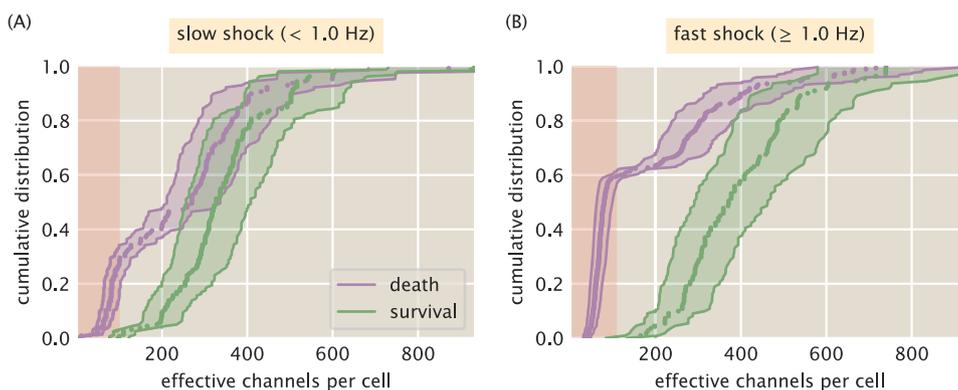


FIG 4 Distributions of survival and death as functions of effective channel number. (A) Empirical cumulative distributions of channel copy numbers separated by survival or death after a slow (<1.0 Hz) osmotic shock. (B) Empirical cumulative distributions for a fast (≥ 1.0 Hz) osmotic shock. Shaded green and purple regions represent the 95% credible region of the effective channel number calculation for each cell. The shaded red stripe signifies the range of channels in which no survival was observed.

higher effective copy numbers, which matches our intuition that including more mechanosensitive channels increases the survival probability.

Prediction of survival probability as a function of channel copy number. There are several ways by which the survival probability can be calculated. The most obvious approach would be to group each individual Shine-Dalgarno mutant as a single bin and compute the average MscL copy number and the survival probability. Binning by strain is the most frequently used approach for such measurements and has provided valuable insight into the qualitative relationship of survival to other physiological factors (4, 8). However, the copy number distribution for each Shine-Dalgarno mutant (Fig. 2B) is remarkably wide and overlaps with those of the other strains. We argue that this coarse-grained binning negates the benefits of performing single-cell measurements, as two strains with different means but overlapping quartiles would be treated as distinctly different distributions.

Another approach would be to pool all data, irrespective of the Shine-Dalgarno mutation, and bin by a defined range of channels. Depending on the width of the bin, this could allow for finer resolution of the quantitative trend, but the choice of the bin width is arbitrary with the *a priori* knowledge that is available. Drawing a narrow bin width can easily restrict the number of observed events to small numbers where the statistical precision of the survival probability is lost. On the other hand, drawing wide bins increases the precision of the estimate but becomes further removed from a true single-cell measurement and represents a population mean, even though it may be a smaller population than binning by the Shine-Dalgarno sequence alone. In both of these approaches, it is difficult to extrapolate the quantitative trend outside the experimentally observed region of channel copy number. Here, we present a method to estimate the probability of survival for any channel copy number, even those that lie outside the experimentally queried range.

To quantify the survival probability while maintaining single-cell resolution, we chose to use a logistic regression model which does not require grouping data into arbitrary bins and treats each cell measurement as an independent experiment. Logistic regression is an inferential method to model the probability of a Boolean or categorical event (such as survival or death) given one or several predictor variables and is commonly used in medical statistics to compute survival rates and dose-response curves (21, 22). The primary assumption of logistic regression is that the log-odds probability of survival p_s is linearly dependent on the predictor variable, in our case the log channels per cell N_c with a dimensionless intercept β_0 and slope β_1 ,

$$\log \frac{p_s}{1-p_s} = \beta_0 + \beta_1 \log N_c \quad (1)$$

Under this assumption of linearity, β_0 is the log-odds probability of survival with no MscL channels. The slope β_1 represents the change in the log-odds probability of survival conveyed by a single channel. As the calculated number of channels in this work spans nearly 3 orders of magnitude, it is better to perform this regression on $\log N_c$, as regressing on N_c directly would give undue weight to lower channel copy numbers due to the sparse sampling of high-copy-number cells. The functional form shown in equation 1 can be derived directly from Bayes' theorem and is shown in Supplement SE. If one knows the values of β_0 and β_1 , the survival probability can be expressed as

$$p_s = \frac{1}{1 + N_c^{-\beta_1} e^{-\beta_0}} \quad (2)$$

In this analysis, we used Bayesian inferential methods to determine the most likely values of the coefficients (described in detail in Fig. S13 and S14 in Supplement SE).

The results of the logistic regression are shown in Fig. 5. We see a slight rightward shift of the survival probability curve under fast shock relative to the case for slow shock, reaffirming the conclusion that survival is also dependent on the rate of osmotic shock (4). This rate dependence has been observed for cells expressing MscL alongside

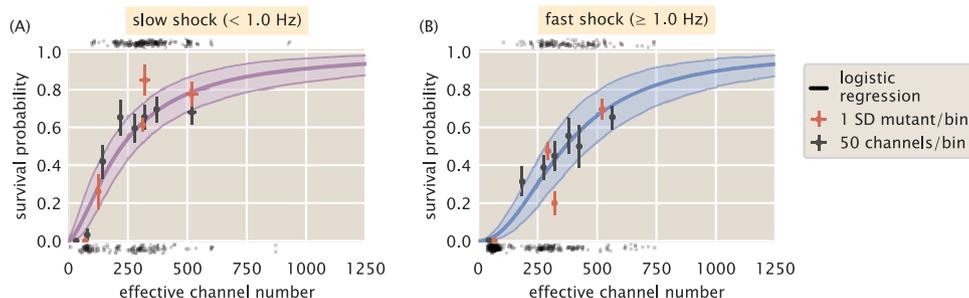


FIG 5 Probability of survival as a function of MscL copy number. (A) Estimated survival probability for survival under slow shock as a function of channel copy number. (B) Estimated survival probability of survival under a fast shock as a function of channel copy number. Solid curves correspond to the most probable survival probability from a 1-dimensional logistic regression. Shaded regions represent the 95% credible regions. Points at the top and bottom of plots represent measurements of individual cells which survived and perished, respectively. The red and black points within the plots correspond to the survival probabilities estimated via binning by Shine-Dalgarno sequence and binning by groups of 50 channels per cell, respectively. Horizontal error bars represent the standard errors of the means from at least 25 measurements. Vertical error bars represent the certainty of the probability estimate given n survival events from N total observations.

other species of mechanosensitive channels but not for MscL alone. This suggests that MscL responds differently to different rates of shock, highlighting the need for further study of rate dependence and the coordination between different species of mechanosensitive channels. The results in Fig. 5 also show that several hundred channels are required to provide appreciable protection from osmotic shock. For a survival probability of 80%, a cell must have approximately 500 to 700 channels per cell for a fast and a slow shock, respectively. The results from the logistic regression are shown as continuous colored curves in Fig. 5; the individual cell measurements separated by survival and death are shown at the top and bottom of each plot, respectively, and are included to provide a sense of sampling density.

Over the explored range of MscL copy numbers, we observed a maximum of 80% survival for any binning method. The remaining 20% survival may be attained when the other species of mechanosensitive channels are expressed alongside MscL. However, it is possible that the flow cell method performed in this work lowers the maximal survival fraction, as the cells are exposed to several, albeit minor, mechanical stresses, such as loading into the flow cell and chemical adherence to the glass surface. To ensure that the results from logistic regression accurately describe the data, we can compare the survival probabilities to those obtained using the binning methods described earlier (Fig. 5, red and black points). Nearly all binned data fall within the error of the prediction (see Materials and Methods for definition of error bars in probability data), suggesting that this approach accurately reflects the survival probability and giving license to extrapolate the estimation of survival probability to regions outside our experimentally explored copy number regime.

Thus far, we have dictated that for a given rate of osmotic shock (i.e., “fast” or “slow”), the survival probability is dependent only on the number of channels. In Fig. S13, we show the results of including other predictor variables, such as area and shock rate alone. In such cases, including other predictors resulted in pathological curves, showing that channel copy number is the most informative of the available predictor variables.

DISCUSSION

One of the most challenging endeavors in the biological sciences is linking the microscopic details of cellular components to the macroscale physiology of the organism. This formidable task has been undertaken repeatedly in the recent history of biology, especially in the era of DNA sequencing and single-molecule biochemistry. For example, the scientific community has been able to connect sickle-cell anemia to a single amino acid substitution in hemoglobin which promotes precipitation under a

change in O₂ partial pressure (23–25). Others have assembled a physical model that quantitatively describes chemosensation in bacteria (26), in which the arbiter of sensory adaptation is the repeated methylation of chemoreceptors (27–30). In the past ~50 years alone, numerous biological and physical models of the many facets of the central dogma have been assembled that give us a sense of the interplay between the genome and physiology. For example, the combination of biochemical experimentation and biophysical models have given us a picture of how gene dosage affects furrow positioning in *Drosophila* (31), how recombination of V(D)J gene segments generates an extraordinarily diverse antibody repertoire (32–34), and how telomere shortening through DNA replication is intrinsically tied to cell senescence (35, 36), to name just a few of many such examples.

By no means are we finished with any of these topics. Rather, it is quite the opposite, in the sense that having a handle on the biophysical knobs that tune the behavior opens the door to a litany of new scientific questions. In the case of mechanosensation and osmoregulation, we have only recently been able to determine some of the basic facts that allow us to approach this fascinating biological phenomenon biophysically. The dependence of survival on mechanosensitive channel abundance is a key quantity that is missing from our collection of critical facts. To our knowledge, this work represents the first attempt to quantitatively control the abundance of a single species of mechanosensitive channel and examine the physiological consequences in terms of survival probability at single-cell resolution. Our results reveal two notable quantities. First, out of the several hundred single-cell measurements, we never observed a cell which had fewer than approximately 100 channels per cell and survived an osmotic shock, irrespective of the shock rate. The second is that between 500 and 700 channels per cell are needed to provide $\geq 80\%$ survival, depending on the shock rate.

Only recently has the relationship between the MscL copy number and the probability of survival been approached experimentally. In the work of van den Berg et al., the authors examined the contribution of MscL to survival in a genetic background where all other known mechanosensitive channels had been deleted from the chromosome and plasmid-borne expression of an MscL-mEos3.2 fusion was tuned through an isopropyl- β -D-thiogalactopyranoside (IPTG)-inducible promoter (8). In that work, they measured the single-cell channel abundance through superresolution microscopy and queried survival through bulk assays. They report a nearly linear relationship between survival and copy number, with approximately 100 channels per cell conveying 100% survival. This number is significantly smaller than our observation of approximately 100 channels as the minimum number needed to convey any observable degree of survival.

The disagreement between the numbers reported in this work and by van den Berg et al. (8) may arise partially from subtle differences in the experimental approach. The primary practical difference is the magnitude of the osmotic shock. van den Berg et al. (8) applied an approximately 600-mosmol downshock in bulk, whereas we applied a 1-osmol downshock, which would lead to lower survival (37). In their work, the uncertainty in both the MscL channel count and survival probability is roughly 30% (Fig. S14 in the supplemental material). Given this uncertainty, it is reasonable to interpret that the number of channels needed for complete protection from osmotic downshock is between 100 and 250 per cell. The uncertainty in determining the number of channels per cell is consistent with the observed width of the channel number distribution of the Shine-Dalgarno sequence mutants used in this work (Fig. 2B). A unique property of the single-cell measurements performed in this work is the direct observation of the survival or death of individual cells. We find that the results of morphological classification and classification through propidium iodide staining agree within 1% (see Supplement SC in the supplemental material). Bulk plating assays, such as are used by van den Berg et al. (8), rely on colony formation and outgrowth to determine survival probability. As is reported in their supplemental information, the precision in this measurement is around 30% (Fig. S14). Accounting for this uncertainty brings both measurements within a few fold, where we still consistently observe lower

survival for a given channel number. This remaining disagreement may be accounted for by systematic uncertainty in both experimental methods.

For example, variations in the lengths of outgrowth, shock rates, and counting statistics could lead to a bias toward higher observed survival rates in ensemble plating assays. During the outgrowth phase, the control sample not exposed to an osmotic shock is allowed to grow for approximately 30 min in a high-salt medium before plating. The shocked cells, however, are allowed to grow in a low-salt medium. We have found that the difference between the growth rates under these two conditions can be appreciable (approximately 35 min versus 20 min, respectively), as can be seen by the results in Fig. S2. Cells that survived an osmotic shock may have a growth advantage relative to the control sample if the shock-induced lag phase is less than the outgrowth, leading to higher observed survival rates (37). This is one possible explanation for the survival rates in excess of 100% that are reported. Cells that survived an osmotic shock may have a growth advantage relative to the normalization sample if the shock-induced lag phase is less than the outgrowth, leading to higher observed survival rates, even surpassing 100%. We have performed these assays ourselves and have observed survival rates above 100% (ranging from 110% to 125%) with an approximately 30% error (see Fig. S3 in the work of Bialecka-Fornal et al. [3]), which we concluded arose from differences in growth rate. We also note that survival rates greater than 100% are observed by van den Berg et al. (Fig. S14) (8). For strains that have survival rates between 80% and 100%, the uncertainty is typically large, making it difficult to make precise statements regarding when full survival is achieved.

It has been shown that there is a strong inverse relationship between the rate of osmotic shock and survival probability (4). Any experiment in which the shock was applied more slowly or quickly than in another would bias toward higher or lower survivability, respectively. The shocks applied in bulk assays are often performed manually, which can be highly variable. We note that in our experiments, we frequently observed cells which did not separate and formed chains of two or more cells (Fig. S9 and S10). In plating assays, it is assumed that colonies arise from a single founding cell; however, a colony formed by a cluster of living and dead cells would be interpreted as a single surviving cell, effectively masking the death of the others in the CFU. This too could bias the measurement toward higher survival rates. Single-cell shock experiments can also have systematic errors which can bias the results toward lower survival rates. Such errors are associated with handling of the cells, such as loading into the flow cell, which may cause shear damage, adhering the cells to the coverslip, and any chemical perturbations introduced by the dye used to measure the shock rate.

Despite these experimental differences, the results of this work and of van den Berg et al. (8) are in agreement that MscL must be present at the level of 100 or more channels per cell in wild-type cells to convey appreciable survival. As both of these works were performed in a strain in which the only mechanosensitive channel was MscL, it remains unknown how the presence of the other channel species would alter the number of MscL needed for complete survival. In our experiments, we observed a maximum survival probability of approximately 80% even with close to 1,000 MscL channels per cell. It is possible that the combined efforts of the six other mechanosensitive channels would make up for some if not all of the remaining 20%. To explore the contribution of another channel to survival, van den Berg et al. (8) also queried the contribution of MscS, another mechanosensitive channel, to survival in the absence of any other species of mechanosensitive channel. It was found that over the explored range of MscS channel copy numbers, the maximum survival rate was approximately 50%, suggesting that different mechanosensitive channels have an upper limit to how much protection they can confer. Both the results of van den Berg et al. (8) and our work show that there is still much to be learned with respect to the interplay between the various species of mechanosensitive channel, as well as their regulation.

Recent work has shown that both the magnitude and the rate of osmotic downshock are important factors in determining cell survival (4). In this work, we show that this finding holds true for a single species of mechanosensitive channel, even at high

TABLE 1 Measured cellular copy numbers of MscL

Reported no. of channels per cell	Method	Reference
480 ± 103 ^a	Western blotting	3
560 ^b	Ribosomal profiling	39
331 ^b	Mass spectrometry	15
583 ^b	Mass spectrometry	16
4 or 5	Electrophysiology	17
10–100	Electrophysiology	13
10–15	Electrophysiology	40

^aMean value ± standard deviation.

^bMscL channel copy number was inferred from the total number of MscL peptides detected.

levels of expression. One might naively expect that this rate-dependent effect would disappear once a certain threshold of channels had been met. Our experiments, however, show that even at nearly 1,000 channels per cell, the predicted survival curves for a slow (<1.0 Hz) and fast (≥1.0 Hz) shock are shifted relative to each other, with the fast shock predicting lower rates of survival. This suggests that either we have not reached this threshold in our experiments or there is more to understand about the relationship between the abundance, channel species, and shock rate.

Some experimental and theoretical treatments suggest that only a few copies of MscL or MscS should be necessary for 100% protection, given our knowledge of the conductance and the maximal water flux through the channel in its open state (11, 38). However, recent proteomic studies have revealed average MscL copy numbers to be in the range of several hundred per cell, depending on the condition, as can be seen by the data in Table 1 (15, 16, 39). Studies focusing solely on MscL have shown similar counts through quantitative Western blotting and fluorescence microscopy (3). Electrophysiology studies have told another story, with copy number estimates ranging between 4 and 100 channels per cell (17, 40). These measurements, however, measure the number of active channels. The factors regulating channel activity in these experiments could include perturbations during the sample preparation or reflect some unknown mechanism of regulation, such as the presence or absence of interacting cofactors (41). The work described here, on the other hand, measures the maximum number of channels that could be active and may be able to explain why the channel abundance is higher than estimated by theoretical means. There remains much more to be learned about the regulation of activity in these systems. As the *in vivo* measurement of protein copy number becomes accessible through novel single-cell and single-molecule methods, we will continue to collect more facts about this fascinating system and hopefully connect the molecular details of mechanosensation with perhaps the most important physiological response—life or death.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The bacterial strains are described in Table S3 in the supplemental material. The parent strain for the mutants used in this study was MJF641 (5) (generously provided by Samantha Miller and Ian Booth), a strain which had all seven mechanosensitive channels deleted. The MscL-sfGFP-coding region from MLG910 (3) was integrated into MJF641 by P1 transduction, creating strain D6LG-Tn10. Selection pressure for MscL integration was created by incorporating an osmotic shock into the transduction protocol, which favored the survival of MscL-expressing strains relative to MJF641 by ~100-fold. Screening for integration candidates was based on the fluorescence expression of plated colonies. Successful integration was verified by sequencing. Attempts to transduce RBS-modified MscL-sfGFP-coding regions became increasingly inefficient as the targeted expression level of MscL was reduced. This was due to the decreasing fluorescence levels and survival rates of the integration candidates. Consequently, Shine-Dalgarno sequence modifications were made by inserting DNA oligonucleotides with lambda red-mediated homologous recombination, i.e., recombineering (42). The oligonucleotides had a designed mutation (Fig. 2) flanked by ~25 bp that matched the targeted MscL region (Table S4). A two-step recombineering process of selection followed by counterselection using a *tetA-sacB* gene fusion cassette (43) was chosen because of its capabilities to integrate with efficiencies comparable to those of P1 transduction and not leave antibiotic resistance markers or scar sequences in the final strain. To prepare strain D6LG-Tn10 for this scheme, the Tn10 transposon containing the *tetA* gene needed to be removed to avoid interference with the *tetA-sacB* cassette. Tn10 was removed from the middle of the *ycjM* gene with the primer Tn10delR (Table S2) by recombineering, creating strain D6LG (SD0). Counterselection against the *tetA* gene was promoted by using agar medium

with fusaric acid (43, 44). The *tetA-sacB* cassette was PCR amplified out of strain XTL298 using primers MscLSPSac and MscLSPSacR (Table S2). The cassette was integrated in place of the spacer region in front of the MscL start codon of D6LG (SD0) by recombineering, creating the intermediate strain D6LTetSac. Positive selection for cassette integration was provided by agar medium with tetracycline. Finally, the RBS-modifying oligonucleotides were integrated by replacing the *tetA-sacB* cassette by recombineering. Counterselection against both *tetA* and *sacB* was ensured by using agar medium with fusaric acid and sucrose (43), creating the Shine-Dalgarno mutant strains used in this work.

Strain cultures were grown in 5 ml of LB Lennox medium with antibiotic (apramycin) overnight at 37°C. The next day, 50 μ l of overnight culture was inoculated into 5 ml of LB Lennox medium with antibiotic and the culture was grown to an optical density at 600 nm (OD_{600}) of \sim 0.25. Subsequently, 500 μ l of that culture was inoculated into 5 ml of LB Lennox medium supplemented with 500 mM NaCl and the culture was regrown to an OD_{600} of \sim 0.25. A 1-ml aliquot was taken and used to load the flow cell.

Flow cell. All experiments were conducted in a home-made flow cell as shown in Fig. 3A. This flow cell has two inlets that allow media of different osmolarities to be exchanged over the course of the experiment. The imaging region is approximately 10 mm wide and 100 μ m in depth. All imaging took place within 1 to 2 cm of the outlet to avoid imaging cells within a nonuniform gradient of osmolarity. The interior of the flow cell was functionalized with a 1:400 dilution of polyethylenimine prior to the addition of cells, with the excess washed away with water. A dilute cell suspension in LB Lennox medium with 500 mM NaCl was loaded into one inlet, while the other was connected to a vial of LB medium with no NaCl. This hypotonic medium was clamped during the loading of the cells.

Once the cells had adhered to the polyethylenimine-coated surface, the excess cells were washed away with the 500 mM NaCl growth medium, which was followed by a small (\sim 20 μ l) air bubble. This air bubble forced the cells to lie flat against the imaging surface, improving the time-lapse imaging. Over the observation period, cells not exposed to an osmotic shock were able to grow for 4 to 6 divisions, showing that the flow cell does not directly impede cell growth.

Imaging conditions. All imaging was performed in a flow cell held at 30°C on a Nikon Ti-Eclipse microscope outfitted with a perfect focus system enclosed in a Haison environmental chamber (approximately 1°C regulation efficiency). The microscope was equipped with a 488-nm laser excitation source (Crystalaser) and a 520/35 laser optimized filter set (Semrock). The images were collected on an Andor iXon EM+ 897 electron-multiplying charge-coupled device (EMCCD) camera, and all microscope and acquisition operations were controlled via the open-source μ Manager microscope control software (45). Once cells were securely mounted onto the surface of the glass coverslip, between 15 and 20 positions containing 5 to 10 cells were marked and the coordinates recorded. At each position, a phase-contrast and a GFP fluorescence image were acquired for segmentation and subsequent measurement of channel copy number. To perform the osmotic shock, LB medium containing no NaCl was pulled into the flow cell through a syringe pump. To monitor the medium exchange, both the high-salt and no-salt LB media were supplemented with a low-affinity version of the calcium-sensitive dye Rhod-2 (250 nM; TEF Labs), which fluoresces when bound to Ca^{2+} . The no-salt medium was also supplemented with 1 μ M $CaCl_2$ to make the medium mildly fluorescent, and the exchange rate was calculated by measuring the fluorescence increase across an illuminated section of one of the positions. These images were collected in real time for the duration of the shock. The difference in measured fluorescence between the preshock images and those at the end of the shock set the scale of a 500 mM NaCl downshock. The rate was calculated by fitting a line to the middle region of this trace. Further details regarding this procedure can be found in the work of Bialecka-Fornal et al. (4).

Image processing. Images were processed using a combination of automated and manual methods. First, the expression of MscL was measured via segmenting individual cells or small clusters of cells in phase contrast and computing the mean pixel value of the fluorescence image for each segmented object. The fluorescence images were passed through several filtering operations which reduced high-frequency noise and corrected for uneven illumination of the excitation wavelength.

Survival or death classification was performed manually using the CellProfiler plug-in for ImageJ software (NIH). A survivor was defined as a cell which was able to undergo at least two division events after the osmotic downshock. Cell death was recognized by stark changes in cell morphology, including loss of phase contrast through ejection of cytoplasmic material, structural decomposition of the cell wall and membrane, and the inability to divide. To confirm that these morphological cues corresponded with cell death, we probed cell viability on a subset of our strains after osmotic shock through staining with propidium iodide, a DNA intercalating dye commonly used to identify dead cells (LIVE/DEAD BacLight bacterial-cell viability assay; Thermo Fisher). We found that our classification based on morphology agreed with that based on staining within 1%. More information regarding these experiments can be found in Supplement SC. Cells which detached from the surface during the postshock growth phase or those which became indistinguishable from other cells due to clustering were not counted as surviving or dead and were removed from the data set completely. A region of the cell was manually marked with 1.0 (survival) or 0.0 (death) by clicking on the image. The *xy* coordinates of the click and the assigned value were saved as an .xml file for that position.

The connection between the segmented cells and their corresponding manual markers was automated. As the manual markings were made on the first phase-contrast image after the osmotic shock, small shifts in the positions of the cell made one-to-one mapping with the segmentation mask nontrivial. The linkages between segmented cell and manual marker were made by computing all pairwise distances between the manual marker and the segmented cell centroid, taking the shortest distance as the true pairing. The linkages were then inspected manually and incorrect mappings were corrected as necessary.

All relevant statistics about the segmented objects, as well as the sample identity, date of acquisition, osmotic shock rate, and camera exposure time, were saved as .csv files for each individual experiment. A more in-depth description of the segmentation procedure and the relevant code can be accessed as a Jupyter Notebook at (http://rpgroup.caltech.edu/mscl_survival).

Calculation of effective channel copy number. To compute the MscL channel copy number, we relied on measuring the fluorescence level of a bacterial strain in which the mean MscL channel copy number was known via fluorescence microscopy (3). *E. coli* strain MLG910, which expresses the MscL-sfGFP fusion protein from the wild-type Shine-Dalgarno sequence, was grown under conditions identical to those described by Bialecka-Fornal et al. (4) in LB Miller medium (BD Medical Sciences) to an OD₆₀₀ of ~0.3. The cells were then diluted 10-fold, immobilized on a rigid 2% agarose substrate, placed on a glass-bottom petri dish, and imaged under the same conditions as described previously.

Images were taken of six biological replicates of MLG910 and were processed identically to those in the osmotic shock experiments. A calibration factor between the average cell fluorescence level and mean MscL copy number was then computed. We assumed that all measured fluorescence ($\langle I_{\text{tot}} \rangle$) (after filtering and background subtraction) was derived from the MscL-sfGFP fusion,

$$\langle I_{\text{tot}} \rangle = \alpha \langle N \rangle \quad (3)$$

in which α is the calibration factor and $\langle N \rangle$ is the mean cellular MscL-sfGFP copy number as reported by Bialecka-Fornal et al. (3). To correct for errors in segmentation, the intensity was computed as an areal density ($\langle I_A \rangle$) and was multiplied by the average cell area ($\langle A \rangle$) of the population. The calibration factor was therefore computed as

$$\alpha = \frac{\langle I_A \rangle \langle A \rangle}{\langle N \rangle} \quad (4)$$

We used Bayesian inferential methods to compute this calibration factor, taking measurement error and replicate-to-replicate variation into account. The resulting average cell area and calibration factor were used to convert the measured cell intensities from the osmotic shock experiments to cell copy number. The details of this inference are described in depth in Supplement SB in the supplemental material.

Logistic regression. We used Bayesian inferential methods to find the most probable values of the coefficients β_0 and β_1 and the appropriate credible regions (described in detail in Supplement SD in the supplemental material). Briefly, we used Markov chain Monte Carlo (MCMC) to sample from the log posterior distribution and took the mean value of the samples for each parameter as the most probable value. MCMC was performed using the Stan probabilistic programming language (46), and all models can be found in the GitHub repository (http://github.com/rpgroup-pboc/mscl_survival).

Calculation of survival probability error. The vertical error bars for the points shown in Fig. 5 represent our uncertainty in the survival probability given our measurement of n survivors out of a total N single-cell measurements. The probability distribution of the survival probability p_s given these measurements can be written using Bayes' theorem as

$$g(p_s | n, N) = \frac{f(n | p_s, N) g(p_s)}{f(n | N)} \quad (5)$$

where g and f represent probability density functions over parameters and data, respectively. The likelihood $f(n | p_s, N)$ represents the probability of measuring n survival events, given a total of N measurements each with a probability of survival p_s . This matches the story for the binomial distribution and can be written as

$$f(n | p_s, N) = \frac{N!}{n!(N-n)!} p_s^n (1-p_s)^{N-n} \quad (6)$$

To maintain maximal ignorance, we can assume that any value for p_s is valid that is in the range [0, 1]. This prior knowledge, represented by $g(p_s)$, can be written as

$$g(p_s) = \begin{cases} 1 & 0 \leq p_s \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

We can also assume maximal ignorance for the total number of survival events we could measure given N observations, $f(n | N)$. Assuming all observations are equally likely, this can be written as

$$f(n | N) = \frac{1}{N+1} \quad (8)$$

where the addition of one comes from the possibility of observing zero survival events. Combining equations 6, 7, and 8, the posterior distribution $g(p_s | n, N)$ is

$$g(p_s | n, N) = \frac{(N+1)!}{n!(N-n)!} p_s^n (1-p_s)^{N-n} \quad (9)$$

The most probable value of p_s , where the posterior probability distribution given by equation 9 is maximized, can be found by computing the point at which the derivative of the log posterior with respect to p_s goes to zero,

$$\frac{d \log g(p_s | n, N)}{d p_s} = \frac{n}{p_s} - \frac{N-n}{1-p_s} = 0 \quad (10)$$

Solving equation 10 for p_s gives the most likely value for the probability,

$$p_s^* = \frac{n}{N} \quad (11)$$

So long as $N \gg np_s^*$, equation 9 can be approximated as a Gaussian distribution with a mean p_s^* and a variance $\sigma_{p_s}^2$. By definition, the variance of a Gaussian distribution is computed as the negative reciprocal of the second derivative of the log posterior evaluated at $p_s = p_s^*$,

$$\sigma_{p_s}^2 = - \left(\frac{d^2 \log g(p_s | n, N)}{dp_s^2} \Big|_{p_s = p_s^*} \right)^{-1} \quad (12)$$

Evaluating equation 12 yields

$$\sigma_{p_s}^2 = \frac{n(N-n)}{N^3} \quad (13)$$

Given equations 11 and 13, the most likely survival probability and estimate of the uncertainty can be expressed as

$$p_s = p_s^* \pm \sigma_{p_s} \quad (14)$$

Data and software availability. All raw image data are freely available and are stored on the CaltechDATA Research Data Repository (47). The raw Markov chain Monte Carlo samples are stored as .csv files on CaltechDATA (48). All processed experimental data and Python and Stan code used in this work are freely available through our GitHub repository (http://github.com/rpgroup-pboc/mscl_survival) (49), accessible through <https://doi.org/10.5281/zenodo.1252524>. The scientific community is invited to fork our repository and open constructive issues.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/JB.00460-18>.

SUPPLEMENTAL FILE 1, PDF file, 4.9 MB.

SUPPLEMENTAL FILE 2, AVI file, 2.4 MB.

ACKNOWLEDGMENTS

We thank Nathan Belliveau, Maja Bialecka-Fornal, Justin Bois, Soichi Hirokawa, Jaspal Landman, Manuel Razo-Mejia, Muir Morrison, and Shyam Saladi for useful advice and discussion. We thank Don Court for strain XTL298, as well as Samantha Miller and Ian Booth at the University of Aberdeen for strain MJF641.

This work was supported by the National Institutes of Health grants number DP1 OD000217 (Director's Pioneer award), R01 GM085286, GM084211-A1, and GM118043-01 and by La Fondation Pierre Gilles de Gennes.

REFERENCES

- Martinac B, Buechner M, Delcour AH, Adler J, Kung C. 1987. Pressure-sensitive ion channel in *Escherichia coli*. Proc Natl Acad Sci U S A 84:2297–2301.
- Bavi N, Cortes DM, Cox CD, Rohde PR, Liu W, Deitmer JW, Bavi O, Strop P, Hill AP, Rees D, Corry B, Perozo E, Martinac B. 2016. The role of MscL amphipathic N terminus indicates a blueprint for bilayer-mediated gating of mechanosensitive channels. Nat Commun 7:11984. <https://doi.org/10.1038/ncomms11984>.
- Bialecka-Fornal M, Lee HJ, DeBerg HA, Gandhi CS, Phillips R. 2012. Single-cell census of mechanosensitive channels in living bacteria. PLoS One 7:e33077. <https://doi.org/10.1371/journal.pone.0033077>.
- Bialecka-Fornal M, Lee HJ, Phillips R. 2015. The rate of osmotic down-shock determines the survival probability of bacterial mechanosensitive channel mutants. J Bacteriol 197:231–237. <https://doi.org/10.1128/JB.02175-14>.
- Edwards MD, Black S, Rasmussen T, Rasmussen A, Stokes NR, Stephen TL, Miller S, Booth IR. 2012. Characterization of three novel mechanosensitive channel activities in *Escherichia coli*. Channels (Austin) 6:272–281. <https://doi.org/10.4161/chan.20998>.
- Naismith JH, Booth IR. 2012. Bacterial mechanosensitive channels—MscS: evolution's solution to creating sensitivity in function. Annu Rev Biophys 41:157–177. <https://doi.org/10.1146/annurev-biophys-101211-113227>.
- Ursell T, Phillips R, Kondev J, Reeves D, Wiggins PA. 2008. The role of lipid bilayer mechanics in mechanosensation, p 37–70. In Kamkin A, Kiseleva I (ed), Mechanosensitivity in cells and tissues. Vol 1. Mechanosensitive ion channels. Springer Science+Business Media, LLC, New York, NY. https://doi.org/10.1007/978-1-4020-6426-5_2.
- van den Berg J, Galbiati H, Rasmussen A, Miller S, Poolman B. 2016. On the mobility, membrane location and functionality of mechanosensitive channels in *Escherichia coli*. Sci Rep 6:32709. <https://doi.org/10.1038/srep32709>.
- Cruickshank CC, Minchin RF, Le Dain AC, Martinac B. 1997. Estimation of the pore size of the large-conductance mechanosensitive ion channel of *Escherichia coli*. Biophys J 73:1925–1931. [https://doi.org/10.1016/S0006-3495\(97\)78223-7](https://doi.org/10.1016/S0006-3495(97)78223-7).
- Haswell ES, Phillips R, Rees DC. 2011. Mechanosensitive channels: what can they do and how do they do it? Structure 19:1356–1369. <https://doi.org/10.1016/j.str.2011.09.005>.
- Louhivuori M, Risselada HJ, van der Giessen E, Marrink SJ. 2010. Release of content through mechano-sensitive gates in pressurized liposomes. Proc Natl Acad Sci U S A 107:19856–19860. <https://doi.org/10.1073/pnas.1001316107>.
- Milo R, Jorgensen P, Moran U, Weber G, Springer M. 2010. BioNumbers—the database of key numbers in molecular and cell biology. Nucleic Acids Res 38(Database issue):D750–D753. <https://doi.org/10.1093/nar/gkp889>.
- Booth IR, Edwards MD, Murray E, Miller S. 2005. The role of bacterial ion channels in cell physiology, p 291–312. In Kubalsi A, Martinac B (ed), Bacterial ion channels and their eukaryotic homologs. American Society for Microbiology, Washington, DC.
- Hase CC, Minchin RF, Kloda A, Martinac B. 1997. Cross-linking studies and membrane localization and assembly of radiolabelled large

- mechanosensitive ion channel (MscL) of *Escherichia coli*. *Biochem Biophys Res Commun* 232:777–782. <https://doi.org/10.1006/bbrc.1997.6370>.
15. Schmidt A, Kochanowski K, Vedelaar S, Ahrné E, Volkmer B, Callipo L, Knoop K, Bauer M, Aebersold R, Heinemann M. 2016. The quantitative and condition-dependent *Escherichia coli* proteome. *Nat Biotechnol* 34:104–110. <https://doi.org/10.1038/nbt.3418>.
 16. Soufi B, Krug K, Harst A, Macek B. 2015. Characterization of the *E. coli* proteome and its modifications during growth and ethanol stress. *Front Microbiol* 6:103. <https://doi.org/10.3389/fmicb.2015.00103>.
 17. Stokes NR, Murray HD, Subramaniam C, Gourse RL, Louis P, Bartlett W, Miller S, Booth IR. 2003. A role for mechanosensitive channels in survival of stationary phase: regulation of channel expression by RpoS. *Proc Natl Acad Sci U S A* 100:15959–15964. <https://doi.org/10.1073/pnas.2536607100>.
 18. Espah Borujeni A, Channarasappa AS, Salis HM. 2014. Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Res* 42:2646–2659. <https://doi.org/10.1093/nar/gkt1139>.
 19. Salis HM, Mirsky EA, Voigt CA. 2009. Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* 27:946–950. <https://doi.org/10.1038/nbt.1568>.
 20. Elowitz MB, Levine AJ, Siggia ED, Swain PS. 2002. Stochastic gene expression in a single cell. *Science* 297:1183–1186. <https://doi.org/10.1126/science.1070919>.
 21. Anderson RP, Jin R, Grunkemeier GL. 2003. Understanding logistic regression analysis in clinical reports: an introduction. *Ann Thorac Surg* 75:753–757. [https://doi.org/10.1016/S0003-4975\(02\)04683-0](https://doi.org/10.1016/S0003-4975(02)04683-0).
 22. Mishra V, Skotak M, Schuetz H, Heller A, Haorah J, Chandra N. 2016. Primary blast causes mild, moderate, severe and lethal TBI with increasing blast overpressures: experimental rat injury model. *Sci Rep* 6:26992. <https://doi.org/10.1038/srep26992>.
 23. Feeling-Taylor AR, Yau S-T, Petsev DN, Nagel RL, Hirsch RE, Vekilov PG. 2004. Crystallization mechanisms of hemoglobin C in the R state. *Biophys J* 87:2621–2629. <https://doi.org/10.1529/biophysj.104.039743>.
 24. Finch JT, Perutz MF, Bertles JF, Dobler J. 1973. Structure of sickled erythrocytes and of sickle-cell hemoglobin fibers. *Proc Natl Acad Sci U S A* 70:718–722.
 25. Perutz MF, Mitchison JM. 1950. State of hæmoglobin in sickle-cell anæmia. *Nature* 166:677–679. <https://doi.org/10.1038/166677a0>.
 26. Berg H, Purcell E. 1977. Physics of chemoreception. *Biophys J* 20:193–219. [https://doi.org/10.1016/S0006-3495\(77\)85544-6](https://doi.org/10.1016/S0006-3495(77)85544-6).
 27. Colin R, Sourjik V. 2017. Emergent properties of bacterial chemotaxis pathway. *Curr Opin Microbiol* 39:24–33. <https://doi.org/10.1016/j.mib.2017.07.004>.
 28. Krembel A, Colin R, Sourjik V. 2015. Importance of multiple methylation sites in *Escherichia coli* chemotaxis. *PLoS One* 10:e0145582. <https://doi.org/10.1371/journal.pone.0145582>.
 29. Krembel AK, Neumann S, Sourjik V. 2015. Universal response-adaptation relation in bacterial chemotaxis. *J Bacteriol* 197:307–313. <https://doi.org/10.1128/JB.02171-14>.
 30. Sourjik V, Berg HC. 2002. Receptor sensitivity in bacterial chemotaxis. *Proc Natl Acad Sci U S A* 99:123–127. <https://doi.org/10.1073/pnas.011589998>.
 31. Liu F, Morrison AH, Gregor T. 2013. Dynamic interpretation of maternal inputs by the *Drosophila* segmentation gene network. *Proc Natl Acad Sci U S A* 110:6724–6729. <https://doi.org/10.1073/pnas.1220912110>.
 32. Lovely GA, Brewster RC, Schatz DG, Baltimore D, Phillips R. 2015. Single-molecule analysis of RAG-mediated V(D)J DNA cleavage. *Proc Natl Acad Sci U S A* 112:E1715–E1723. <https://doi.org/10.1073/pnas.1503477112>.
 33. Schatz DG, Baltimore D. 2004. Uncovering the V(D)J recombinase. *Cell* 116:S103–S108. [https://doi.org/10.1016/S0092-8674\(04\)00042-X](https://doi.org/10.1016/S0092-8674(04)00042-X).
 34. Schatz DG, Ji Y. 2011. Recombination centres and the orchestration of V(D)J recombination. *Nat Rev Immunol* 11:251–263. <https://doi.org/10.1038/nri2941>.
 35. Herbig U, Jobling WA, Chen BP, Chen DJ, Sedivy JM. 2004. Telomere shortening triggers senescence of human cells through a pathway involving ATM, p53, and p21CIP1, but not p16INK4a. *Mol Cell* 14:501–513. [https://doi.org/10.1016/S1097-2765\(04\)00256-4](https://doi.org/10.1016/S1097-2765(04)00256-4).
 36. Victorelli S, Passos JF. 2017. Telomeres and cell senescence—size matters not. *EBioMedicine* 21:14–20. <https://doi.org/10.1016/j.ebiom.2017.03.027>.
 37. Levina N, Totemeyer S, Stokes NR, Louis P, Jones MA, Booth IR. 1999. Protection of *Escherichia coli* cells against extreme turgor by activation of MscS and MscL mechanosensitive channels: identification of genes required for MscS activity. *EMBO J* 18:1730–1737. <https://doi.org/10.1093/emboj/18.7.1730>.
 38. Booth IR. 2014. Bacterial mechanosensitive channels: progress towards an understanding of their roles in cell physiology. *Curr Opin Microbiol* 18:16–22. <https://doi.org/10.1016/j.mib.2014.01.005>.
 39. Li G-W, Burkhardt D, Gross C, Weissman JS. 2014. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* 157:624–635. <https://doi.org/10.1016/j.cell.2014.02.033>.
 40. Blount P, Sukharev SI, Moe PC, Martinac B, Kung C. 1999. Mechanosensitive channels of bacteria. *Methods Enzymol* 294:458–482. [https://doi.org/10.1016/S0076-6879\(99\)94027-2](https://doi.org/10.1016/S0076-6879(99)94027-2).
 41. Schumann U, Edwards MD, Rasmussen T, Bartlett W, van West P, Booth IR. 2010. YbdG in *Escherichia coli* is a threshold-setting mechanosensitive channel with MscM activity. *Proc Natl Acad Sci U S A* 107:12664–12669. <https://doi.org/10.1073/pnas.1001405107>.
 42. Sharan SK, Thomason LC, Kuznetsov SG, Court DL. 2009. Recombineering: a homologous recombination-based method of genetic engineering. *Nat Protoc* 4:206–223. <https://doi.org/10.1038/nprot.2008.227>.
 43. Li X-T, Thomason LC, Sawitzke JA, Costantino N, Court DL. 2013. Positive and negative selection using the tetA-sacB cassette: Recombineering and P1 transduction in *Escherichia coli*. *Nucleic Acids Res* 41:e204. <https://doi.org/10.1093/nar/gkt1075>.
 44. Bochner BR, Huang H-C, Schieven GL, Ames BN. 1980. Positive selection for loss of tetracycline resistance. *J Bacteriol* 143:926–933.
 45. Edelstein AD, Tsuchida MA, Amodaj N, Pinkard H, Vale RD, Stuurman N. 2014. Advanced methods of microscope control using μ Manager software. *J Biol Methods* 1:10. <https://doi.org/10.14440/jbm.2014.36>.
 46. Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P, Riddell A. 2017. Stan: a probabilistic programming language. *J Stat Softw* 76:JSSv076i01. <https://doi.org/10.18637/jss.v076.i01>.
 47. Chure G, Lee HJ, Phillips R. 16 May 2018. Microscopy images analyzed in “Connecting the dots between mechanosensitive channel abundance, osmotic shock, and survival at single-cell resolution.” CaltechDATA. <https://doi.org/10.22002/D1.941>.
 48. Chure G, Lee HJ, Phillips R. 16 May 2018. MCMC samples for channel quantification and logistic regression described in “Connecting the dots between mechanosensitive channel abundance, osmotic shock, and survival at single-cell resolution.” CaltechDATA. <https://doi.org/10.22002/D1.942>.
 49. Chure G, Lee HJ, Rasmussen A, Phillips R. 3 June 2018. RGroup-PBoC/mscL_survival: initial submission. Github repository for “Connecting the dots between mechanosensitive channel abundance, osmotic shock, and survival at single-cell resolution.” <https://doi.org/10.5281/zenodo.1252523>.

Supplemental information for “Connecting the dots between mechanosensitive channel abundance, osmotic shock, and survival at single-cell resolution”

Griffin Chure^{a,†}, Heun Jin Lee^{b,†}, Akiko Rasmussen^c, and Rob Phillips^{a,b,*}

^aDepartment of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA

^bDepartment of Applied Physics, California Institute of Technology, Pasadena, CA, USA

^cSchool of Medicine, Medical Sciences and Nutrition, Institute of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD United Kingdom

† contributed equally

* Send correspondence to phillips@pboc.caltech.edu

Supplement A: Experimental validation of MscL-sfGFP

Despite revolutionizing modern cell biology, tagging proteins with fluorophores can lead to myriad deleterious effects such as mislocalization, abrogation of functionality, or even cytotoxicity. In this section, we examine the stability and functionality of the MscL-sfGFP construct used in this work.

Comparing functionality of wild-type and fluorescently tagged MscL

To quantitatively compare the functionality between the wild-type MscL and MscL-sfGFP, patch-clamp electrophysiology experiments were conducted on each channel. Patch-clamp recordings were performed on membrane patches derived from giant protoplasts which were prepared as previously described (1). In brief, cells were grown in Luria-Bertani (LB) medium with 0.06 mg/ml cephalixin for 2.5 hours. The elongated cells were then collected by centrifugation and digested by 0.2 mg/ml lysozyme to form giant protoplasts.

Excised, inside-out patches were analyzed at a membrane potential of -20 mV with pipette and bath solutions containing 200 mM KCl, 90 mM MgCl₂, 10 mM CaCl₂, and 5 mM HEPES buffer at pH 7. All data were acquired at a sampling rate of 50 kHz with 5 kHz filtration using an AxoPatch 200B amplifier and pClamp software (Molecular Devices). The pressure threshold for activation a single MscS channel (blue stripe in Fig. S1) was compared to that of single MscL channels (yellow strip in Fig. S1). The pressure threshold for activation of the MscL channels was referenced against the

activation threshold of MscS to determine the pressure ratio (PL:PS) for gating as previously described (2). Recordings of the transmembrane current were made of three individual patches with an average PL:PS ratio of 1.56 for MscL-sfGFP. This ratio quantitatively agrees with the PL:PS ratio of 1.54 measured in a strain (MJF429 from the Booth laboratory) which expresses the wild-type MscL protein from the chromosome. The average transient current change from MscL openings (Fig. S1 shaded yellow region) is 75 pA, corresponding to a single channel conductance of 3.7 nS, comparable to the reported values of wild-type MscL. The agreement between these two strains indicates that there is negligible difference in functionality between MscL and MscL-sfGFP, allowing us to make physiological conclusions of the wild-type channel from our experiments.

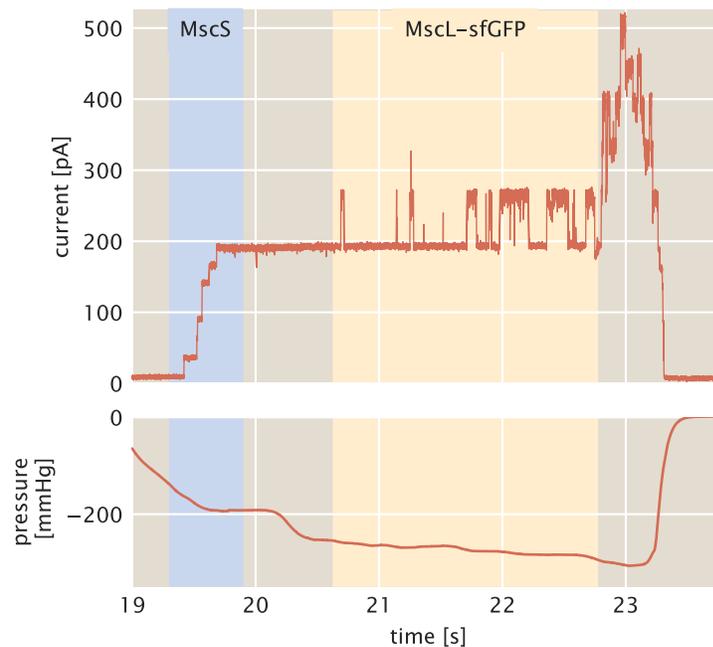


FIG S1 Characteristic MscL-sfGFP conductance obtained through patch-clamp electrophysiology. Top panel presents a characteristic measurement of channel current obtained through a patch-clamp electrophysiology measurement of bacterial protoplasts. The bottom panel shows the applied pressure through the micropipette to facilitate opening of the mechanosensitive channels. The blue shaded region indicates opening of the mechanosensitive channel of small conductance (MscS). The shaded yellow region represents opening of single MscL channels. These regions were used to compute the PL:PS ratio.

Maturation time of MscL-sfGFP

Reliable quantification of the channel copy number is paramount to this work. As such, it is important to verify that the detected fluorescence per cell accurately represents the total cellular MscL

copy number. We have made the assumption that the total fluorescence per represents all MscL-sfGFP channels present. However, it is possible that there are more channels present per cell but are not detected as the fluorophores have not properly matured. This potential error becomes more significant with longer maturation times of the fluorophore as the mean expression level changes with the growth phase of the culture. With a maturation time much longer than the typical cell division time, it is possible that the measured channel copy number represents only a fraction of the total number inherited over generations.

In our earlier work, we quantified the MscL-sfGFP channel copy number using fluorescence microscopy as well as with quantitative Western blotting. We found that these two methods agreed within 20% of the mean value, often with the counts resulting from microscopy being slightly larger than those measured through Western blotting (3). This strongly suggests that a negligible amount of channels are not observed due to inactive fluorophores.

Despite these suggestive data, we directly measured the maturation time of the superfolder GFP protein. We constructed a chromosomal integration of sfGFP expressed from a promoter under regulation from plasmid-borne TetR (*E. coli* MG1655 K12 $\Delta lacIZYA ybcN::sfGFP$). These cells were allowed to grow in LB supplemented with 500 mM NaCl held at 37°C to an OD_{600nm} of approximately 0.3. At this time, transcription and translation of the sfGFP gene was induced by addition of 10 ng/mL of anhydrous tetracycline. This expression was allowed to occur for three minutes before the addition of 100 µg/mL of kanamycin, ceasing proper protein synthesis. Three minutes of expression was chosen to provide enough time for transcription and translation. The sfGFP variant used in this work is 1155 base pairs. We can assume that the rate for transcription is 42 nucleotides per second (BNID 108488)(4), meaning approximately 28 seconds are needed to transcribe the gene. The translation rate is on the order of 10 amino acids per second, (12 - 42 amino acids / s, BNID 100059)(4). This means that 39 seconds are needed to complete translation. In total, approximately one minute is needed to complete expression of the genes. These numbers are not known for LB supplemented with 500 mM NaCl but may be reduced. For this reason, we extended the length of induction to three minutes before translation was ceased.

The excess anhydrous tetracycline was removed from the culture through centrifugation and washing with one volume of LB supplemented with 500 mM NaCl and 100 µg/mL kanamycin at 37°C. The maturation of sfGFP was then monitored through flow cytometry by measuring the mean expression of 100,000 cells every 60 to 90 seconds. The result of these measurements are shown in Fig. S2.

We observe complete maturation of the protein within 20 minutes after translation of the sfGFP gene was ceased. While the growth rate in LB + 500mM NaCl varies depending on the expression of MscL-sfGFP, we typically observe doubling times between 30 and 40 minutes, as indicated by a yellow stripe in Fig. S2A. To examine the “best case” scenario for cell growth in this medium, we measured the

growth rate of the same *E. coli* strain used to measure the fluorophore maturation time (Fig. S2 B). We observed a doubling time of 35 ± 1 min, which falls in the middle of the yellow stripe shown in Fig. S2 A. These data, coupled with our previous quantification of MscL copy number using independent methods, suggests that the fluorescence measurements made in this work reflect the total amount of MscL protein expressed.

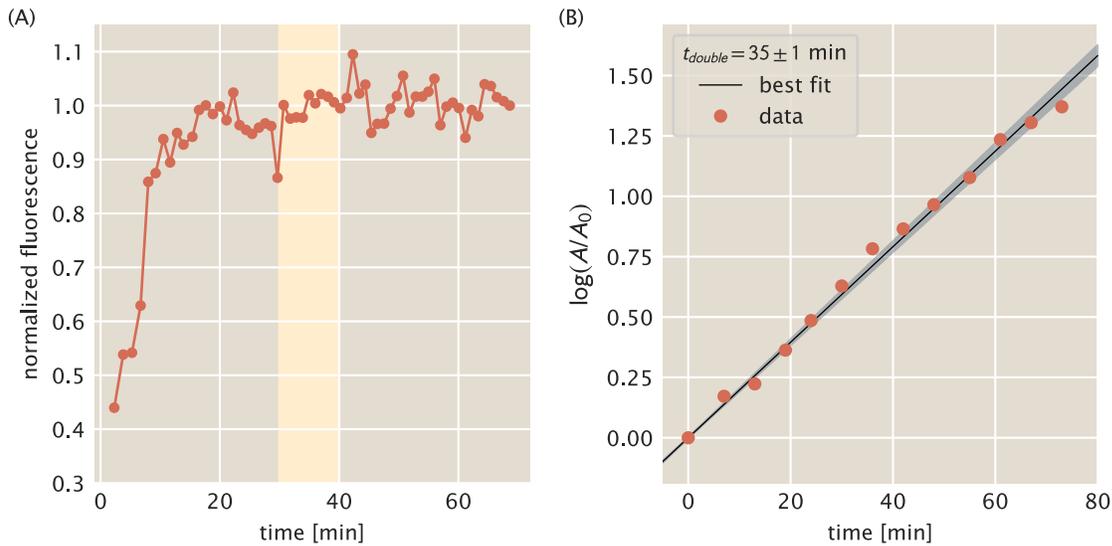


FIG S2 Measurement of sfGFP maturation as a function of time through flow cytometry. (A) Measurement of sfGFP fluorescence intensity as a function of time after cessation of protein translation. Points and connected lines indicate means of gated flow cytometry intensity distributions. Yellow stripe indicates the range of doubling times observed for the various RBS mutant strains described in this work (B) Growth curve of *E. coli* MG1655 cells in LB + 500mM NaCl. Red points indicate individual absorbance measurements. Line of best fit is shown in black with the uncertainty shown in shaded gray. The measured doubling time was 35 ± 1 min.

Quantifying Presence of MscL-sfGFP aggregates

van den berg et al. (2016) (5) showed that high levels of expression of MscL-mEos3.2 resulted in aggregates of channels, altering the physiology. To ensure that our method of measurement does not compromise our ability to draw physiological conclusions, it is important to quantify the extent of this phenomenon in our data as well as any bias it may impart on our analysis. We do indeed see fluorescent puncta in our data, yet it is possible these arise from simple statistical organization along the membrane. In van den Berg et al. 2016 (5), puncta were imaged using super-resolution microscopy, allowing tracking of their movement and calculation of a diffusion coefficient. Unfortunately, our

data is limited to single snapshots, prohibiting us from using diffusion as an identifying property of aggregation. We are therefore restricted to using statistical measures to characterize their abundance in the data.

To quantify the abundance of puncta, we analyzed a set of images from our highest expressing Shine-Dalgarno mutant (SD1) along with one of our lowest (SD4). Rather than just quantifying the mean pixel intensity of each cell, we calculated the coefficient of variation in intensity which serves as a measure of spatial uniformity. If the fluorescent proteins were very well behaved and showed no aggregation, one would naïvely expect the variation in intensity of the cell to be small with the intensity being relatively uniform across the entire segmented area. However, if there were aggregates, one would observe the formation of fluorescent puncta which would result in larger variation. In our reanalysis, we found that a very small proportion of cells of our highest expressing strain showed a large degree of variation in intensity (Fig. S3). Inspection of these images revealed that there were apparent fluorescent puncta which could be aggregates of the sfGFP tagged MscL proteins. These cells constitute approximately 10% of the total data set.

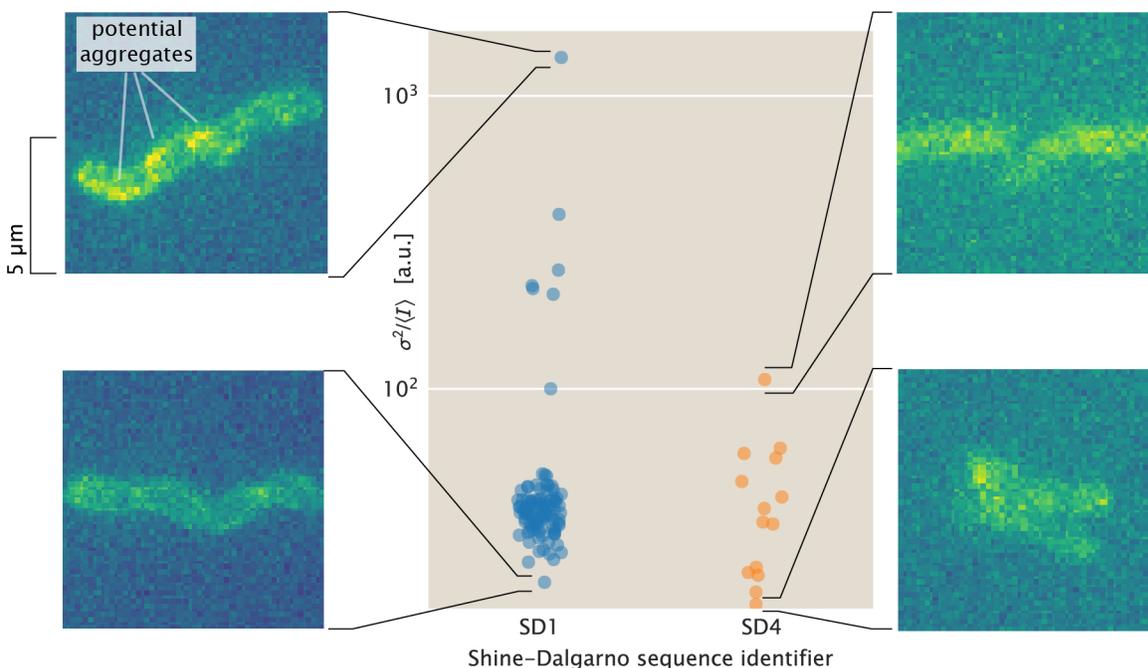


FIG S3 Distribution of potential channel aggregates under high and low expression. The noise of the measured pixel intensity is plotted with respect to the Shine-Dalgarno sequence modification. Example images are shown in false-color and are linked to their corresponding noise measurements with thin black lines. Color scale is relative to each image and cannot be compared between cells.

However, it is possible that the observed puncta are not aggregates but are rather the result of density fluctuations, where several channels happen to diffuse within a distance comparable to diffraction limited spot. We can test if this null hypothesis could explain our observation by making a simple stochastic model. Any channels within about 250 nm of each other would appear as a single fluorescent spot. We can make a simple estimate of the likelihood of observing a given number of MscL channels in a diffraction limited spot by coarse graining the cell membrane into 250 nm by 250 nm bins, as is shown in Fig. S4A. Suppose that we have a $4 \mu\text{m}^2$ sheet of cell membrane (an area similar to that observed in our experiments) split into 64 boxes each 250 nm on a side. Assuming that the closed MscL channel is approximately 10 nm in diameter, up to 625 pentameric channels can theoretically fit in one of these lattice sites. For our purposes it is fair to assume as null hypothesis that each lattice site has an equal probability of being occupied by an MscL channel. Using the mean expression value of our highest expressing strain (500 channels per cell), we can compute the probability distribution for number of MscL channels per lattice site, as is shown in Fig. S4B. We would expect to find seven MscL channels on average per site, which would all appear to be within the same diffraction limited spot. From our data, we find that on average there are 17 MscL channels per punctum, constituting approximately 3% of the total cellular channel copy number (Fig. S4C). The probability distribution shown in Fig. S4B predicts that approximately 5% of the lattice sites will have 15 or more MscL channels, which agrees with our experimental measurement of 3%. It is therefore unclear whether the observed puncta in the high-expressing cells are the result of aggregation of protein or merely a consequence of the statistics of partitioning.

Regardless, these cells rarely appear in our data, suggesting that any pathological consequences of punctate cells bear little weight in our conclusions regarding channel abundance dependent survival. The electrophysiology trace shown in Fig. S1 suggests that sfGFP tagged channels function identically to the wild-type untagged version in terms of conductance and gating tension (6). It has been previously shown that even wild-type MscL can form clusters in reconstituted membranes which can result in a hampered gating tension, although van den Berg et al. (5) propose their data does not suggest such cluster formation. It is therefore plausible that if the putative puncta observed in our data are aggregates, there may be none to little consequence when it comes to surviving an osmotic shock.

Supplement B: Standard Candle Calibration

To estimate the single-cell MscL abundance via microscopy, we needed to determine a calibration factor that could translate arbitrary fluorescence units to protein copy number. To compute this calibration factor, we relied on *a priori* knowledge of the mean copy number of MscL-sfGFP for a particular bacterial strain in specific growth conditions. In Bialecka-Fornal et al. 2012 (3), the average MscL copy number

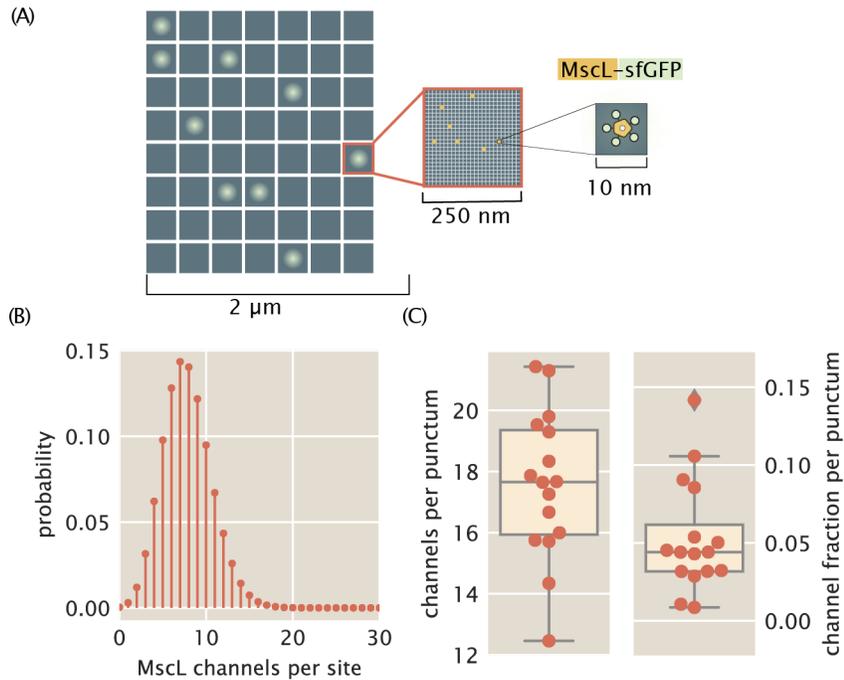


FIG S4 Formation of diffraction limited puncta from statistical positioning of MscL channels. (A) A $4\mu\text{m}^2$ sheet of membrane split into 64 boxes, each with a 250 nm edge length. All channels within one of these boxes would appear as a single fluorescent punctum. Each box can be split into 625 individual sites with a width of 10 nm (middle), each of which can accommodate a single MscL-sfGFP pentameric channel (right). (B) Probability distribution of number of MscL channels per 250 nm edge length box. Total cellular number of channels was taken as 500. (C) Observed distribution of channels per punctum (left) and fraction of channels found in each punctum (right). Individual measurements are shown in red. The box represents the interquartile region, centerline corresponds to the median, and whiskers extend to 1.5 times the maximum and minimum interquartile region.

for a population of cells expressing an MscL-sfGFP fusion (*E. coli* K-12 MG1655 $\phi(mscL\text{-}sfGFP)$) cells was measured using quantitative Western blotting and single-molecule photobleaching assays. By growing this strain in identical growth and imaging conditions, we can make an approximate measure of this calibration factor. In this section, we derive a statistical model for estimating the most-likely value of this calibration factor and its associated error.

Definition of a calibration factor

We assume that all detected fluorescence signal from a particular cell is derived from the MscL-sfGFP protein, after background subtraction and correction for autofluorescence. The arbitrary units of fluorescence can be directly related to the protein copy number via a calibration factor α ,

$$I_{\text{tot}} = \alpha N_{\text{tot}}, \quad (1)$$

where I_{tot} is the total cell fluorescence and N_{tot} is the total number of MscL proteins per cell. Bialecka-Fornal et al. report the average cell MscL copy number for the population rather than the distribution. Knowing only the mean, we can rewrite Eq. 1 as

$$\langle I_{\text{tot}} \rangle = \alpha \langle N_{\text{tot}} \rangle, \quad (2)$$

assuming that α is a constant value that does not change from cell to cell or fluorophore to fluorophore.

The experiments presented in this work were performed using non-synchronously growing cultures. As there is a uniform distribution of growth phases in the culture, the cell size distribution is broad the the extremes being small, newborn cells and large cells in the process of division. As described in the main text, the cell size distribution of a population is broadened further by modulating the MscL copy number with low copy numbers resulting in aberrant cell morphology. To speak in the terms of an effective channel copy number, we relate the average areal intensity of the population to the average cell size,

$$\langle I_{\text{tot}} \rangle = \langle I_A \rangle \langle A \rangle = \alpha \langle N_{\text{tot}} \rangle, \quad (3)$$

where $\langle I_A \rangle$ is the average areal intensity in arbitrary units per pixel of the population and $\langle A \rangle$ is the average area of a segmented cell. As only one focal plane was imaged in these experiments, we could not compute an appropriate volume for each cell given the highly aberrant morphology. We therefore opted to use the projected two-dimensional area of each cell as a proxy for cell size. Given this set of measurements, the calibration factor can be computed as

$$\alpha = \frac{\langle I_A \rangle \langle A \rangle}{\langle N_{\text{tot}} \rangle}. \quad (4)$$

While it is tempting to use Eq. 4 directly, there are multiple sources of error that are important to propagate through the final calculation. The most obvious error to include is the measurement error

reported in Bialecka-Fornal et al. 2012 for the average MscL channel count (3). There are also slight variations in expression across biological replicates that arise from a myriad of day-to-day differences. Rather than abstracting all sources of error away into a systematic error budget, we used an inferential model derived from Bayes' theorem that allows for the computation of the probability distribution of α .

Estimation of α for a single biological replicate

A single data set consists of several hundred single-cell measurements of intensity, area of the segmentation mask, and other morphological quantities. The areal density I_A is computed by dividing the total cell fluorescence by the cell area A . We are interested in computing the probability distributions for the calibration factor α , the average cell area $\langle A \rangle$, and the mean number of channels per cell $\langle N_{\text{tot}} \rangle$ for the data set as a whole given only I_A and A . Using Bayes' theorem, the probability distribution for these parameters given a single cell measurement, hereafter called the posterior distribution, can be written as

$$g(\alpha, \langle A \rangle, \langle N_{\text{tot}} \rangle | A, I_A) = \frac{f(A, I_A | \alpha, \langle A \rangle, \langle N_{\text{tot}} \rangle)g(\alpha, \langle A \rangle, \langle N_{\text{tot}} \rangle)}{f(\alpha, I_A)}, \quad (5)$$

where g and f represent probability density functions over parameters and data, respectively. The term $f(A, I_A | \alpha, \langle A \rangle, \langle N_{\text{tot}} \rangle)$ in the numerator represents the likelihood of observing the areal intensity I_A and area A of a cell for a given values of α , $\langle A \rangle$, and $\langle N_{\text{tot}} \rangle$. The second term in the numerator $g(\alpha, \langle A \rangle, \langle N_{\text{tot}} \rangle)$ captures all prior knowledge we have regarding the possible values of these parameters knowing nothing about the measured data. The denominator, $f(I_A, A)$ captures the probability of observing the data knowing nothing about the parameter values. This term, in our case, serves simply as a normalization constant and is neglected for the remainder of this section.

To determine the appropriate functional form for the likelihood and prior, we must make some assumptions regarding the biological processes that generate them. As there are many independent processes that regulate the timing of cell division and cell growth, such as DNA replication and peptidoglycan synthesis, it is reasonable to assume that for a given culture the distribution of cell size would be normally distributed with a mean of $\langle A \rangle$ and a variance $\sigma_{\langle A \rangle}$. Mathematically, we can write this as

$$f(A | \langle A \rangle, \sigma_{\langle A \rangle}) \propto \frac{1}{\sigma_{\langle A \rangle}} \exp \left[-\frac{(A - \langle A \rangle)^2}{2\sigma_{\langle A \rangle}^2} \right], \quad (6)$$

where the proportionality results from dropping normalization constants for notational simplicity.

While total cell intensity is intrinsically dependent on the cell area the areal intensity I_A is independent of cell size. The myriad processes leading to the detected fluorescence, such as translation and proper protein folding, are largely independent, allowing us to assume a normal distribution for I_A as well with a mean $\langle I_A \rangle$ and a variance $\sigma_{I_A}^2$. However, we do not have knowledge of the average areal intensity

for the standard candle strain *a priori*. This can be calculated knowing the calibration factor, total MsCL channel copy number, and the average cell area as

$$I_A = \frac{\alpha \langle N_{\text{tot}} \rangle}{\langle A \rangle}. \quad (7)$$

Using Eq. 7 to calculate the expected areal intensity for the population, we can write the likelihood as a Gaussian distribution,

$$f(I_A | \alpha, \langle A \rangle, \langle N_{\text{tot}} \rangle, \sigma_{I_A}) \propto \frac{1}{\sigma_{I_A}} \exp \left[-\frac{\left(I_A - \frac{\alpha \langle N_{\text{tot}} \rangle}{\langle A \rangle} \right)^2}{2\sigma_{I_A}^2} \right]. \quad (8)$$

With these two likelihoods in hand, we are tasked with determining the appropriate priors. As we have assumed normal distributions for the likelihoods of $\langle A \rangle$ and I_A , we have included two additional parameters, $\sigma_{\langle A \rangle}$ and σ_{I_A} , each requiring their own prior probability distribution. It is common practice to assume maximum ignorance for these variances and use a Jeffreys prior (7),

$$g(\sigma_{\langle A \rangle}, \sigma_{I_A}) = \frac{1}{\sigma_{\langle A \rangle} \sigma_{I_A}}. \quad (9)$$

The next obvious prior to consider is for the average channel copy number $\langle N_{\text{tot}} \rangle$, which comes from Bialecka-Fornal et al. 2012. In this work, they report a mean μ_N and variance σ_N^2 , allowing us to assume a normal distribution for the prior,

$$g(\langle N_{\text{tot}} \rangle | \mu_N, \sigma_N) \propto \frac{1}{\sigma_N} \exp \left[-\frac{(\langle N_{\text{tot}} \rangle - \mu_N)^2}{2\sigma_N^2} \right]. \quad (10)$$

For α and $\langle A \rangle$, we have some knowledge of what these parameters can and cannot be. For example, we know that neither of these parameters can be negative. As we have been careful to not overexpose the microscopy images, we can say that the maximum value of α would be the bit-depth of our camera. Similarly, it is impossible to segment a single cell with an area larger than our camera's field of view (although there are biological limitations to size below this extreme). To remain maximally uninformative, we can assume that the parameter values are uniformly distributed between these bounds, allowing us to state

$$g(\alpha) = \begin{cases} \frac{1}{\alpha_{\text{max}} - \alpha_{\text{min}}} & \alpha_{\text{min}} \leq \alpha \leq \alpha_{\text{max}} \\ 0 & \text{otherwise} \end{cases}, \quad (11)$$

for α and

$$g(\langle A \rangle) = \begin{cases} \frac{1}{\langle A \rangle_{\text{max}} - \langle A \rangle_{\text{min}}} & \langle A \rangle_{\text{min}} \leq \langle A \rangle \leq \langle A \rangle_{\text{max}} \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

for $\langle A \rangle$.

Piecing Eq. 6 through Eq. 12 together generates a complete posterior probability distribution for the parameters given a single cell measurement. This can be generalized to a set of k single cell measurements as

$$g(\alpha, \langle A \rangle, \langle N_{\text{tot}} \rangle, \sigma_{I_A}, \sigma_{\langle A \rangle} \mid [I_A, A], \mu_N, \sigma_N) \propto \frac{1}{(\alpha_{\text{max}} - \alpha_{\text{min}})(\langle A \rangle_{\text{max}} - \langle A \rangle_{\text{min}})} \frac{1}{(\sigma_{I_A} \sigma_{\langle A \rangle})^{k+1}} \times \frac{1}{\sigma_N} \exp \left[-\frac{(\langle N_{\text{tot}} \rangle - \mu_N)^2}{2\sigma_N^2} \right] \prod_i^k \exp \left[-\frac{(A^{(i)} - \langle A \rangle)^2}{2\sigma_{\langle A \rangle}^2} - \frac{\left(I_A^{(i)} - \frac{\alpha \langle N_{\text{tot}} \rangle}{\langle A \rangle} \right)^2}{2\sigma_{I_A}^2} \right], \quad (13)$$

where $[I_A, A]$ represents the set of k single-cell measurements.

As small variations in the day-to-day details of cell growth and sample preparation can alter the final channel count of the standard candle strain, it is imperative to perform more than a single biological replicate. However, properly propagating the error across replicates is non trivial. One option would be to pool together all measurements of n biological replicates and evaluate the posterior given in Eq. 13. However, this by definition assumes that there is no difference between replicates. Another option would be to perform this analysis on each biological replicate individually and then compute a mean and standard deviation of the resulting most-likely parameter estimates for α and $\langle A \rangle$. While this is a better approach than simply pooling all data together, it suffers a bias from giving each replicate equal weight, skewing the estimate of the most-likely parameter value if one replicate is markedly brighter or dimmer than the others. Given this type of data and a limited number of biological replicates ($n = 6$ in this work), we chose to extend the Bayesian analysis presented in this section to model the posterior probability distribution for α and $\langle A \rangle$ as a hierarchical process in which α and $\langle A \rangle$ for each replicate is drawn from the same distribution.

A hierarchical model for estimating α

In the previous section, we assumed maximally uninformative priors for the most-likely values of α and $\langle A \rangle$. While this is a fair approach to take, we are not completely ignorant with regard to how these values are distributed across biological replicates. A major assumption of our model is that the most-likely value of α and $\langle A \rangle$ for each biological replicate are comparable, so long as the experimental error between them is minimized. In other words, we are assuming that the most-likely value for each parameter for each replicate is drawn from the same distribution. While each replicate may have a unique value, they are all related to one another. Unfortunately, proper sampling of this distribution requires an extensive amount of experimental work, making inferential approaches more attractive.

This approach, often called a multi-level or hierarchical model, is schematized in Fig. S5. Here, we use an informative prior for α and $\langle A \rangle$ for each biological replicate. This informative prior probability distribution can be described by summary statistics, often called hyper-parameters, which are then

treated as the “true” value and are used to calculate the channel copy number. As this approach allows us to get a picture of the probability distribution of the hyper-parameters, we are able to report a point estimate for the most-likely value along with an error estimate that captures all known sources of variation.

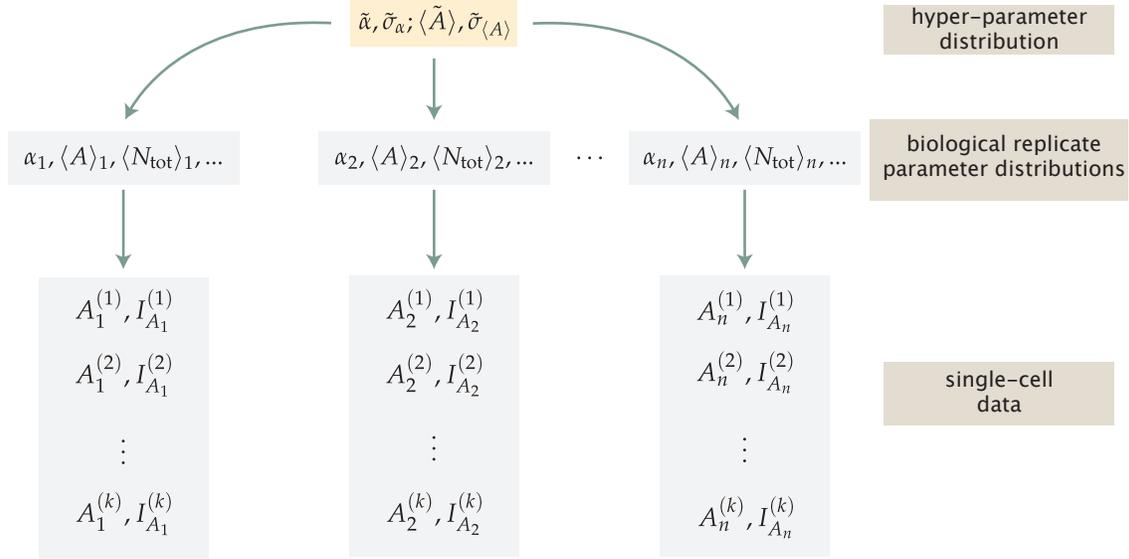


FIG S5 Schematic of hierarchical model structure. The hyper-parameter probability distributions (top panel) are used as an informative prior for the most-likely parameter values for each biological replicate (middle panel). The single-cell measurements of cell area and areal intensity (bottom panel) are used as data in the evaluation of the likelihood.

The choice for the functional form for the informative prior is often not obvious and can require other experimental approaches or back-of-the-envelope estimates to approximate. Each experiment in this work was carefully constructed to minimize the day-to-day variation. This involved adhering to well-controlled growth temperatures and media composition, harvesting of cells at comparable optical densities, and ensuring identical imaging parameters. As the experimental variation is minimized, we can use our knowledge of the underlying biological processes to guess at the approximate functional form. For similar reasons presented in the previous section, cell size is controlled by a myriad of independent processes. As each replicate is independent of another, it is reasonable to assume a normal distribution for the average cell area for each replicate. This normal distribution is described by a mean $\langle \tilde{A} \rangle$ and variance $\tilde{\sigma}_{\langle A \rangle}$. Therefore, the prior for $\langle A \rangle$ for n biological replicates can be written as

$$g(\langle A \rangle | \langle \tilde{A} \rangle, \tilde{\sigma}_{\langle A \rangle}) \propto \frac{1}{\tilde{\sigma}_{\langle A \rangle}^n} \prod_{j=1}^n \exp \left[-\frac{(\langle A \rangle_j - \langle \tilde{A} \rangle)^2}{2\tilde{\sigma}_{\langle A \rangle}^2} \right]. \quad (14)$$

In a similar manner, we can assume that the calibration factor for each replicate is normally distributed

with a mean $\tilde{\alpha}$ and variance $\tilde{\sigma}_\alpha$,

$$g(\alpha | \tilde{\alpha}, \tilde{\sigma}_\alpha) \propto \frac{1}{\tilde{\sigma}_\alpha^n} \prod_{j=1}^n \exp \left[-\frac{(\alpha_j - \tilde{\alpha})^2}{2\tilde{\sigma}_\alpha^2} \right]. \quad (15)$$

With the inclusion of two more normal distributions, we have introduced four new parameters, each of which needing their own prior. However, our knowledge of the reasonable values for the hyper-parameters has not changed from those described for a single replicate. We can therefore use the same maximally uninformative Jeffreys priors given in Eq. 9 for the variances and the uniform distributions given in Eq. 11 and Eq. 12 for the means. Stitching all of this work together generates the full posterior probability distribution for the best-estimate of $\tilde{\alpha}$ and $\langle \tilde{A} \rangle$ shown in Eq. 2 given n replicates of k single cell measurements,

$$g(\tilde{\alpha}, \tilde{\sigma}_\alpha, \langle \tilde{A} \rangle, \tilde{\sigma}_{\langle A \rangle}, \{ \langle N_{\text{tot}} \rangle, \langle A \rangle, \alpha, \sigma_{I_A} \} | [I_A, A], \mu_N, \sigma_N) \propto \frac{1}{(\tilde{\alpha}_{\text{max}} - \tilde{\alpha}_{\text{min}})(\langle \tilde{A} \rangle_{\text{max}} - \langle \tilde{A} \rangle_{\text{min}}) \sigma_N^n (\tilde{\sigma}_\alpha \tilde{\sigma}_{\langle A \rangle})^{n+1}} \times \prod_{j=1}^n \exp \left[-\frac{(\langle N \rangle_j^{(i)} - \mu_N)^2}{2\sigma_N^2} - \frac{(\alpha_j - \tilde{\alpha})^2}{2\tilde{\sigma}_\alpha^2} - \frac{(\langle A \rangle_j - \langle \tilde{A} \rangle)^2}{2\tilde{\sigma}_{\langle A \rangle}^2} \right] \times \frac{1}{(\sigma_{I_A} \sigma_{\langle A \rangle})^{k_j+1}} \prod_{i=1}^{k_j} \exp \left[-\frac{(A_j^{(i)} - \langle A \rangle_j)^2}{2\sigma_{\langle A \rangle}^{(i)2}} - \frac{(I_{A_j}^{(i)} - \frac{\alpha_j \langle N_{\text{tot}} \rangle_j}{\langle A \rangle_j})^2}{2\sigma_{I_{A_j}}^{(i)2}} \right], \quad (16)$$

where the braces $\{ \dots \}$ represent the set of parameters for biological replicates and the brackets $[\dots]$ correspond to the set of single-cell measurements for a given replicate.

While Eq. 16 is not analytically solvable, it can be easily sampled using Markov chain Monte Carlo (MCMC). The results of the MCMC sampling for $\tilde{\alpha}$ and $\langle \tilde{A} \rangle$ can be seen in Fig. S6. From this approach, we found the most-likely parameter values of 3300_{-700}^{+700} a.u. per MscL channel and $5.4_{-0.5}^{+0.4} \mu\text{m}^2$ for $\tilde{\alpha}$ and $\langle \tilde{A} \rangle$, respectively. Here, we've reported the median value of the posterior distribution for each parameter with the upper and lower bound of the 95% credible region as superscript and subscript, respectively. These values and associated errors were used in the calculation of channel copy number.

Effect of correction

The posterior distributions for α and $\langle A \rangle$ shown in Fig. S6 were used directly to compute the most-likely channel copy number for each measurement of the Shine-Dalgarno mutant strains, as is described in the coming section (*Logistic Regression*). The importance of this correction can be seen in Fig. S7. Cells with low abundance of MscL channels exhibit notable morphological defects, as illustrated in Fig. S7A. While these would all be considered single cells, the two-dimensional area of each may be comparable to two or three wild-type cells. For all of the Shine-Dalgarno mutants, the distribution of projected cell

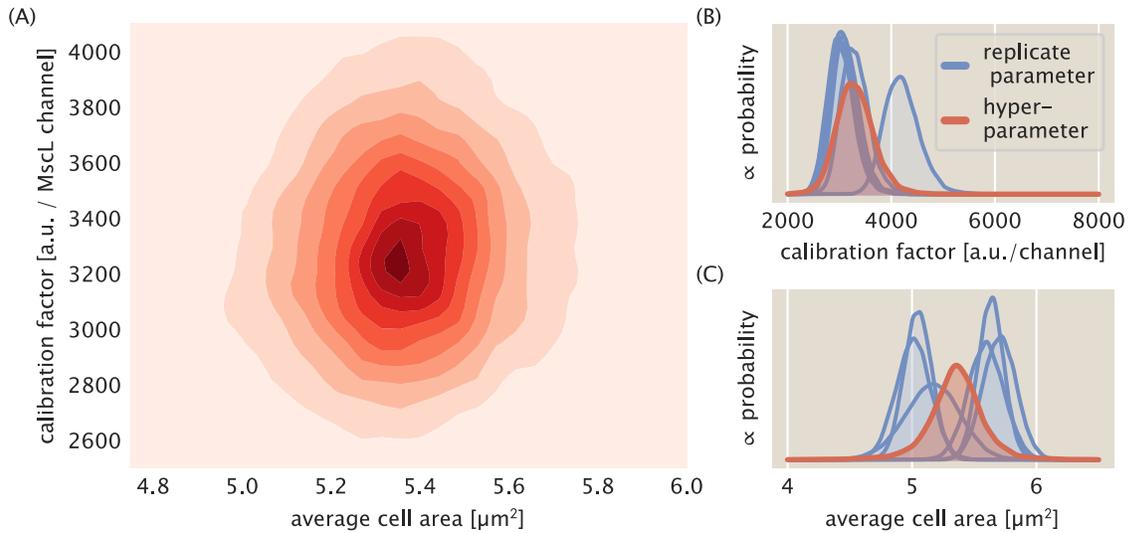


FIG S6 Posterior distributions for hyper-parameters and replicate parameters. (A) The posterior probability distribution for $\tilde{\alpha}$ and $\langle \tilde{A} \rangle$. Probability increases from light to dark red. The replicate parameter (blue) and hyper-parameter (red) marginalized posterior probability distributions for α (B) and $\langle A \rangle$ (C).

area has a long tail, with the extremes reaching $35 \mu\text{m}^2$ per cell (Fig. S7B). Calculating the total number of channels per cell does nothing to decouple this correlation between cell area and measured cell intensity. Fig. S7C shows the correlation between cell area and the total number of channels without normalizing to an average cell size $\langle A \rangle$ differentiated by their survival after an osmotic down-shock. This correlation is removed by calculating an effective channel copy number shown in Fig. S7D.

Choice of growth medium

To measure the calibration factor, we chose to grow the MLG910 *E. coli* strain in LB Miller medium as was performed in Bialecka-Fornal et al. 2012 (3), despite the fact that the Shine-Dalgarno sequence mutants used in this work were grown in LB supplemented with 500 mM NaCl. While a variety of growth media were tested in Bialecka-Fornal et al. 2012, the MscL-sfGFP copy number was not measured in this high-salt medium. To test if the calibration factor was dependent on the growth medium, we compared the estimated calibration factor in LB to that estimated in M9 supplemented with 0.5% glucose. These two media have starkly different compositions, serving as a good probe of the sensitivity of the calibration factor to intracellular chemistry.

E. coli strain MLG910 expressing an MscL-sfGFP fusion was grown in LB Miller, LB + 500 mM NaCl, and in M9 + 0.5% glucose to early exponential phase ($\text{OD}_{600\text{nm}} \sim 0.3$) and were imaged on a

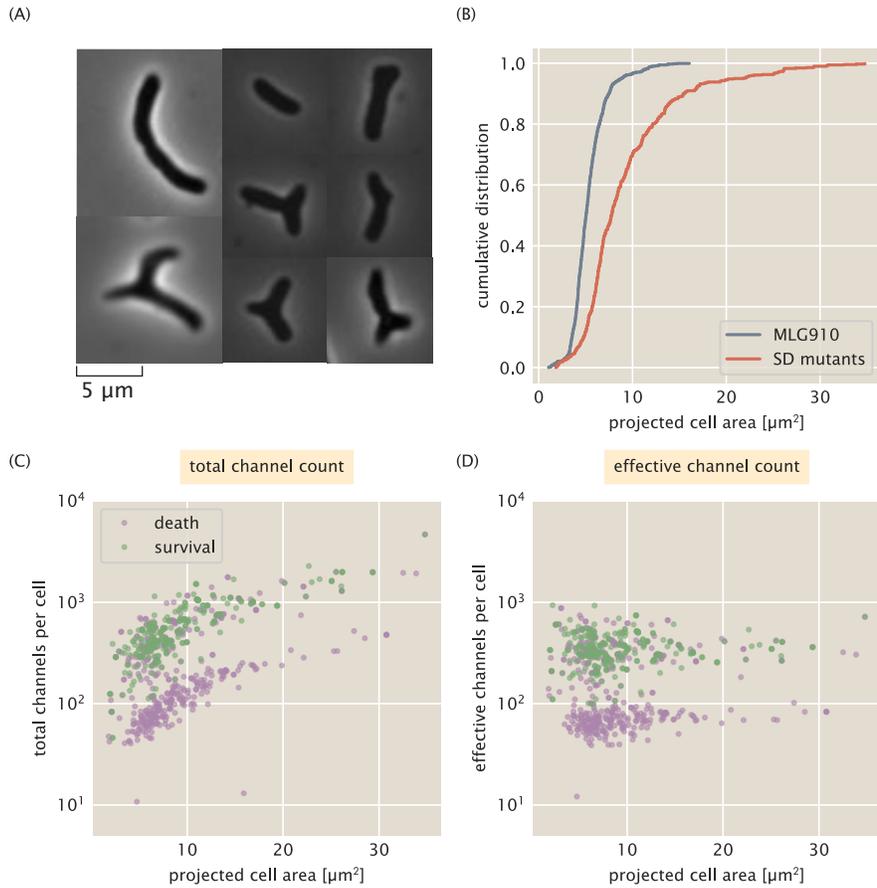


FIG S7 Influence of area correction for Shine-Dalgarno mutants. (A) Representative images of aberrant cell morphologies found in low-expressing Shine-Dalgarno mutants. (B) Empirical cumulative distribution of two-dimensional projected cell area for the standard candle strain MLG910 (gray line) and for all Shine-Dalgarno mutants (red line). (C) The correlation between channel copy number and cell area without the area correction. (D) The correlation between effective channel copy number and cell area with the area correction applied.

rigid agarose substrate. The cumulative intensity distributions for these growth media are shown in Fig. S8. There is a notable difference in the fluorescence distribution between LB Miller (blue line) and M9 + 0.5% glucose (red), as is expected given the estimated channel copy numbers reported in (3). The fluorescence distribution from MLG910 grown in LB + 500 mM NaCl (green line) is similar, yet not identical, to that grown in LB Miller broth. Using the count of 340 ± 68 and 466 ± 64 channels per cell for MLG910 grown in LB Miller and M9 + 0.5% glucose, respectively, coupled with the measurements shown in Fig. S8 we computed the posterior probability distribution of the calibration factor for each medium (Fig. S8 B). We found that these posterior distributions largely overlap, although the modes are distinctly different. We note that the calibration factor obtained from LB Miller completely overlaps that obtained from the M9 + 0.5% glucose data, suggesting that the error propagated through the final channel calculation accounts for any difference between the growth medium chosen.

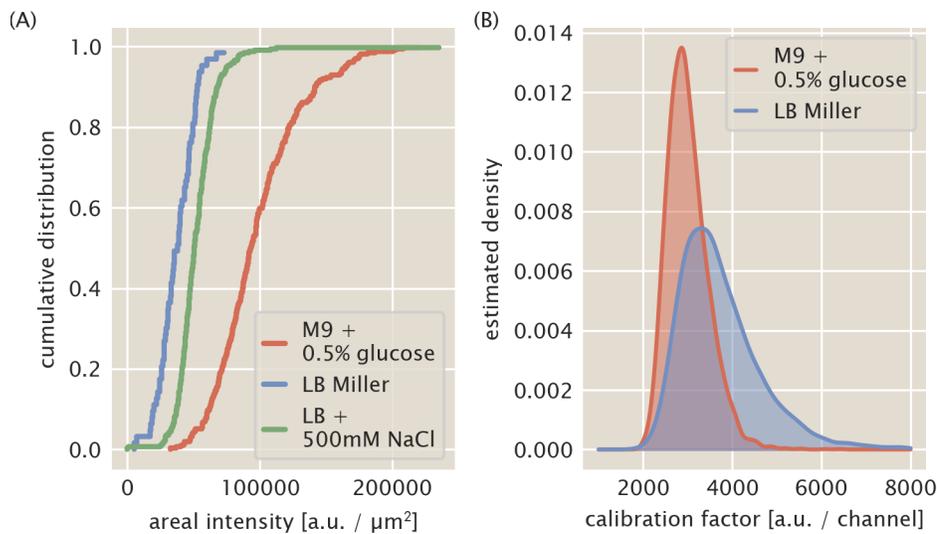


FIG S8 MLG910 intensity measurements and calibration factor estimation in different growth media. (A) Cumulative distribution of measured areal intensity (arbitrary units per square micron) measured for a collection of MLG910 cells grown in LB Miller medium (blue, medium used for calibration factor in this work), M9 + 0.5% glucose (red), and LB + 500mM NaCl (green, medium used for osmotic shock experiments). (B) Posterior probability distributions for the calibration factor (arbitrary units per MscL channel) estimated using MLG910 as a standard candle grown in LB miller (blue) and M9 + 0.5% glucose (red).

Supplement C: Classification of Cell Fate

We defined a survival event as a cell that went on to divide at least twice in the several hours following the applied osmotic shock. In nearly all of our experiments, cells which did not survive an osmotic shock exhibited necrosis with loss of phase contrast, extensive blebbing and bursting of the membrane, and the presence of dark aggregates at the cell poles. An example field across time is shown below in Fig. S9 where the cells are necrotic. We have also included a supplementary video of this field (Video S1). On occasion, we observed cells which did not obviously display the aforementioned death criteria yet did not undergo one or two division events. These cells were not counted in our experiments and were not included in the final tally of survival versus death. Across our 2822 single cell measurements, such “no call” classifications were observed only 83 times, constituting only 3% of the total cell measurements. A breakdown of all classification types and their respective abundances can be seen in Table S1.

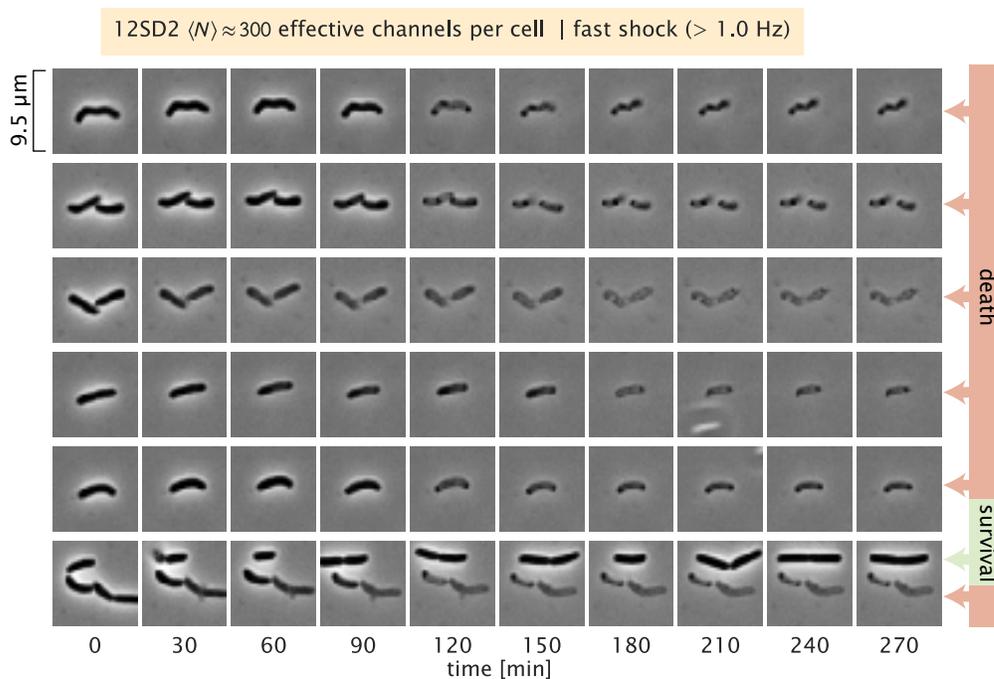


FIG S9 Time lapse of a representative field after osmotic shock and the resulting classifications. Each row shows an individual cell or pair of neighboring cells over time after the application of a fast osmotic shock. Cells classified as dead are denoted by red arrows. The lone surviving cell in this field (bottom row, top $\frac{1}{4}$ of image) is marked in green. A video file of a similar field can be found in the supplementary material as Video S1.

TABLE S1 Cell fate classifications and their relative abundances in the complete data set. “Dead-on-Arrival” classification was assigned to cells that appeared to be dead before application of osmotic shock. “No Call” was assigned to cells to whom fate could not be applied. “Death” was assigned to cells which displayed either obvious necrosis or did not undergo at least two division events. “Survival” was assigned to cells which underwent at least two division events after osmotic shock. Neither “No Call” nor “Dead-On-Arrival” classifications were included in the final analysis.

Classification	Number of Observations	Percentage of Measurements
Dead-On-Arrival	11	0.4%
No Call	83	3%
Death	1246	44%
Survival	1482	53%

To assess the validity of our morphology-based classification scheme, we performed a subset of the osmotic shock experiments described in the manuscript using propidium iodide staining to mark cells which had compromised membranes, identifying them as dead. Briefly, cells expressing on average ~ 80 MscL channels per cell were grown in LB + 500 mM NaCl to an OD_{600nm} of approximately 0.25. The cells were then mounted in the flow cell as described in the Materials and Methods in the main text and subjected to a large osmotic shock. After the shock, the cells were monitored for two hours. The propidium iodide stain (LIVE/DEAD BacLight Bacterial Cell Viability Staining, Thermo Fisher) was then passed into the flow chamber and imaged. An example of image of the phase contrast and propidium iodide fluorescence images are shown in Fig. S10. We note that cells matching our death criteria, meaning loss of phase contrast and visible distortion of the cell membrane, were strongly marked with propidium iodide, confirming that these cells were dead. The few example of “no call” classification where survival or death could not be determined from morphology alone showed that these cells were in fact dead (see highlighted row in Fig. S10). Cells that went on to divide two or more times in this period were not significantly stained with propidium iodide, confirming their viability and effectiveness of the stain itself. Given this data set, we compared the classification breakdown using our morphology-based method with the conclusive results from the propidium iodide staining (Table S2). We found that the two approaches to defining death agreed within 1%. This agreement leads us to believe that our definition of cell survival as morphological regularity and sustained cell growth is sufficiently accurate to draw physiological conclusions from our experiments.

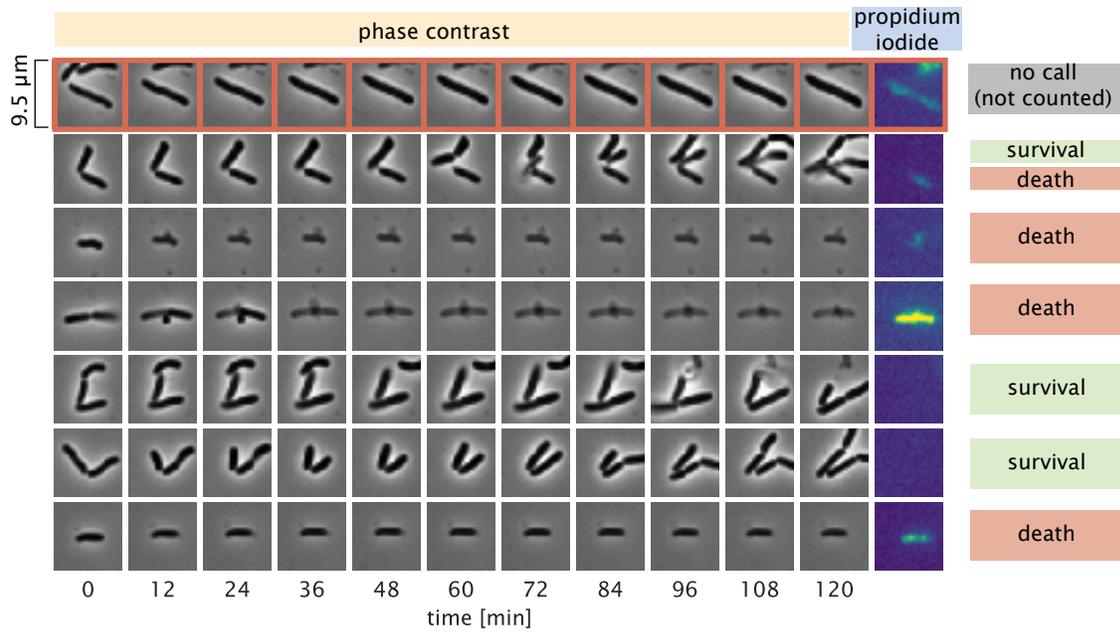


FIG S10 Representative images of propidium iodide staining after a strong osmotic shock. Phase contrast images of individual or pairs of cells as a function of time (columns). The final column corresponds to fluorescence from propidium iodide. Bright fluorescence indicates intercalation with DNA indicating cell death. Classification of survival based only from morphology is shown as text in the final column. Highlighted row indicates a “no call” event where morphology alone could not be used to determine survival or death.

TABLE S2 Comparison of morphology-based and dye-based survival classification. One cell was identified as a “No Call” in propidium iodide due to absence of propidium iodide staining despite extensive necrosis.

Classification	Observations via Morphology	Observations via Propidium Iodide Staining
Dead-On-Arrival	184	185
No Call	2	1
Survival	5	5

Supplement D: Logistic Regression

In this work, we were interested in computing the survival probability under a large hypo-osmotic shock as a function of MscL channel number. As the channel copy number distributions for each Shine-Dalgarno sequence mutant were broad and overlapping, we chose to calculate the survival probability through logistic regression – a method that requires no binning of the data providing the least biased estimate of survival probability. Logistic regression is a technique that has been used in medical statistics since the late 1950’s to describe diverse phenomena such as dose response curves, criminal recidivism, and survival probabilities for patients after treatment (8–10). It has also found much use in machine learning to tune a binary or categorical response given a continuous input (11–13).

In this section, we derive a statistical model for estimating the most-likely values for the coefficients β_0 and β_1 and use Bayes’ theorem to provide an interpretation for the statistical meaning.

Bayesian parameter estimation of β_0 and β_1

The central challenge of this work is to estimate the probability of survival p_s given only a measure of the total number of MscL channels in that cell. In other words, for a given measurement of N_c channels, we want to know likelihood that a cell would survive an osmotic shock. Using Bayes’ theorem, we can write a statistical model for the survival probability as

$$g(p_s | N_c) = \frac{f(N_c | p_s)g(p_s)}{f(N_c)}, \quad (17)$$

where g and f represent probability density functions over parameters and data, respectively. The posterior probability distribution $g(p_s | N_c)$ describes the probability of p_s given a specific number of channels N_c . This distribution is dependent on the likelihood of observing N_c channels assuming a value of p_s multiplied by all prior knowledge we have about knowing nothing about the data, $g(s)$.

The denominator $f(N_c)$ in Eq. 17 captures all knowledge we have about the available values of N_c , knowing nothing about the true survival probability. As this term acts as a normalization constant, we will neglect it in the following calculations for convenience.

To begin, we must come up with a statistical model that describes the experimental measurable in our experiment – survival or death. As this is a binary response, we can consider each measurement as a Bernoulli trial with a probability of success matching our probability of survival p_s ,

$$f(s | p_s) = p_s^s (1 - p_s)^{1-s}, \quad (18)$$

where s is the binary response of 1 or 0 for survival and death, respectively. As is stated in the introduction to this section, we decided to use a logistic function to describe the survival probability. We assume that the log-odds of survival is linear with respect to the effective channel copy number N_c as

$$\log \frac{p_s}{1 - p_s} = \beta_0 + \beta_1 N_c, \quad (19)$$

where β_0 and β_1 are coefficients which describe the survival probability in the absence of channels and the increase in log-odds of survival conveyed by a single channel. The rationale behind this interpretation is presented in the following section, *A Bayesian interpretation of β_0 and β_1* . Using this assumption, we can solve for the survival probability p_s as,

$$p_s = \frac{1}{1 + e^{-\beta_0 - \beta_1 N_c}}. \quad (20)$$

With a functional form for the survival probability, the likelihood stated in Eq. 17 can be restated as

$$f(N_c, s | \beta_0, \beta_1) = \left(\frac{1}{1 + e^{-\beta_0 - \beta_1 N_c}} \right)^s \left(1 - \frac{1}{1 + e^{-\beta_0 - \beta_1 N_c}} \right)^{1-s}. \quad (21)$$

As we have now introduced two parameters, β_0 , and β_1 , we must provide some description of our prior knowledge regarding their values. As is typically the case, we know nothing about the values for β_0 and β_1 . These parameters are allowed to take any value, so long as it is a real number. Since all values are allowable, we can assume a flat distribution where any value has an equally likely probability. This value of this constant probability is not necessary for our calculation and is ignored. For a set of k single-cell measurements, we can write the posterior probability distribution stated in Eq. 17 as

$$g(\beta_0, \beta_1 | N_c, s) = \prod_{i=1}^n \left(\frac{1}{1 + e^{-\beta_0 - \beta_1 N_c^{(i)}}} \right)^{s^{(i)}} \left(1 - \frac{1}{1 + e^{-\beta_0 - \beta_1 N_c^{(i)}}} \right)^{1-s^{(i)}} \quad (22)$$

Implicitly stated in Eq. 22 is absolute knowledge of the channel copy number N_c . However, as is described in *Standard Candle Calibration*, we must convert from a measured areal sfGFP intensity I_A to a effective channel copy number,

$$N_c = \frac{I_A \langle \tilde{A} \rangle}{\tilde{\alpha}}, \quad (23)$$

where $\langle \tilde{A} \rangle$ is the average cell area of the standard candle strain and $\tilde{\alpha}$ is the most-likely value for the calibration factor between arbitrary units and protein copy number. In *Standard Candle Calibration*, we detailed a process for generating an estimate for the most-likely value of $\langle \tilde{A} \rangle$ and $\tilde{\alpha}$. Given these estimates, we can include an informative prior for each value. From the Markov chain Monte Carlo samples shown in Fig. S6, the posterior distribution for each parameter is approximately Gaussian. By approximating them as Gaussian distributions, we can assign an informative prior for each as

$$g(\alpha | \tilde{\alpha}, \tilde{\sigma}_\alpha) \propto \frac{1}{\tilde{\sigma}_\alpha^k} \prod_{i=1}^k \exp \left[-\frac{(\alpha_i - \tilde{\alpha})^2}{2\tilde{\sigma}_\alpha^2} \right] \quad (24)$$

for the calibration factor for each cell and

$$g(\langle A \rangle | \langle \tilde{A} \rangle, \tilde{\sigma}_{\langle A \rangle}) = \frac{1}{\tilde{\sigma}_{\langle A \rangle}^k} \prod_{i=1}^k \exp \left[-\frac{(\langle A \rangle_i - \langle \tilde{A} \rangle)^2}{2\tilde{\sigma}_{\langle A \rangle}^2} \right], \quad (25)$$

where $\tilde{\sigma}_\alpha$ and $\tilde{\sigma}_{\langle A \rangle}$ represent the variance from approximating each posterior as a Gaussian. The proportionality for each prior arises from the neglecting of normalization constants for notational convenience.

Given Eq. 21 through Eq. 25, the complete posterior distribution for estimating the most likely values of β_0 and β_1 can be written as

$$g(\beta_0, \beta_1 | [I_A, s], \langle \tilde{A} \rangle, \tilde{\sigma}_{\langle A \rangle}, \tilde{\alpha}, \tilde{\sigma}_\alpha) \propto \frac{1}{(\tilde{\sigma}_\alpha \tilde{\sigma}_{\langle A \rangle})^k} \prod_{i=1}^k \left(1 + \exp \left[-\beta_0 - \beta_1 \frac{I_{A_i} \langle A \rangle_i}{\alpha_i} \right] \right)^{-s_i} \times \left(1 - \left(1 + \exp \left[-\beta_0 - \beta_1 \frac{I_{A_i} \langle A \rangle_i}{\alpha_i} \right] \right)^{-1} \right)^{1-s_i} \exp \left[-\frac{(\langle A \rangle_i - \langle \tilde{A} \rangle)^2}{2\tilde{\sigma}_{\langle A \rangle}^2} - \frac{(\alpha_i - \tilde{\alpha})^2}{2\tilde{\sigma}_\alpha^2} \right]. \quad (26)$$

As this posterior distribution is not solvable analytically, we used Markov chain Monte Carlo to draw samples out of this distribution, using the log of the effective channel number as described in the main text. The posterior distributions for β_0 and β_1 for both slow and fast shock rate data can be seen in Fig. S11

A Bayesian interpretation of β_0 and β_1

The assumption of a linear relationship between the log-odds of survival and the predictor variable N_c appears to be arbitrary and is presented without justification. However, this relationship is directly connected to the manner in which Bayes' theorem updates the posterior probability distribution upon the observation of new data. In following section, we will demonstrate this connection using the relationship between survival and channel copy number. However, this description is general and can be applied to any logistic regression model so long as the response variable is binary. This connection was shown briefly by Allen Downey in 2014 and has been expanded upon in this work (14).

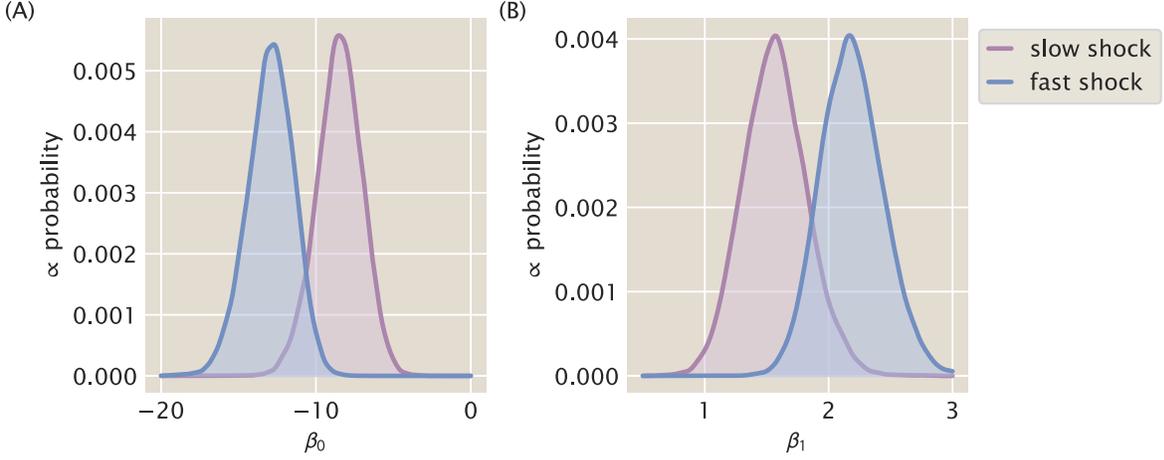


FIG S11 Posterior distributions for logistic regression coefficients evaluated for fast and slow shock rates. (A) Kernel density estimates of the posterior distribution for β_0 for fast (blue) and slow (purple) shock rates. (B) Kernel density estimates of posterior distribution for β_1 .

The probability of observing a survival event s given a measurement of N_c channels can be stated using Bayes' theorem as

$$g(s | N_c) = \frac{f(N_c | s)g(s)}{f(N_c)}. \quad (27)$$

where g and f represent probability density functions over parameters and data respectively. The posterior distribution $g(s | N_c)$ is the quantity of interest and implicitly related to the probability of survival. The likelihood $g(N_c | s)$ tells us the probability of observing N_c channels in this cell given that it survives. The quantity $g(s)$ captures all *a priori* knowledge we have regarding the probability of this cell surviving and the denominator $f(N_c)$ tells us the converse – the probability of observing N_c cells irrespective of the survival outcome.

Proper calculation of Eq. 27 requires that we have knowledge of $f(N_c)$, which is difficult to estimate. While we are able to give appropriate bounds on this term, such as a requirement of positivity and some knowledge of the maximum membrane packing density, it is not so obvious to determine the distribution between these bounds. Given this difficulty, it's easier to compute the odds of survival $\mathcal{O}(s | N_c)$, the probability of survival s relative to death d ,

$$\mathcal{O}(s | N_c) = \frac{g(s | N_c)}{g(d | N_c)} = \frac{f(N_c | s)g(s)}{f(N_c | d)g(d)}, \quad (28)$$

where $f(N_c)$ is cancelled. The only stipulation on the possible value of the odds is that it must be a positive value. As we would like to equally weigh odds less than one as those of several hundred or thousand, it is more convenient to compute the log-odds, given as

$$\log \mathcal{O}(s | N_c) = \log \frac{g(s)}{g(d)} + \log \frac{f(N_c | s)}{f(N_c | d)}. \quad (29)$$

Computing the log-transform reveals two interesting quantities. The first term is the ratio of the priors and tells us the *a priori* knowledge of the odds of survival irrespective of the number of channels. As we have no reason to think that survival is more likely than death, this ratio goes to unity. The second term is the log likelihood ratio and tells us how likely we are to observe a given channel copy number N_c given the cell survives relative to when it dies.

For each channel copy number, we can evaluate Eq. 29 to measure the log-odds of survival. If we start with zero channels per cell, we can write the log-odds of survival as

$$\log \mathcal{O}(s | N_c = 0) = \log \frac{g(s)}{g(d)} + \log \frac{f(N_c = 0 | s)}{f(N_c = 0 | d)}. \quad (30)$$

For a channel copy number of one, the odds of survival is

$$\log \mathcal{O}(s | N_c = 1) = \log \frac{g(s)}{g(d)} + \log \frac{f(N_c = 1 | s)}{f(N_c = 1 | d)}. \quad (31)$$

In both Eq. 30 and Eq. 31, the log of our *a priori* knowledge of survival versus death remains. The only factor that is changing is log likelihood ratio. We can be more general in our language and say that the log-odds of survival is increased by the difference in the log-odds conveyed by addition of a single channel. We can rewrite the log likelihood ratio in a more general form as

$$\log \frac{f(N_c | s)}{f(N_c | d)} = \log \frac{f(N_c = 0 | s)}{f(N_c = 0 | d)} + N_c \left[\log \frac{f(N_c = 1 | s)}{f(N_c = 1 | d)} - \log \frac{f(N_c = 0 | s)}{f(N_c = 0 | d)} \right], \quad (32)$$

where we are now only considering the case in which $N_c \in [0, 1]$. The bracketed term in Eq. 32 is the log of the odds of survival given a single channel relative to the odds of survival given no channels. Mathematically, this odds-ratio can be expressed as

$$\log \mathcal{OR}_{N_c}(s) = \log \frac{\frac{f(N_c=1|s)g(s)}{f(N_c=1|d)g(d)}}{\frac{f(N_c=0|s)g(s)}{f(N_c=0|d)g(d)}} = \log \frac{f(N_c = 1 | s)}{f(N_c = 1 | d)} - \log \frac{f(N_c = 0 | s)}{f(N_c = 0 | d)}. \quad (33)$$

Eq. 33 is mathematically equivalent to the bracketed term shown in Eq. 32.

We can now begin to staple these pieces together to arrive at an expression for the log odds of survival. Combining Eq. 32 with Eq. 29 yields

$$\log \mathcal{O}(s | N_c) = \log \frac{g(s)}{g(d)} + \log \frac{f(N_c = 0 | s)}{f(N_c = 0 | d)} + N_c \left[\log \frac{f(N_c = 1 | s)}{f(N_c = 1 | d)} - \log \frac{f(N_c = 0 | s)}{f(N_c = 0 | d)} \right]. \quad (34)$$

Using our knowledge that the bracketed term is the log odds-ratio and the first two terms represents the log-odds of survival with $N_c = 0$, we conclude with

$$\log \mathcal{O}(s | N_c) = \log \mathcal{O}(s | N_c = 0) + N_c \log \mathcal{OR}_{N_c}(s). \quad (35)$$

This result can be directly compared to Eq. 1 presented in the main text,

$$\log \frac{p_s}{1 - p_s} = \beta_0 + \beta_1 N_c, \quad (36)$$

which allows for an interpretation of the seemingly arbitrary coefficients β_0 and β_1 . The intercept term, β_0 , captures the log-odds of survival with no MscL channels. The slope, β_1 , describes the log odds-ratio of survival which a single channel relative to the odds of survival with no channels at all. While we have examined this considering only two possible channel copy numbers (1 and 0), the relationship between them is linear. We can therefore generalize this for any MscL copy number as the increase in the log-odds of survival is constant for addition of a single channel.

Other properties as predictor variables

The previous two sections discuss in detail the logic and practice behind the application of logistic regression to cell survival data using only the effective channel copy number as the predictor of survival. However, there are a variety of properties that could rightly be used as predictor variables, such as cell area and shock rate. As is stipulated in our standard candle calibration, there should be no correlation between survival and cell area. Fig. S12A and B show the logistic regression performed on the cell area. We see for both slow and fast shock groups, there is little change in survival probability with changing cell area and the wide credible regions allow for both positive and negative correlation between survival and area. The appearance of a bottle neck in the notably wide credible regions is a result of a large fraction of the measurements being tightly distributed about a mean value. Fig. S12C shows the predicted survival probability as a function of the the shock rate. There is a slight decrease in survivability as a function of increasing shock rate, however the width of the credible region allows for slightly positive or slightly negative correlation. While we have presented logistic regression in this section as a one-dimensional method, Eq. 19 can be generalized to n predictor variables x as

$$\log \frac{p_s}{1-p_s} = \beta_0 + \sum_i^n \beta_i x_i. \quad (37)$$

Using this generalization, we can use both shock rate and the effective channel copy number as predictor variables. The resulting two-dimensional surface of survival probability is shown in Fig. S12D. As is suggested by Fig. S12C, the magnitude of change in survivability as the shock rate is increased is smaller than that along increasing channel copy number, supporting our conclusion that for MscL alone, the copy number is the most important variable in determining survival.

Supplement E: Shock Classification

Its been previously shown that the rate of hypo-osmotic shock dictates the survival probability (15). To investigate how a single channel contributes to survival, we queried survival at several shock rates with varying MscL copy number. In the main text of this work, we separated our experiments into

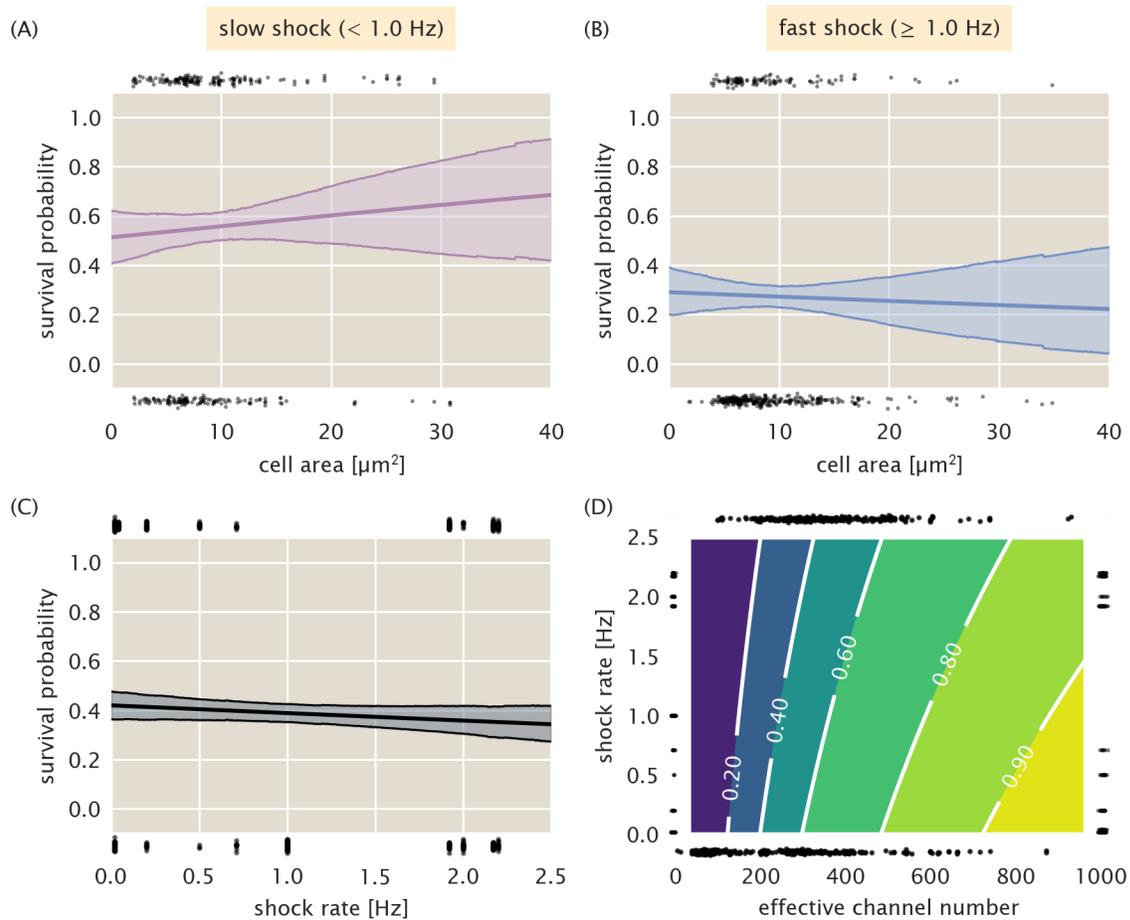


FIG S12 Survival probability estimation using alternative predictor variables. (A) Estimated survival probability as a function of cell area for the slow shock group. (B) Estimated survival probability as a function of cell area for the fast shock group. (C) Estimated survival probability as a function shock rate. Black points at top and bottom of plots represent single-cell measurements of cells who survived and perished, respectively. Shaded regions in (A) - (C) represent the 95% credible region. (D) Surface of estimated survival probability using both shock rate and effective channel number as predictor variables. Black points at left and right of plot represent single-cell measurements of cells which survived and died, respectively, sorted by shock rate. Points at top and bottom of plot represent survival and death sorted by their effective channel copy number. Labeled contours correspond to the survival probability.

arbitrary bins of “fast” (≥ 1.0 Hz) and “slow” (< 1.0 Hz) shock rates. In this section, we discuss our rationale for coarse graining our data into these two groupings.

As is discussed in the main text and in the supplemental section *Logistic Regression*, we used a bin-free method of estimating the survival probability given the MscL channel copy number as a predictor variable. While this method requires no binning of the data, it requires a data set that sufficiently covers the physiological range of channel copy number to accurately allow prediction of survivability. Fig. S13 shows the results of the logistic regression treating each shock rate as an individual data set. The most striking feature of the plots shown in Fig. S13 is the inconsistent behavior of the predicted survivability from shock rate to shock rate. The appearance of bottle necks in the credible regions for some shock rates (0.2Hz, 0.5Hz, 2.00Hz, and 2.20 Hz) appear due to a high density of measurements within a narrow range of the channel copy number at the narrowest point in the bottle neck. While this results in a seemingly accurate prediction of the survival probability at that point, the lack of data in other copy number regimes severely limits our extrapolation outside of the copy number range of that data set. Other shock rates (0.018 Hz, 0.04 Hz, and 1.00 Hz) demonstrate completely pathological survival probability curves due to either complete survival or complete death of the population.

Ideally, we would like to have a wide range of MscL channel copy numbers at each shock rate shown in Fig. S13. However, the low-throughput nature of these single-cell measurements prohibits completion of this within a reasonable time frame. It is also unlikely that thoroughly dissecting the shock rate dependence will change the overall finding from our work that several hundred MscL channels are needed to convey survival under hypo-osmotic stress.

Given the data shown in Fig. S13, we can try to combine the data sets into several bins. Fig. S14 shows the data presented in Fig. S13 separated into “slow” (< 0.5 Hz, A), “intermediate” (0.5 - 1.0 Hz, B), and “fast” (> 1.0 Hz, C) shock groups. Using these groupings, the full range of MscL channel copy numbers are covered for each case, with the intermediate shock rate sparsely sampling copy numbers greater than 200 channels per cell. In all three of these cases, the same qualitative story is told – several hundred channels per cell are necessary for an appreciable level of survival when subjected to an osmotic shock. This argument is strengthened when examining the predicted survival probability by considering all shock rates as a single group, shown in Fig. S14D. This treatment tells nearly the same quantitative and qualitative story as the three rate grouping shown in this section and the two rate grouping presented in the main text. While there does appear to be a dependence on the shock rate for survival when only MscL is expressed, the effect is relatively weak with overlapping credible regions for the logistic regression across the all curves. To account for the sparse sampling of high copy numbers observed in the intermediate shock group, we split this set and partitioned the measurements into either the “slow” (< 1.0 Hz) or “fast” (≥ 1.0 Hz) groups presented in the main text of this work.

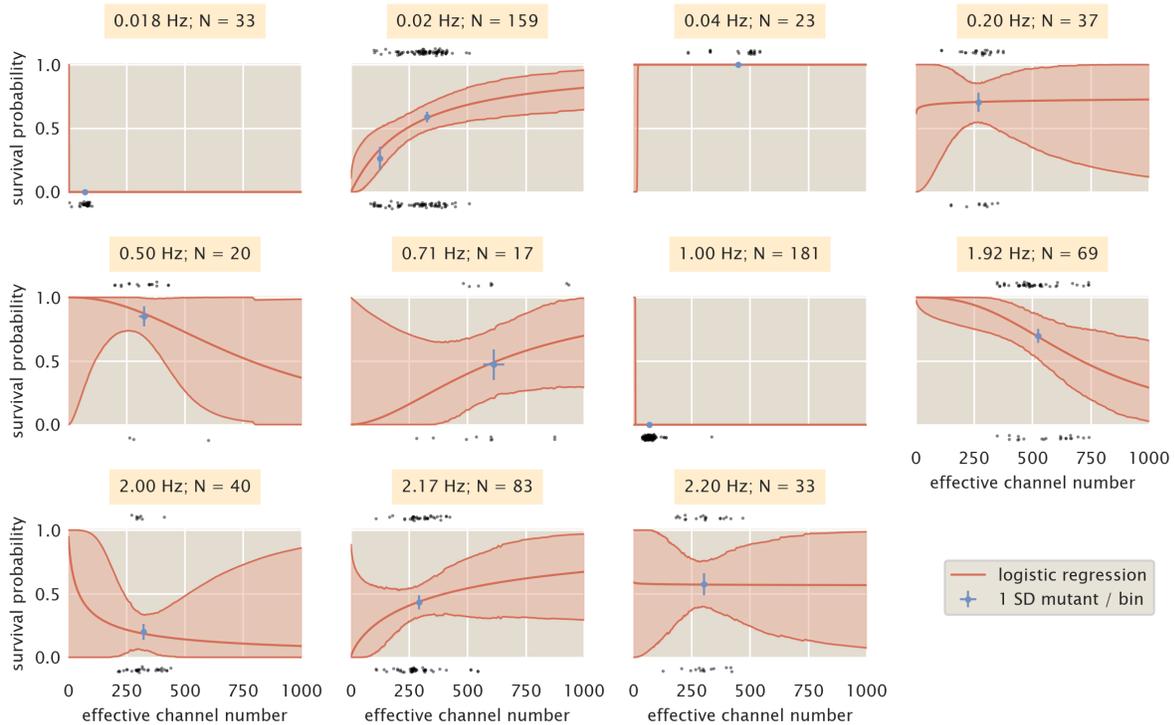


FIG S13 Binning by individual shock rates. Survival probability estimates from logistic regression (red lines) and the computed survival probability for all SD mutants subjected to that shock rate (blue points). Black points at top and bottom of each plot correspond to single cell measurements of survival (top) and death (bottom). Red shaded regions signify the 95% credible region of the logistic regression. Horizontal error bars of blue points are the standard error of the mean channel copy number. Vertical error bars of blue points correspond to the uncertainty in survival probability by observing n survival events from N single-cell measurements.

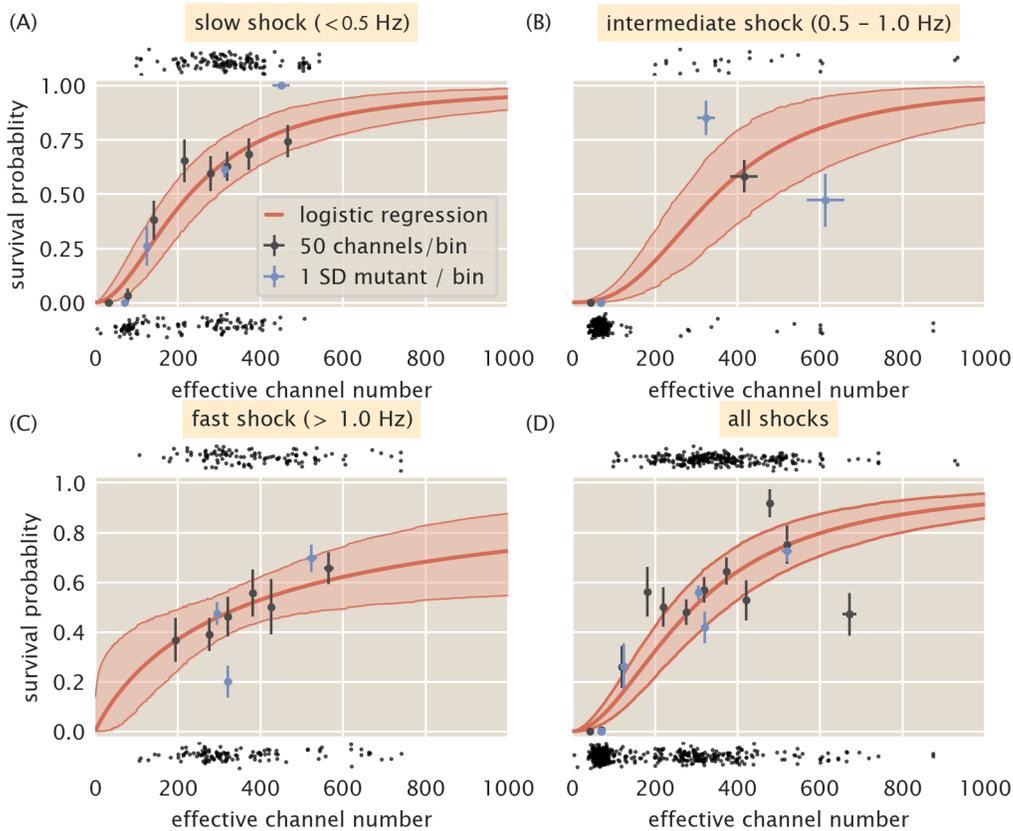


FIG S14 Coarse graining shock rates into different groups. Estimated survival probability curve for slow (A), intermediate (B), and fast (C) shock rates. (D) Estimated survival probability curve from pooling all data together, ignoring varying shock rates. Red shaded regions correspond to the 95% credible region of the survival probability estimated via logistic regression. Black points at top and bottom of each plot represent single-cell measurements of cells which survived and died, respectively. Black points and error bars represent survival probability calculations from bins of 50 channels per cell. Blue points represent the survival probability for a given Shine-Dalgarno mutant. Horizontal error bars are the standard error of the mean with at least 25 measurements and vertical error bars signifies the uncertainty in the survival probability from observing n survival events out of N total measurements.

Supplement F: Comparison of Survival Probability with van den Berg et al. 2016

In van den Berg et al. 2016, the authors report a 100% survival rate at approximately 100 channels per cell. While the number of mechanosensitive channels per cell was quantified at the level of single cells, the survival probability was measured in bulk using ensemble plating assays. The results of these experiments considering the contribution of MscL to survival is shown in Figure 5 of their work, although with out displayed uncertainty in the survival probability. Figure S6B of their work shows the approximate error in survival probability through ensemble plating assays for three different strains (Fig. S15A), which is approximately 30%. Using this approximate error and the data shown in their Figure 5B, we have reproduced this plot with error bars in both measured dimensions (Fig. S15B). This plot shows that even when the mean survival probability is 100%, the variation in the measured survival probability is large, extending as low as ~70%. This variation is likely born from a multitude of experimental steps including time of outgrowth, variation in shock rate, plating efficiency, and counting errors. As our experimental approach directly measures the survival/death of individual cells, we remove many sources of error that would arise from an ensemble approach, albeit at lower throughput. While it is possible that the discrepancy between van den Berg et al. 2016 and the work under review could arise from other unknown factors, we believe that single-cell experiments introduce the fewest sources of error.

Supplement G: Strains

Strain name	Genotype	Reference
MJF641	Frag1, $\Delta mscL::cm$, $\Delta mscS$, $\Delta mscK::kan$, $\Delta ybdG::apr$, $\Delta ynaI$, $\Delta yjeP$, $\Delta ybiO$, $ycjM::Tn10$	(16)
MLG910	MG1655, $\Delta mscL::\phi mscL-sfGFP$, $\Delta galK::kan$, $\Delta lacI$, $\Delta lacZY A$	(3)
D6LG-Tn10	Frag1, $\Delta mscL::\phi mscL-sfGFP$, $\Delta mscS$, $\Delta mscK::kan$, $\Delta ybdG::apr$, $\Delta ynaI$, $\Delta yjeP$, $\Delta ybiO$, $ycjM::Tn10$	This work
D6LG (SD0)	Frag1, $\Delta mscL::\phi mscL-sfGFP$, $\Delta mscS$, $\Delta mscK::kan$, $\Delta ybdG::apr$, $\Delta ynaI$, $\Delta yjeP$, $\Delta ybiO$	This work
XTL298	CC4231, $araD::tetA-sacB-amp$	(17)
D6LTetSac	Frag1, $mscL-sfGFP::tetA-sacB$, $\Delta mscS$, $\Delta mscK::kan$, $\Delta ybdG::apr$, $\Delta ynaI$, $\Delta yjeP$, $\Delta ybiO$	This work
D6LG (SD1)	Frag1, $\Delta mscL::\phi mscL-sfGFP$, $\Delta mscS$, $\Delta mscK::kan$, $\Delta ybdG::apr$, $\Delta ynaI$, $\Delta yjeP$, $\Delta ybiO$	This work

Strain name	Genotype	Reference
D6LG (SD2)	Frag1, $\Delta mscL :: \phi mscL\text{-sfGFP}$, $\Delta mscS$, $\Delta mscK::kan$, $\Delta ybdG::apr$, $\Delta ynaI$, $\Delta yjeP$, $\Delta ybiO$	This work
D6LG (SD4)	Frag1, $\Delta mscL :: \phi mscL\text{-sfGFP}$, $\Delta mscS$, $\Delta mscK::kan$, $\Delta ybdG::apr$, $\Delta ynaI$, $\Delta yjeP$, $\Delta ybiO$	This work
D6LG (SD6)	Frag1, $\Delta mscL :: \phi mscL\text{-sfGFP}$, $\Delta mscS$, $\Delta mscK::kan$, $\Delta ybdG::apr$, $\Delta ynaI$, $\Delta yjeP$, $\Delta ybiO$	This work
D6LG (12SD2)	Frag1, $\Delta mscL :: \phi mscL\text{-sfGFP}$, $\Delta mscS$, $\Delta mscK::kan$, $\Delta ybdG::apr$, $\Delta ynaI$, $\Delta yjeP$, $\Delta ybiO$	This work
D6LG (16SD0)	Frag1, $\Delta mscL :: \phi mscL\text{-sfGFP}$, $\Delta mscS$, $\Delta mscK::kan$, $\Delta ybdG::apr$, $\Delta ynaI$, $\Delta yjeP$, $\Delta ybiO$	This work

:*Escherichia coli* strains used in this work.

TABLE S4 Oligonucleotide sequences used in this work. Bold and italics correspond to Shine-Dalgarno sequence modifications and AT hairpin insertion modifications, respectively. Double bar || indicates a transposon insertion site.

Primer Name	Sequence (5' → 3')
<i>Tn10delR</i>	taaagccaacggcatccaggcggacataactcagca cctttcgcaaggtaacagagtaaaacatccaccat
<i>MscLSPSac</i>	gaaaatggcttaacatttgttagacttatggttgcg cttcat aggag TCCTAATTTTGTGACACTCTATC
<i>MscLSPSacR</i>	accacgttcccgcgcacatcgcaaattcgcgaaat tctttaataatgctcatATCAAAGGGAAAACGTCCATA
<i>MscL-SD1R</i>	atcgcaaattcgcgaaattctttaataatgctcat gttatt ctcctc atgaagccgacaaccataagcttaacaaa
<i>MscL-SD2R</i>	atcgcaaattcgcgaaattctttaataatgctcatgttatt tcccct atgaagccgacaaccataagcttaacaaa
<i>MscL-SD4R</i>	atcgcaaattcgcgaaattctttaataatgctcat gttatt cctgct atgaagccgacaaccataagcttaacaaa
<i>MscL-SD6R</i>	atcgcaaattcgcgaaattctttaataatgctcat gttatt gctcgt atgaagccgacaaccataagcttaacaaa
<i>MscL-12SD2R</i>	atcgcaaattcgcgaaattctttaataatgctcat atatatatatat tcccct atgaagccgacaaccataagcttaacaaa

Primer Name	Sequence (5' → 3')
<i>MscL-16SD0R</i>	atcgcaaattcgcgaaattctttaataaatgctcat atatatatatatatat ctccct atgaagccgacaaccataagtctaacaaa

References

1. Blount P, Sukharev SI, Moe PC, Martinac B, Kung C. 1999. Mechanosensitive channels of bacteria. *Methods in Enzymology* 294:458–482.
2. Blount P, Sukharev SI, Schroeder MJ, Nagle SK, Kung C. 1996. Single residue substitutions that change the gating properties of a mechanosensitive channel in *Escherichia coli*. *Proc Natl Acad Sci U S A* 93:11652–7.
3. Bialecka-Fornal M, Lee HJ, DeBerg HA, Gandhi CS, Phillips R. 2012. Single-Cell Census of Mechanosensitive Channels in Living Bacteria. *PLoS ONE* 7:e33077.
4. Milo R, Jorgensen P, Moran U, Weber G, Springer M. 2010. BioNumbers the database of key numbers in molecular and cell biology. *Nucleic Acids Research* 38:D750–D753.
5. van den Berg J, Galbiati H, Rasmussen A, Miller S, Poolman B. 2016. On the mobility, membrane location and functionality of mechanosensitive channels in *Escherichia coli*. *Scientific Reports* 6.
6. Grage SL, Keleshian AM, Turdeladze T, Battle AR, Tay WC, May RP, Holt SA, Contera SA, Haertlein M, Moulin M, Pal P, Rohde PR, Forsyth VT, Watts A, Huang KC, Ulrich AS, Martinac B. 2011. Bilayer-mediated clustering and functional interaction of MscL channels. *Biophys J* 100:1252–60.
7. Sivia D, Skilling J. 2006. *Data Analysis: A Bayesian Tutorial* Second Edition. Oxford University Press, Oxford, New York.
8. Anderson RP, Jin R, Grunkemeier GL. 2003. Understanding logistic regression analysis in clinical reports: An introduction. *The Annals of Thoracic Surgery* 75:753–757.
9. Mishra V, Skotak M, Schuetz H, Heller A, Haorah J, Chandra N. 2016. Primary blast causes mild, moderate, severe and lethal TBI with increasing blast overpressures: Experimental rat injury model. *Scientific Reports* 6:26992.
10. Stahler GJ, Mennis J, Belenko S, Welsh WN, Hiller ML, Zajac G. 2013. Predicting Recidivism For Released State Prison Offenders. *Criminal justice and behavior* 40:690–711.

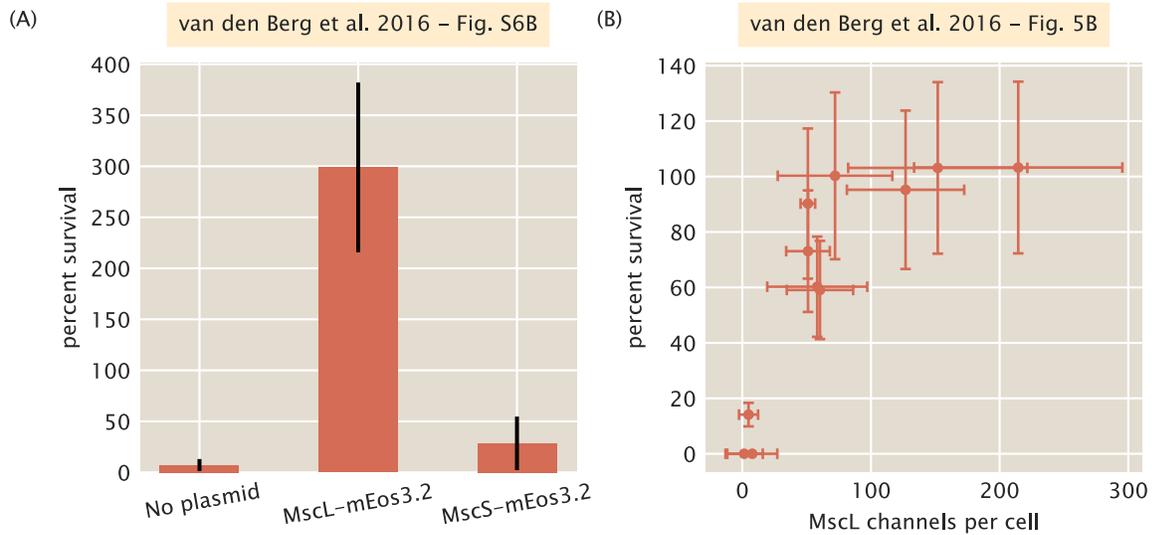


FIG S15 MscL abundance vs survival data reported in van den Berg et al. 2016 with included error. (A) Reported survival probabilities of a strain lacking all mechanosensitive channels (“no plasmid”), plasmid borne MscL-mEos3.2, and plasmid borne MscS-mEos3.2. Approximate reported errors for MscL-mEos3.2 survival probability is 30%. (B) The measurement of survival probability as a function of MscL channel copy number was obtained from Figure 5B in van den Berg et al 2016. Errors in channel copy number represent the standard deviation of several biological replicates (present in original figure) while the error in survival probability is taken as ~ 30%.

11. Cheng W, Hüllermeier E. 2009. Combining instance-based learning and logistic regression for multilabel classification. *Machine Learning* 76:211–225.
12. Dreiseitl S, Ohno-Machado L. 2002. Logistic regression and artificial neural network classification models: A methodology review. *Journal of Biomedical Informatics* 35:352–359.
13. Ng AY, Jordan MI. On Discriminative vs. Generative Classifiers: A comparison of logistic regression and naive Bayes 8.
14. Downey A. 2014. Probably Overthinking It: Bayes’s theorem and logistic regression. *Probably Overthinking It*.
15. Bialecka-Fornal M, Lee HJ, Phillips R. 2015. The Rate of Osmotic Downshock Determines the Survival Probability of Bacterial Mechanosensitive Channel Mutants. *Journal of Bacteriology* 197:231–237.
16. Edwards MD, Black S, Rasmussen T, Rasmussen A, Stokes NR, Stephen TL, Miller S, Booth IR. Characterization of three novel mechanosensitive channel activities in *Escherichia coli*. *Channels (Austin)* 6:272–81.
17. Li X-t, Thomason LC, Sawitzke JA, Costantino N, Court DL. 2013. Positive and negative selection using the tetA-sacB cassette: Recombineering and P1 transduction in *Escherichia coli*. *Nucleic acids research* 41:e204–e204.