nature biotechnology

Tunable protein degradation in bacteria

D Ewen Cameron[1–3] & James J Collins[1–4]

Tunable control of protein degradation in bacteria would provide a powerful research tool. Here we use components of the *Mesoplasma florum* transfer-messenger RNA system to create a synthetic degradation system that provides both independent control of steady-state protein level and inducible degradation of targeted proteins in *Escherichia coli*. We demonstrate application of this system in synthetic circuit development and control of core bacterial processes and antibacterial targets, and we transfer the system to *Lactococcus lactis* to establish its broad functionality in bacteria. We create a 238-member library of tagged essential proteins in *E. coli* that can serve as both a research tool to study essential gene function and an applied system for antibiotic discovery. Our synthetic protein degradation system is modular, does not require disruption of host systems and can be transferred to diverse bacteria with minimal modification.

Exogenous control of protein biosynthesis through transcriptional and translational regulation has been well established[1–4], but robust and tunable control of protein degradation in bacteria remains elusive. Controlled protein degradation would provide biologists with the ability to probe protein function without disrupting the transcriptional and translational regulation that control the expression of its cognate gene and would provide biological engineers with a tool to develop more complex synthetic gene circuits.

Protein degradation in bacteria occurs in part through the transfer-messenger RNA (tmRNA) system, which uses C-terminal fusion of the ssrA peptide to direct proteins to the endogenous ClpXP and ClpAP proteases for rapid degradation in *E. coli*[5]. Variants of the *E. coli* ssrA tag (*ec*-ssrA) are commonly used to modify the degradation rate of attached proteins in both bacteria and eukaryotes, but these tags do not provide inducible control of degradation. Recently developed inducible eukaryotic systems rely on degradation machinery not present in bacteria[6,7], and bacterial systems such as the one developed by Davis *et al.*[8] require disruption of the endogenous tmRNA system and are therefore not easily transferred to other organisms.

Here we present a synthetic degradation system based on the Gram-positive *M. florum* tmRNA system that does not rely on host degradation systems and can function in a wide range of bacteria. Gur and Sauer[9] showed that the *M. florum* ssrA tag (*mf*-ssrA) is degraded by its endogenous Lon protease (*mf*-Lon) but not by *E. coli* Lon or ClpXP, and *mf*-Lon does not recognize or degrade *ec*-ssrA, providing a protease and cognate degradation tag with orthogonal functionality in *E. coli*.

We renamed the *mf*-ssrA tag pdt (for protein degradation tag) to minimize confusion with the *E. coli* ssrA tag and incorporated it into a GFP-based test platform for inducible protein degradation in *E. coli* (**Fig. 1a**). To first engineer pdt variants that modify steady-state GFP levels in the absence of *mf*-Lon expression, we chose to target pdt residues 24–27 for mutagenesis because of the region's partial homology with the *ec*-ssrA ClpA binding site[5] and altered GFP-pdt stability in the *clpA*, *clpX* and *clpP* deletion strains (**Supplementary Figs. 1** and **2**). We identified several pdt variants (denoted with numbers and referred to as the pdt number variants) that altered GFP steady-state levels and maintained near–wild-type GFP degradation rates after *mf*-Lon expression (**Fig. 1b**). Notably, untagged GFP remained largely unaffected by *mf*-Lon expression, whereas the wild-type GFP-pdt fusion was reduced to 3% of its initial levels, confirming the specificity of pdt-mediated *mf*-Lon degradation seen by Gur and Sauer[9] for LacZ degradation. Sequence analysis of the identified pdt number variants showed that a majority contained multiple arginine or glutamine residues in the mutagenized region, and none of them contained negatively charged residues known to disrupt *mf*-Lon recognition[9] (**Supplementary Table 1**).

We used flow cytometry to further characterize *mf*-Lon–mediated GFP-pdt degradation and found that the pdt number variants displayed temporal degradation dynamics that were similar to those of wild-type pdt, reducing GFP levels to 1–5% of the initial levels within 4 h (**Fig. 1c**). GFP degradation did not occur in the absence of either *mf*-Lon or the pdt tag, and the tight monomodal shift in the fluorescent population distribution showed that degradation occurred across all cells in the experimental population (**Supplementary Fig. 3**).

We next sought to identify pdt variants (denoted with letters and referred to as the pdt letter variants) that alter *mf*-Lon–dependent degradation but not recognition by endogenous *E. coli* proteases. We used GFP-pdt#3 as the parental tag and targeted pdt residues 13–15 for mutagenesis because this region is essential for Lon-mediated degradation in *Mycoplasma pneumoniae*[10] and does not show homology to known ClpA, ClpX or SspB binding sites[5]. We identified pdt variants that maintained steady-state GFP levels and displayed a range of *mf*-Lon–dependent degradation rates (**Fig. 1d**).

To determine whether these letter variants could be combined with other number variants to produce hybrid tags with predictable control over both steady-state protein level and induced degradation

[1]Howard Hughes Medical Institute, Boston University, Boston, Massachusetts, USA. [2]Center of Synthetic Biology, Boston University, Boston, Massachusetts, USA. [3]Department of Biomedical Engineering, Boston University, Boston, Massachusetts, USA. [4]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, Massachusetts, USA. Correspondence should be addressed to J.J.C. (jcollins@bu.edu).

Received 1 August; accepted 23 September; published online 17 November 2014; doi:10.1038/nbt.3053

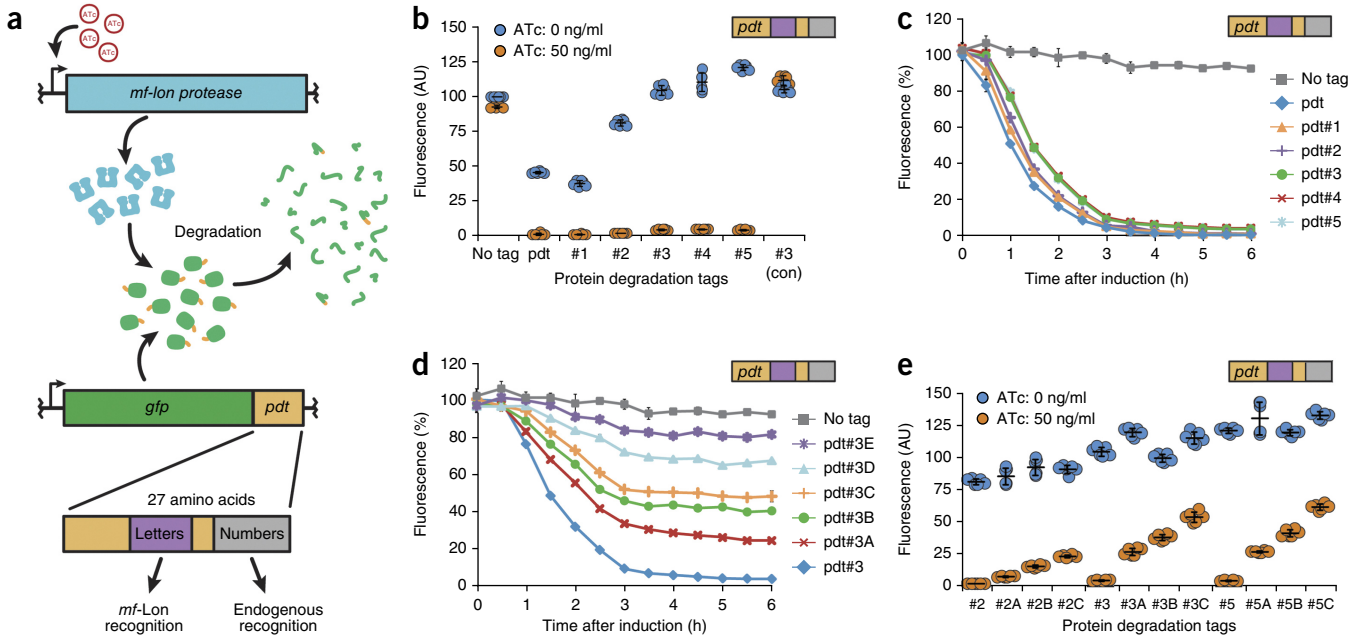

Figure 1 Protein degradation tag characterization. (**a**) Schematic of the tunable protein degradation system in which ATc-induced *mf*-Lon expression allows the protease to degrade constitutively expressed GFP in a pdt-dependent manner. Mutations in two pdt regions produced tag variants with altered recognition by *mf*-Lon (letters) or endogenous *E. coli* proteases (numbers). (**b**) Dot plot of the pdt number variants showing altered steady-state levels. Fluorescence was measured by flow cytometry 6 h after ATc induction of cells in exponential growth. As an experimental control, the pdt#3 variant was tested in a strain that did not contain the *mf*-Lon expression cassette (#3 con). Fluorescence units are arbitrary (AU), with untagged GFP set to 100, and show the mean of six biological replicates. $P < 0.001$ for ATc induction of each pdt variant, except #3 con, using a paired-sample, one-tailed *t*-test. (**c**,**d**) Flow cytometry measurements of GFP degradation after *mf*-Lon induction with 50 ng/ml ATc. Data show the geometric mean fluorescence of at least 5,000 cells as a percentage of the uninduced control for each pdt variant. (**c**) Pdt number variants maintain similar *mf*-Lon–mediated degradation dynamics. (**d**) Pdt letter variants display altered *mf*-Lon–mediated degradation rates. (**e**) Dot plot of hybrid pdt variants. Strains expressing the indicated GFP-pdt fusion were measured by flow cytometry 6 h after ATc induction. Fluorescence units are arbitrary, with untagged GFP set to 100, and show the mean of six biological replicates. $P < 0.001$ for ATc induction of each pdt variant using a paired-sample, one-tailed *t*-test. The data throughout the figure show the mean of at least three biological replicates, and all error bars show the s.d.

rate, we created a panel of hybrid pdt variants and measured GFP fluorescence in the presence and absence of *mf*-Lon induction. When combined with the number variants pdt#2 and pdt#5, the letter variants displayed the same rank order of degradation rates that we initially identified using pdt#3 (**Fig. 1e**). In the absence of *mf*-Lon induction, the hybrid pdt variants also showed steady-state levels that conformed largely to the level dictated by the number variant used, although there was marked variation in some hybrid tag combinations, suggesting partial recognition of the letter variant region by *E. coli* proteases.

To determine whether this GFP-pdt characterization can be used to predict pdt-mediated degradation of other protein targets, we placed pdt variants on the fluorescent protein mCherry and measured degradation after *mf*-Lon induction. The letter variants produced mCherry degradation dynamics that correlated strongly with GFP degradation, displaying linear regression with an R^2 value of 0.98 (**Fig. 2a**). The slope of the regression line (1.09) and its *y* intercept (−0.01) suggest that *mf*-Lon–mediated degradation of GFP and mCherry occurred at similar relative rates for all of the pdt letter variants tested. Pdt number variants also showed strong correlation for mCherry and GFP, with a linear regression R^2 value of 0.97 (**Supplementary Fig. 4a**).

To determine whether this targeted degradation system can function in other bacteria, we transferred the inducible protease and pdt variants to *L. lactis*, an industrially important Gram-positive bacterium that is phylogenetically distant from *E. coli* (a Gram-negative bacterium). *mf*-Lon expression from the inducible nisA promoter[11] in *L. lactis* resulted in efficient pdt-mediated degradation of mCherry (**Fig. 2b**), and the relative degradation strength of the pdt letter variants in *L. lactis* correlated well with their corresponding strength in *E. coli* ($R^2 = 0.91$) (**Fig. 2c**). As expected for pdt number variants that affect recognition by *E. coli*–specific proteases, steady-state levels of mCherry number variants showed only weak correlation between *L. lactis* and *E. coli* (**Supplementary Fig. 4b**).

Targeted protein degradation is dependent on not only the target protein and the pdt variant but also *mf*-Lon expression levels, providing an additional mechanism to control target protein levels. Transcriptional control of *mf*-Lon, based on anhydrotetracycline (ATc) induction of the P_{LtetO} promoter, provided a well-defined range of *mf*-Lon expression levels, as defined by targeted GFP-pdt#3 degradation (**Fig. 2d**). To enable post-translational control of *mf*-Lon, we fused variants of the *ec*-ssrA tag to *mf*-Lon and measured their effect on GFP-pdt#3 degradation. The *ec*-AAV variant caused a marked shift in *mf*-Lon–dependent GFP-pdt#3 degradation throughout the range of ATc levels tested, whereas the weaker *ec*-ASV variant had only a small effect. An inactivating mutation in the conserved active site of the *mf*-Lon proteolytic domain (S692A) fully blocked *mf*-Lon–mediated GFP-pdt#3 degradation.

To demonstrate the use of this system to control engineered genetic circuits, we used pdt fusions to provide post-translational control of a transcription-based toggle switch[12]. LacI and TetR were engineered previously to form a bistable circuit based on reciprocal repression, and concomitant regulation of GFP and mCherry allows facile fluorescence-based identification of the toggle switch state (**Fig. 3a**)[13]. We fused pdt#3 to the C terminus of LacI in the toggle circuit and used

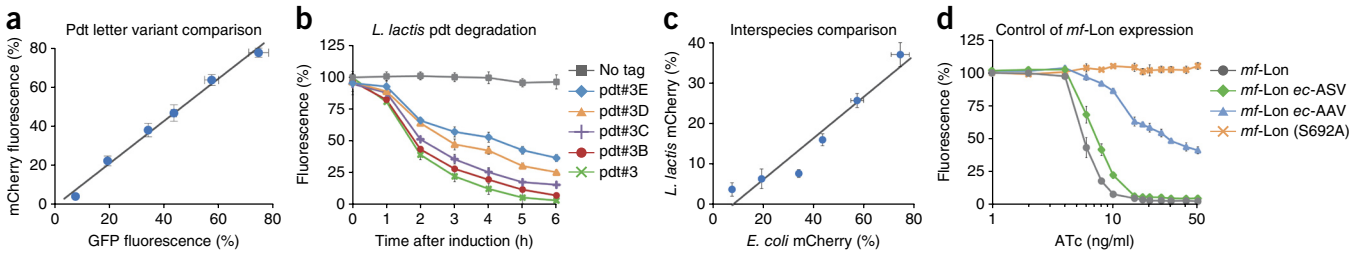

Figure 2 Pdt system characterization. (**a**) Comparative analysis of pdt-mediated degradation of mCherry and GFP. Pdt letter variants were fused to GFP and mCherry, and the percentage fluorescence remaining after *mf*-Lon induction is shown (50 ng/ml ATc for 6 h). Fluorescent data were collected by flow cytometry, and the pdt variants shown are pdt#3, pdt#3A, pdt#3B, pdt#3C, pdt#3D and pdt#3E, listed in order of increasing percentage fluorescence. The best-fit line by linear regression is $y = 1.09x - 0.01$, with an R^2 value of 0.98 and standard errors of 0.03 and 0.01 for the slope and *y* intercept, respectively. (**b**) Pdt-dependent degradation of mCherry in *L. lactis*. Nisin-induced *mf*-Lon expression in *L. lactis* causes pdt-dependent mCherry degradation. Data show the geometric mean fluorescence as a percentage of the fluorescence of uninduced cells. Nisin induction was at 3 ng/ml. (**c**) Comparative analysis of the pdt letter variants in *E. coli* and *L. lactis*. Pdt letter variants were fused to mCherry, and the percentage fluorescence remaining after *mf*-Lon induction is shown (6 h induction; *E. coli*: 50 ng/ml ATc; *L. lactis*: 3 ng/ml nisin). Fluorescent data were collected by flow cytometry, and the pdt variants shown are pdt#3, pdt#3A, pdt#3B, pdt#3C, pdt#3D and pdt#3E, listed in order of increasing percentage fluorescence. The best-fit line by linear regression is $y = 0.51x - 0.04$, with an R^2 value of 0.91 and standard errors of 0.03 and 0.02 for the slope and *y* intercept, respectively. (**d**) Transcription-based and post-translation–based control of *mf*-Lon–mediated pdt degradation. Inducible transcription provides control of *mf*-Lon–mediated degradation of GFP-pdt#3 across a range of ATc induction levels. Fusion of the *E. coli* ssrA tag variants *ec*-AAV and *ec*-ASV to *mf*-Lon shift the GFP degradation profile, and inactivation of *mf*-Lon protease activity (S692A) blocks GFP degradation. Data were collected 6 h after ATc induction using GFP-pdt#3 as the degradation target. The data throughout the figure show the mean of at least three biological replicates, and all error bars show the s.d.

the arabinose-inducible P_{BAD} promoter[1] to drive *mf*-Lon expression from a second plasmid. After *mf*-Lon induction, the circuit containing LacI-pdt#3 switched from the $LacI^+GFP^+$ state to the $TetR^+mCherry^+$ state within 8 h of *mf*-Lon induction, whereas the untagged circuit remained unchanged (**Fig. 3b**). Moreover, substitution of LacI-pdt#3 with the hybrid tags pdt#3A and pdt#3B provided temporal control over the circuit switch rate (**Fig. 3c**), and pdt fusions to TetR enabled *mf*-Lon to switch the toggle in the opposite direction (**Supplementary Fig. 5**). Notably, the LacI-pdt circuits maintained transcription-based bistability in the absence of *mf*-Lon induction, demonstrating the ability of the system to leave existing regulatory networks intact (**Supplementary Fig. 6**).

The relatively slow switch rate for this degradation-dependent toggle switch reflects the time needed for arabinose-induced *mf*-Lon expression and LacI-pdt#3 degradation, and the additional 3 hours it takes compared to ATc or isopropyl β-D-thiogalactoside (IPTG) induction[13] correlates well with the time needed for inducible degradation of GFP-pdt#3 by *mf*-Lon (**Fig. 1d**). As described below, induced *mf*-Lon degradation of endogenous proteins takes as little as 45 min, probably reflecting the relatively low abundance of most endogenous proteins compared to LacI and GFP, which are strongly expressed from multicopy plasmids, as is typical for synthetic gene circuits.

A major goal in microbial biotechnology is to develop tools to control and manipulate endogenous bacterial systems, so we next sought to target native *E. coli* pathways for control by our system. We developed a modified recombineering method to insert pdt tags into the *E. coli* genome (**Fig. 4a**) and began by targeting MurA, an essential enzyme involved in peptidoglycan biosynthesis[14] whose depletion causes cell lysis that is measurable by a drop in optical density. The *murA-pdt#1* genomic fusion caused observable cell lysis 45 min after *mf*-Lon induction (**Fig. 4b**), and the delayed phenotypic response of the hybrid variants pdt#1A and pdt#1B correlates well with the temporal delay seen for letter variants in the toggle switch and GFP degradation assays (**Figs. 1d** and **3c**). Notably, cells containing *murA-pdt* fusions displayed the same growth rate as wild-type cells in the absence of *mf*-Lon induction, demonstrating that the pdt variants do not interfere with MurA function or regulation (**Supplementary Fig. 7a**). Sequence analysis of cells with *murA-pdt#1* that escaped ATc-induced cell lysis showed mutations in *mf-lon* or its promoter that block its expression.

We next targeted FtsZ, a tubulin homolog that forms the ring structure necessary for cell septation after genome replication[15]. *mf*-Lon induction caused distinct filamentation in *ftsZ-pdt#5* cells but not wild-type cells within 3 h of ATc induction (**Fig. 4c**). The pdt#5 fusion had no discernible effect on FtsZ function under noninducing conditions (0 h images), and its growth rate was identical to that of wild-type cells (**Supplementary Fig. 7b**). We then targeted CheZ, a member of the chemotaxis signaling system whose disruption prevents directed flagellar motility[16]. In a disk diffusion assay on motility agar, *mf*-Lon induction caused bacteria containing *cheZ-pdt#5* to lose chemotactic motility (**Fig. 4d**), but bacteria that did not contain the *cheZ-pdt#5* fusion or did not express *mf*-Lon maintained normal chemotactic motility, confirming the specificity of ATc-induced *mf*-Lon degradation of CheZ-pdt#5.

We sought to determine whether this system could serve as a tool for antibiotic discovery and characterization. Large-scale chemical and natural product libraries can be screened efficiently for antimicrobial activity, but the next step of target identification is difficult and represents a major bottleneck in antibiotic development[17]. To identify compounds that target a specific protein for inhibition, one method is to reduce expression of the desired protein target to levels that induce a hypersensitive phenotype, in which antimicrobial compounds that inhibit the targeted protein show increased potency[18]. To demonstrate target-specific antibiotic hypersensitivity using our system, we returned our focus to MurA, the known target of fosfomycin[19]. We used the weak hybrid tag pdt#1D to inducibly degrade MurA and tested a range of ATc concentrations to identify *mf*-Lon induction conditions that produce a small *murA-pdt#1D*–dependent growth defect (4 ng/ml ATc; **Supplementary Fig. 8**). Under this low induction condition, cells that contained *murA-pdt#1D* displayed increased sensitivity to fosfomycin, whereas untagged or uninduced cells remained equally sensitive to the MurA-specific antibiotic (**Fig. 4e**).

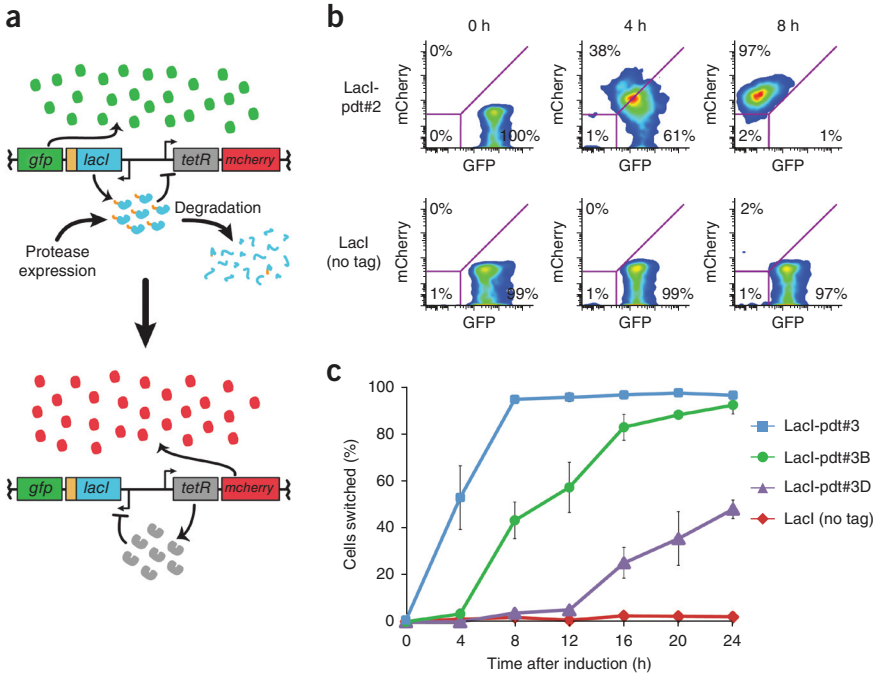

Figure 3 Protease-driven control of a synthetic toggle switch. (**a**) Schematic of the synthetic toggle switch in which reciprocal transcriptional repression by TetR and LacI form a bistable circuit. GFP and mCherry serve as fluorescent reporters for the LacI$^+$ and TetR$^+$ toggle states, respectively. Addition of a pdt tag to LacI enables a protease-driven switch from the GFP$^+$ to the mCherry$^+$ state. (**b**) Flow cytometry scatter plots showing GFP and mCherry fluorescence 0, 4 and 8 h after *mf*-Lon expression from the inducible promoter P_{BAD}. Degradation of LacI-pdt#3 causes the toggle to switch from the GFP$^+$ state to the mCherry$^+$ state by 8 h, whereas the untagged toggle remains in the GFP$^+$ state. Purple lines indicate the gate parameters used to define the GFP$^+$ and mCherry$^+$ states: cells bounded in the lower left quadrant are considered negative for both GFP and mCherry. (**c**) The percentage of cells in the mCherry$^+$ state after *mf*-Lon induction with 1 mM arabinose. Data collected by flow cytometry were measured using the parameters in **b** and represent the mean of three biological replicates. For all LacI-pdt variants, $P < 0.001$ when compared to untagged LacI at 24 h after induction using a paired-sample, one-tailed *t*-test. All error bars show the s.d. Data showing that uninduced strains did not shift to mCherry$^+$ are shown in **Supplementary Figure 6**.

The slow pace of antibiotic discovery in recent years has led to alternative approaches to drug development, including an increased focus on chemical adjuvants that boost the potency of antibiotics currently in clinical use. Inducible control of adjuvant targets with our system could provide a method to both identify new cellular targets for adjuvant development and screen for chemical adjuvants that inhibit a desired protein. As a proof-of-principle demonstration, we used our platform to control endogenous levels of RecA, a DNA damage repair protein that helps protect cells against norfloxacin and is being actively pursued as an antibiotic adjuvant target[20]. After *mf*-Lon induction, cells containing a *recA-pdt#3* fusion became hypersensitive to norfloxacin, and the induced phenotype was nearly identical to that of a *recA* deletion strain (**Fig. 4f**). As described above for MurA-dependent fosfomycin sensitivity, intermediate-level RecA degradation could be used to screen for molecules that inhibit RecA.

Current classes of antibiotics target only a small subset of essential cellular proteins for inhibition, and emerging resistance to these antibiotics has led to an urgent search for additional targets for antibiotic development. To aid in this search, we targeted every essential gene in *E. coli* for pdt fusion to create a strain library that could be screened for antibiotic hypersensitivity and degradation-dependent cell death. Using high-throughput electroporation, we recovered pdt#1 insertions in 238 of the 305 essential genes targeted, and we arrayed the strains in multiwell format to form the Essential Protein Degradation (EPD) library (**Supplementary Table 2**). Despite repeated attempts, we could not recover pdt insertions in the remaining 67 genes, including 28 ribosomal genes, possibly because of pdt interference with protein function, altered target protein stability caused by the pdt#1 variant or a polar effect of the insertion cassette on downstream genes.

To establish the functionality of the EPD library and search for potential antibiotic targets, we screened for growth inhibition and cell death after ATc-induced *mf*-Lon expression. 112 strains (46%) showed a reduction in colony forming units (CFUs) to less than 10% of uninduced controls, including 54 that showed at least a 6-log drop, and this reduction correlated well with their observed growth defect as measured by optical density (**Fig. 4g** and **Supplementary Table 2**). As has been similarly proposed by Wei *et al.*[21] in *Mycobacterium smegmatis*, these degradation-sensitive strains represent a list of potential drug targets, and as described above for MurA, they also provide a means to screen for target-specific antibiotics. For EPD library members with no inducible growth defect, possible explanations include pdt inaccessibility due to protein folding, multiprotein complex formation, export from the cytoplasm or post-translational C-terminal processing. High expression or proteolytic stability of the target protein may also prevent *mf*-Lon from degrading it to levels that are low enough to affect cell viability.

Antibiotics that disrupt bacterial membrane integrity are particularly desirable, as they not only kill cells but may also enhance the potency of other antibiotics that are excluded from their cytoplasmic targets by intrinsic or acquired resistance mechanisms. To identify protein targets that affect membrane integrity, we screened the EPD library using propidium iodide (PI), a hydrophilic nucleic acid intercalating agent that cannot normally permeate the inner membrane of *E. coli*. We assayed strains by high-throughput flow cytometry 2 hours after *mf*-Lon induction and quantified the percentage of cells that showed cytoplasmic PI staining. Targeted degradation of phospholipid and peptidoglycan biogenesis enzymes such as PlsB and MurC caused a distinct membrane permeability phenotype, and in particular, isoprenoid biosynthesis enzymes, including Dxr, Dxs, IspD, IspG, IspH and IspU, showed high membrane permeability after ATc induction (**Fig. 4h** and **Supplementary Table 2**). These data correlate well with the antimicrobial activity of the Dxr inhibitor fosmidomycin and several IspU inhibitors[22] and suggest that additional isoprenoid biosynthesis enzymes may serve as effective antibiotic targets, especially because the pathway is not found in humans. Other interesting PI$^+$ strains include the carbonic anhydrase Can and the NAD kinase NadK (**Fig. 4h**), suggesting that they have an essential role in maintaining membrane integrity and could also be effective antibiotic targets.

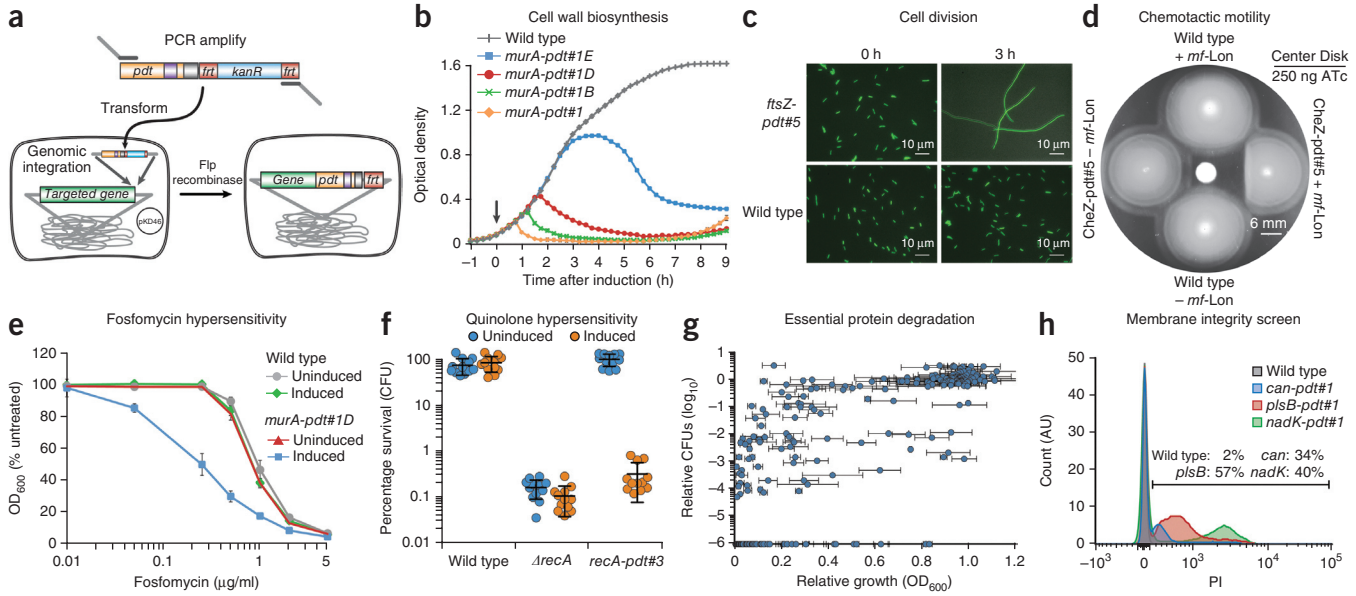

Figure 4 Tunable control of endogenous bacterial systems and antibacterial targets. (**a**) Schematic of our recombineering method for genomic insertion of pdt variants, adapted from Datsenko and Wanner[30]. Red recombinase–assisted insertion of the desired pdt variant is followed by Flp recombinase–mediated excision of the accompanying *kanR* cassette. The resulting insertion contains the C-terminal pdt variant fusion and a 106-bp scar, including the remaining FRT site. (**b**) Growth of strains after protease-driven depletion of MurA. Protease induction during early exponential-phase growth (arrow) causes cells containing *murA-pdt#1* to lyse within 1 h, as measured by optical density at 600 nm (OD_{600}). Cells containing the weakened pdt letter variants show a delayed response. Error bars show the s.d. from the mean of six biological replicates. Data showing wild-type growth of uninduced cells are shown in **Supplementary Figure 7**. (**c**) Differential interference contrast fluorescence overlay images of cells after ATc induction for 3 h. Cells containing *ftsZ-pdt#5* form filaments, whereas untagged wild-type bacteria maintain normal length. The fluorescence micrograph overlay showing constitutive GFP expression serves as a visual aid. (**d**) Disk diffusion assay on a chemotactic motility plate showing loss of chemotactic motility due to pdt-dependent CheZ degradation. Cells were stabbed into the chemotaxis plate after addition of 250 ng ATc to the center disk. (**e**) Cells containing *murA-pdt#1D* show increased sensitivity to fosfomycin after simultaneous induction with 4 ng/ml ATc (induced). OD_{600} measurements were taken 4 h after ATc and fosfomycin treatment and are presented as a percentage of the OD_{600} of cells not exposed to fosfomycin (untreated). In the *murA-pdt#1D* strain, $P < 0.001$ when comparing uninduced and induced cells for fosfomycin concentrations between 0.05 and 2 µg/ml using a paired-sample, one-tailed *t*-test. Additional data are shown in **Supplementary Figure 8**. (**f**) Pdt-dependent degradation of RecA causes hypersensitivity to the quinolone norfloxacin that matches the known hypersensitivity of a *recA* deletion strain (*ΔrecA*). Where indicated, cells were induced with 50 ng/ml ATc for 2 h before treatment with norfloxacin (25 ng/ml) for 2 h. Survival was measured by CFUs and is presented as a percentage of CFUs measured immediately before norfloxacin treatment. $P < 0.001$ for ATc induction of the *recA-pdt#3* strain using a paired-sample, one-tailed *t*-test. (**g**) Scatter plot displaying the relative growth and CFU count of EPD library members after targeted *mf*-Lon degradation. Growth and CFU measurements are displayed as a ratio of induced to uninduced cells at 4 h after ATc induction (50 ng/ml). CFU data points were placed at 1.0×10^{-6}, the limit of detection, when colonies were not recovered from the induced well. Error bars (**e**–**g**) show the s.d. from the mean of three biological replicates. CFU ratios represent results from a single experiment. Additional details are shown in **Supplementary Table 2.** (**h**) Histogram of PI staining for EPD library members. PI was measured by flow cytometry 2 h after induction with 50 ng/ml ATc. The percentage of cells that are PI^+ is displayed. The data are normalized to mode and are representative of three biological replicates.

The synthetic degradation system presented here is facile and modular, comprising a single protease gene and a small peptide tag that provides control over both the steady-state level and the inducible degradation rate of attached proteins. This system represents an important advance in synthetic biology, where control over protein level will provide an additional regulatory mechanism to aid in complex circuit design[23,24]. As we demonstrate here for a transcription-based toggle switch, the existing synthetic circuitry can be readily modified with our system to enable post-translational control while leaving the original regulatory framework intact, and use of our system to integrate multiple synthetic circuits can easily be envisioned[25,26]. Recent work by Huang *et al.*[27] also used *mf*-Lon to create a toggle switch, and Prindle *et al.*[28] used competition for the endogenous ClpXP protease to couple synthetic oscillatory circuits in *E. coli*, further demonstrating the utility of protease-driven control in engineered systems.

This platform provides tunable control of endogenous bacterial systems and can serve as both a basic research tool and an applied system to aid in antibiotic discovery. Single-step genomic insertion provides a simple and efficient method to target pdt fusions to *E. coli* genes, and the ability to control endogenous systems without disrupting the existing regulatory networks should be particularly useful in metabolic engineering, where dynamic control of *mf*-Lon may be used to regulate flux through targeted metabolic pathways (**Supplementary Fig. 9**)[29]. For the EPD library presented here, the system's inducible protease effectively converts the inducer into an 'antibiotic' that can specifically target nearly 80% of the essential proteins in *E. coli*. By this perspective, the EPD library may facilitate screens for synergistic and antagonistic relationships among existing and potential antibiotic targets and may also serve as a basic research tool to study essential gene function. As we observed for *murA-pdt#1* (**Fig. 4b**), the selective pressure caused by essential protein degradation may result in the outgrowth of cells containing protease mutations, potentially limiting the experimental window after *mf*-Lon induction.

Mf-Lon and pdt are based on a unique tmRNA system found only in *Mycoplasma*[5], so as we show here for *L. lactis*, the system should be transferable to many other organisms, including microbial pathogens such as *Staphylococcus aureus* and *Listeria monocytogenes*, where

protocols for genomic integration and inducible protein expression are well established. As in *E. coli*, the degradation system may best be used to study essential genes, but the inducible nature of the system may also provide a means to study the temporal function of virulence genes during infection. Additional pdt number variants may be needed to control degradation by endogenous proteases in these organisms, but the pdt letter variants that determine *mf*-Lon degradation rates should function as predicted in this study.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. GenBank: KM521209, KM521211, KM521207, KM521210, KM521212 and KM521208.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We are grateful to C. Bashor, M. Lobritz, M. Khalil and D. Dwyer for helpful discussions and critical reviews of the manuscript. This work was supported by funding from the Office of Naval Research (ONR) MURI Program, Defense Threat Reduction Agency grant HDTRA1-14-1-0006 and the Howard Hughes Medical Institute.

AUTHOR CONTRIBUTIONS

D.E.C. and J.J.C. conceived the study, analyzed data and wrote the paper. D.E.C. designed and performed the experiments.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Guzman, L.M., Belin, D., Carson, M.J. & Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* **177**, 4121–4130 (1995).
2. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1–I2 regulatory elements. *Nucleic Acids Res.* **25**, 1203–1210 (1997).
3. Isaacs, F.J. *et al.* Engineered riboregulators enable post-transcriptional control of gene expression. *Nat. Biotechnol.* **22**, 841–847 (2004).
4. Lou, C., Stanton, B., Chen, Y.J., Munsky, B. & Voigt, C.A. Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat. Biotechnol.* **30**, 1137–1142 (2012).
5. Janssen, B.D. & Hayes, C.S. The tmRNA ribosome-rescue system. *Adv. Protein Chem. Struct. Biol.* **86**, 151–191 (2012).
6. Neklesa, T.K. *et al.* Small-molecule hydrophobic tagging-induced degradation of HaloTag fusion proteins. *Nat. Chem. Biol.* **7**, 538–543 (2011).
7. Bonger, K.M., Chen, L.C., Liu, C.W. & Wandless, T.J. Small-molecule displacement of a cryptic degron causes conditional protein degradation. *Nat. Chem. Biol.* **7**, 531–537 (2011).
8. Davis, J.H., Baker, T.A. & Sauer, R.T. Small-molecule control of protein degradation using split adaptors. *ACS Chem. Biol.* **6**, 1205–1213 (2011).
9. Gur, E. & Sauer, R.T. Evolution of the ssrA degradation tag in *Mycoplasma*: specificity switch to a different protease. *Proc. Natl. Acad. Sci. USA* **105**, 16113–16118 (2008).
10. Ge, Z. & Karzai, A.W. Co-evolution of multipartite interactions between an extended tmRNA tag and a robust Lon protease in *Mycoplasma*. *Mol. Microbiol.* **74**, 1083–1099 (2009).
11. Mierau, I. & Kleerebezem, M. 10 years of the nisin-controlled gene expression system (NICE) in *Lactococcus lactis*. *Appl. Microbiol. Biotechnol.* **68**, 705–717 (2005).
12. Gardner, T.S., Cantor, C.R. & Collins, J.J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
13. Litcofsky, K.D., Afeyan, R.B., Krom, R.J., Khalil, A.S. & Collins, J.J. Iterative plug-and-play methodology for constructing and modifying synthetic gene networks. *Nat. Methods* **9**, 1077–1080 (2012).
14. Brown, E.D., Vivas, E.I., Walsh, C.T. & Kolter, R. MurA (MurZ), the enzyme that catalyzes the first committed step in peptidoglycan biosynthesis, is essential in *Escherichia coli*. *J. Bacteriol.* **177**, 4194–4197 (1995).
15. Adams, D.W. & Errington, J. Bacterial cell division: assembly, maintenance and disassembly of the Z ring. *Nat. Rev. Microbiol.* **7**, 642–653 (2009).
16. Silversmith, R.E. Auxiliary phosphatases in two-component signal transduction. *Curr. Opin. Microbiol.* **13**, 177–183 (2010).
17. Roemer, T. & Boone, C. Systems-level antimicrobial drug and drug synergy discovery. *Nat. Chem. Biol.* **9**, 222–231 (2013).
18. DeVito, J.A. *et al.* An array of target-specific screening strains for antibacterial discovery. *Nat. Biotechnol.* **20**, 478–483 (2002).
19. Kim, D.H. *et al.* Characterization of a Cys115 to Asp substitution in the *Escherichia coli* cell wall biosynthetic enzyme UDP-GlcNAc enolpyruvyl transferase (MurA) that confers resistance to inactivation by the antibiotic fosfomycin. *Biochemistry* **35**, 4923–4928 (1996).
20. Peterson, E.J., Janzen, W.P., Kireev, D. & Singleton, S.F. High-throughput screening for RecA inhibitors using a transcreener adenosine 5′-O-diphosphate assay. *Assay Drug Dev. Technol.* **10**, 260–268 (2012).
21. Wei, J.R. *et al.* Depletion of antibiotic targets has widely varying effects on growth. *Proc. Natl. Acad. Sci. USA* **108**, 4176–4181 (2011).
22. Zhu, W. *et al.* Antibacterial drug leads targeting isoprenoid biosynthesis. *Proc. Natl. Acad. Sci. USA* **110**, 123–128 (2013).
23. Cameron, D.E., Bashor, C.J. & Collins, J.J. A brief history of synthetic biology. *Nat. Rev. Microbiol.* **12**, 381–390 (2014).
24. Weber, W. & Fussenegger, M. Emerging biomedical applications of synthetic biology. *Nat. Rev. Genet.* **13**, 21–35 (2012).
25. Slusarczyk, A.L., Lin, A. & Weiss, R. Foundations for the design and implementation of synthetic genetic circuits. *Nat. Rev. Genet.* **13**, 406–420 (2012).
26. Callura, J.M., Cantor, C.R. & Collins, J.J. Genetic switchboard for synthetic biology applications. *Proc. Natl. Acad. Sci. USA* **109**, 5850–5855 (2012).
27. Huang, D., Holtz, W.J. & Maharbiz, M.M. A genetic bistable switch utilizing nonlinear protein degradation. *J. Biol. Eng.* **6**, 9 (2012).
28. Prindle, A. *et al.* Rapid and tunable post-translational coupling of genetic circuits. *Nature* **508**, 387–391 (2014).
29. Holtz, W.J. & Keasling, J.D. Engineering static and dynamic control of synthetic pathways. *Cell* **140**, 19–23 (2010).
30. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* **97**, 6640–6645 (2000).

ONLINE METHODS

Strains and reagents. The *E. coli* K-12 derivative strain MG1655Pro (F-, λ-, Sp^r, lacI, tetR) published previously[2,31] was used as the wild-type strain in all cases except for the synthetic toggle experiments, where we used MG1655*ΔlacIΔaraBAD*. Unless otherwise noted, *E. coli* were grown in Luria broth (LB) at 30 °C with shaking, and *mf*-Lon expression was induced with 50 ng/ml ATc. *L. lactis* strain NZ9000 (ref. 11) was used for all *L. lactis* experiments and was grown in M17 broth containing 0.5% glucose. The antibiotics carbenicillin (100 μg/ml), kanamycin (30 μg/ml) and erythromycin (10 μg/ml) were added to the media when appropriate.

Strain construction. The parental strains for all experiments were *E. coli* MG1655 (ATCC 47076) and *L. lactis* NZ9000 (ref. 11). MG1655*ΔlacIΔaraBAD* was created through P1 phage transduction of *lacI::kanR* from the Keio collection[32] into MG1655, and Red recombinase–mediated homologous recombination was used to create the in-frame deletion of *araBAD* according to published methods[30]. Flp recombinase, expressed on pECA102 (10 mM arabinose for 4 h), was used to remove the *kanR* cassette in each case. DH5αλpir[33] was used for cloning. Endogenous *E. coli* protease deletions were constructed by P1 phage transduction of the corresponding mutations from the Keio collection into MG1655Pro[31] followed by *kanR* cassette removal as detailed above. To construct the *mf*-Lon expression cassette, we codon optimized *mf-lon* for expression in *E. coli*, forward engineered a strong ribosome binding site to enable high expression[34] and cloned the cassette into pZE11. We used overlapping PCR to add a transcriptional terminator 5′ of the P_{LtetO} promoter in this plasmid and then cloned this expression cassette, which includes the 5′ terminator, P_{LtetO} promoter, *mf-lon* gene and 3′ terminator, into pWM91-lacZ, a derivative of pWM91 (ref. 35) that includes a 1-kb lacZ targeting region. The resulting plasmid, pECL275 (GenBank accession number KM521209), was introduced into MG1655Pro by conjugation from Sm10λpir, and single integrants were selected on carbenicillin plates, grown in rich medium for 8 h and then selected on plates containing 1% tryptone, 0.5% yeast extract, 8% sucrose and 1.5% agar to select for plasmid excision. The resulting colonies were screened by PCR for the *mf*-Lon expression cassette. *Mf*-Lon variants that contain *ec*-AAV and *ec*-ASV fusions and the S692A point mutation were constructed in the same manner.

***E. coli*–based degradation platform.** The *mf-lon* gene was codon optimized for *E. coli* expression, placed under control of the P_{LtetO} promoter and integrated into the *lacZ* locus along with 5′ and 3′ transcriptional terminators to block unwanted *mf*-Lon expression from any proximal genomic promoters. The GFP variant GFPmut3b[36] was used for all GFP expression. To express GFPmut3b and mCherry-pdt#3 fusions, the constitutive P_{lacIq} promoter[37] was fused to GFP and mCherry and inserted into pZE21-MCS[2] to form pZE27GFP3 and pZE27MC3, respectively. For these plasmids, the last number and letter indicate the pdt variant used (for example, pZE27GFP5B contains pdt#5B).

PDT mutagenesis screens. Pdt mutant libraries were created by PCR using primers containing randomized nucleotides at the indicated pdt codons. Strains containing the GFP-pdt mutants were individually picked into 96-well plates and measured by plate fluorimetry during exponential-phase growth. Strains that exhibited the desired GFP degradation dynamics after induction with ATc were further characterized by flow cytometry.

Synthetic toggle switch. The plasmid pKDL071R8, based on pKDL071 (ref. 13), was altered to contain a weakened *tetR* ribosome binding site to enhance toggle bistability in the minimal medium conditions used. This plasmid served as the parental strain for all LacI-pdt toggle switch experiments. Pdt#3 was fused to *lacI* by overlapping PCR and cloned into pKDL071R8 to make pECJ3 (GenBank accession number KM521210). To express *mf*-Lon, the arabinose-inducible P_{BAD} promoter was fused to *mf*-Lon and cloned into pZA11 to form pZA16mflon. Cells containing the toggle switch and *mf*-Lon expression plasmids were grown in 200 μl in 96-well round-bottom plates at 37° in M9 minimal medium containing 0.2% glycerol and 0.05% casamino acids, and care was taken to maintain exponential growth throughout the experiment. Cells were grown for 6 h with either 30 ng/ml ATc or 500 μM IPTG to induce cells into the GFP^+ or $mCherry^+$ state, respectively. Cells were diluted 1:1,000 into noninducing medium and allowed to grow for an additional 12 h. To induce *mf*-Lon, cells were grown at 37° with shaking and 1 mM arabinose and passaged every 4 h (~1:10 dilution) into medium containing the same inducing conditions. At each time point, cells were fixed with 1% paraformaldehyde in PBS and stored at 4° for up to 5 d. Cells that did not contain the toggle switch plasmid were used to define the $GFP^-mCherry^-$ state shown in **Figure 3b**.

***L. lactis*–based degradation platform.** The plasmid pECGMC was created to enable expression of *mf*-Lon and mCherry in *L. lactis*. Based on the plasmid pZE11-MCS[2], pECGMC contains the ColE1 origin of replication and the *ampR* ampicillin resistance cassette to enable cloning in *E. coli*, as well as the AMβ1 origin of replication and the *ermR* erythromycin resistance cassette from pVE5523 (ref. 38) to enable replication and selection in *L. lactis*. In pECGMC, mCherry is expressed from the constitutive P32 promoter[39], and *mf*-Lon is expressed from the inducible promoter P_{nisA}[11] with 3 ng/ml nisin. Plasmids containing pdt fusions to mCherry are labeled according to the pdt variant used (for example, pECGMC3 contains pdt#3). Plasmids were transformed into *L. lactis* according to established protocols[40], and the sequence of pECGMC3 was deposited in GenBank under accession number KM521211.

Genomic insertion of pdt variants. The pECT plasmids were created to serve as a template for PCR amplification of the pdt variant cassettes shown in **Figure 4a**. The *kanR* cassette and surrounding FRT sites from pKD13 (ref. 32) were cloned into pWM91 (ref. 35) to generate pECT, which was further named according to the cloned pdt variant (for example, pECT3A contains pdt#3A). To generate pECA102, Flp recombinase was cloned into pBAD24 (ref. 1) using KpnI, and the constitutively expressed *sacB* cassette was subsequently amplified from pWM91 (ref. 35) and cloned into the plasmid using partial MluI and SalI digestion. Plasmids pECT3 and pECA102 were deposited in GenBank under accession numbers KM521207 and KM521212, respectively. To generate PCR products for genomic integration, pdt variants were amplified from their pECT plasmid using primers P1 and P2 that contained additional 42-base 5′ extensions with homology to the C terminus and the immediate 3′ untranslated region of the targeted gene, respectively. Of note, the endogenous gene's stop codon should not be included in the P1 5′ extension. The base P1 and P2 primer sequences and full-length primers used to target *murA*, *ftsZ*, *cheZ* and *recA* are as follows: P1: GCGGCGAACAAAAACGAA; P2: TTATGTAGGCTGGAGCTGC; P1-*murA*: CTGCGCGCTTTAGGTGCAAATATTGAGCGTGTGAAAGGCGAAGCGGCGAACAAAAACGAA; P2-*murA*: CTGGCGGTAGCCCCGCGAACGGGGCTGCCAGCTCTCAGACGATTATGTAGGCTGGAGCTGC; P1-*ftsZ*: GATTATCTGGATATCCCAGCATTCCTGCGTAAGCAAGCTGATGCGGCGAACAAAAACGAA; P2-*ftsZ*: GTTTAGCACAAAGAGCCTCGAAACCCAAATTCCAGTCAATTCTTATGTAGGCTGGAGCTGC; P1-*cheZ*: AGTCAGGATCAGGTGGACGATTTGTTGGATAGTCTTGGATTTGCGGCGAACAAAAACGAA; P2-*cheZ*: CCGCCTGATATGACGTGGTCACGCCACATCAGGCAATACAAATTATGTAGGCTGGAGCTGC; P1-*recA*: GTAGATGATAGCGAAGGCGTAGCAGAAACTAACGAAGATTTTGCGGCGAACAAAAACGAA; and P2-*recA*: AAAAGGGCCGCAGATGCGACCCTTGTGTATCAAACAAGACGATTATGTAGGCTGGAGCTGC. PCR products were transformed into *E. coli* containing pKD46 using published methods[30], and successful genomic pdt insertions were verified by PCR. Plasmid pECA102 was used to remove the *kanR* cassette and was subsequently cured by selection on LB plates containing 8% sucrose.

EPD library. The EPD library was created using pECM1, a derivative of pECT1 that includes MmeI sites on both sides of the insertion cassette, resulting in an extended pdt tag that includes the additional N-terminal amino acids VG. The full pdt#1 tag is VGAANKNEENTNEVPTFMLNAGQANRLQL. Genomic integration was performed according to the protocol described above for individual pdt insertions, except that electroporation was performed using the Gemini X2 System (BTX) with an HT100 plate handler, and the kanamycin cassette was not removed. The list of targeted essential genes was collated manually from Ecogene[41] and Baba *et al.*[32], and a list of the primers used can be found in **Supplementary Table 3**. Plasmid pECM1 was deposited in GenBank under accession number KM521208.

doi:10.1038/nbt.3053

Flow cytometry. Data for GFP and mCherry degradation dynamics and PI staining were collected using an LSRFortessa cell analyzer equipped with a High Throughput Sampler (BD Biosciences), and data for the synthetic toggle switch were collected using a FACSAriaII flow cytometer (BD Biosciences). For each GFP and mCherry measurement, cells were fixed in 1% paraformaldehyde, held at 4° for up to 5 d and then diluted 1:10 in PBS for analysis. At least 5,000 cells were collected for each measurement, and FloJo (Treestar) was used for data analysis.

Plate fluorimetry and optical density. Fluorescence and optical density measurements were made with a SpectraMax M5 microplate reader (Molecular Devices) using excitation and emission wavelengths of 488 nm and 520 nm, respectively, with an emission filter cutoff at 515 nm. Optical density was measured at 600 nm (OD_{600}). All measurements were made in 200 µl in 96-well flat-bottom plates.

MurA-induced lysis growth conditions. Strains were grown in 200 µl LB in 96-well flat-bottom plates with lids at 30 °C with shaking in a SpectraMax M5 plate reader. OD_{600} measurements were taken every 15 min and were normalized using measurements from medium-only wells. Wells on the perimeter of the plate were filled with water and were not used for bacterial growth.

FtsZ microscopy. Differential interference contrast and fluorescence microscopy images were taken with a Nikon Eclipse Ti microscope using a 100× objective and a Coldsnap HQ2 charge-coupled device (CCD) camera (Photometrics) operated with NIS-Elements Advanced Research 3.2 software. Cells in exponential growth in liquid cultures were induced with ATc, grown for 3 h at 30 °C, placed on a 300 µl pad containing PBS and 0.75% low-melt agarose (Boston Bioproducts) and imaged immediately.

Chemotactic motility plates. Cells in exponential growth were stabbed into soft agar plates containing 1% tryptone, 0.5% NaCl and 0.3% agar. ATc dissolved in 10 µl water was added to sterile 6-mm disks in the center of the plates immediately before bacterial inoculation. Plates were incubated for 18 h at 30 °C before imaging with a Gel Logic 6000 Pro (Carestream).

Hypersensitivity assay. Fosfomycin and ATc were added simultaneously at the indicated concentrations to cells in exponential growth in 96-well flat-bottom plates, and OD_{600} was measured after 4 h, with medium-only wells serving as absorbance controls. For targeted RecA degradation, cells in exponential growth were induced with 50 ng/ml ATc for 2 h, if indicated, before treatment with 25 ng/ml norfloxacin for 2 h. Cells were then serially diluted in PBS and spotted on LB plates without selection, and visible CFUs were counted after incubation overnight at 30 °C.

CFU assay. Cells with and without ATc induction were grown with shaking in 96-well round-bottom plates for 4 h and then serially diluted in PBS before plating on LB agar without selection for growth overnight at 30 °C.

PI assay. Cells in exponential growth were induced with ATc for 4 h in 96-well round-bottom plates, diluted 1:100 in PBS containing 5 µM PI and analyzed by flow cytometry within 20 min. The threshold for PI^+ cells was defined using untagged wild-type cells and is presented as a percentage of the total whole-cell population for each strain.

Statistical analyses. Linear regression models used the ordinary least-squares approach to determine the best-fit line, coefficient of determination (R^2) and s.d. All observed *y*-axis data were used to calculate the best-fit lines. *P* values were calculated using a paired-sample, one-tailed *t*-test.

31. Callura, J.M., Dwyer, D.J., Isaacs, F.J., Cantor, C.R. & Collins, J.J. Tracking, tuning, and terminating microbial physiology using synthetic riboregulators. *Proc. Natl. Acad. Sci. USA* **107**, 15898–15903 (2010).
32. Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
33. Metcalf, W.W., Jiang, W. & Wanner, B.L. Use of the rep technique for allele replacement to construct new *Escherichia coli* hosts for maintenance of R6Kγ origin plasmids at different copy numbers. *Gene* **138**, 1–7 (1994).
34. Salis, H.M., Mirsky, E.A. & Voigt, C.A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27**, 946–950 (2009).
35. Metcalf, W.W. *et al.* Conditionally replicative and conjugative plasmids carrying lacZα for cloning, mutagenesis, and allele replacement in bacteria. *Plasmid* **35**, 1–13 (1996).
36. Cormack, B.P., Valdivia, R.H. & Falkow, S. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33–38 (1996).
37. Müller-Hill, B., Crapo, L. & Gilbert, W. Mutants that make more lac repressor. *Proc. Natl. Acad. Sci. USA* **59**, 1259–1264 (1968).
38. Dieye, Y., Usai, S., Clier, F., Gruss, A. & Piard, J.C. Design of a protein-targeting system for lactic acid bacteria. *J. Bacteriol.* **183**, 4157–4166 (2001).
39. van de Guchte, M., Kok, J. & Venema, G. Gene expression in *Lactococcus lactis*. *FEMS Microbiol. Rev.* **8**, 73–92 (1992).
40. Holo, H. & Nes, I.F. Transformation of *Lactococcus* by electroporation. *Methods Mol. Biol.* **47**, 195–199 (1995).
41. Zhou, J. & Rudd, K.E. EcoGene 3.0. *Nucleic Acids Res.* **41**, D613–D624 (2013).