SI Appendix

Aoyama, Raina et al.

Bacterial Strains and Plasmids. Bacterial strains, plasmids, and oligonucleotides used in this study are listed in Dataset S1. *E. coli* strains are derivatives of wild-type MG1655 (F- lambda- *ilvG- rfb*-50 *rph*-1). Tagged strains were generated by λ Red–mediated recombineering (1) using NM400 and the oligonucleotides listed in Appendix Table S3. Alleles marked by antibiotic markers were moved between strains by P1 transduction. When necessary, kanamycin resistance cassettes were excised from the chromosome by FLP-mediated recombination (2). All mutations and fusions were confirmed by sequencing. Descriptions of plasmids used in this study are in Appendix Table S2. All SpfP derivatives were overexpressed from pKK177-3 (a derivative of pKK223-3 lacking the region between at BamHI at 256 and PvuII at 1,945), while Spot 42 WT and the Spot 42_{STOP} mutant were overexpressed from the pRI derivative of pKK177-3 in which an EcoRI site was introduced at the P_{tac} +1 site (3). The Spot 42 and SpfP derivatives were ordered as primers listed in Table S3, annealed and cloned into pKK177-3 or pRI digested with the EcoRI and HindIII restriction enzymes. All CRP derivatives were overexpressed from pACYC184 (containing the p15a origin of replication to allowing the pACYC184 and pKK177-3 plasmids and their derivatives to replicate in the same cells). CRP was amplified from the chromosome and ligated into pACYC184 between the HindIII and BamHI sites. Site-directed mutagenesis was carried out using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent). All inserts were confirmed by sequencing.

Hydroxylamine Mutagenesis. pACYC-CRP DNA (10 µg in 20 µl of H_2O) was mixed with 100 µl of 0.5 M K_2HPO_4, 5mM EDTA solution (pH 6) and 80 µl of hydroxylamine (pH 6). The mixture was incubated at 65˚C for 45 min before the treated DNA was purified using a NucleoSpin Gel and PCR cleanup kit (Macherey-Nagel). The mutagenized DNA was transformed into Δ*spf* Δ*crp* double deletion cells carrying pKK-SpfP and transformants were plated onto M63 media supplemented with 0.2% galactose. Plasmid DNA isolated from large colonies that arose overnight were submitted for sequencing.

Bacterial Growth Conditions. Unless indicated otherwise, bacterial strains were grown with shaking at 250 rpm at 37˚C in either LB rich or M63 minimal medium supplemented with 0.001% vitamin B1 and glucose, galactose or maltose at 0.2%, and sorbitol at 0.4%. Where indicated, media contained antibiotics with the following concentrations: ampicillin (100 µg/ml), chloramphenicol (25 µg/ml), and kanamycin (30 µg/ml).

Growth Assay. Colonies of Δ*spf::kan* (GSO433) transformed with indicated plasmids were grown overnight in LB with ampicillin and diluted to OD_{600}=0.05 in M63 minimal medium supplemented with 0.001% vitamin B1 and either 0.2% glucose or 0.2% galactose. OD_{600} was measured after 16h.

Immunoblot Analysis. Aliquots (1 ml) of cells were grown in the indicated medium were collected and the cell pellet was resuspended in 1XP BS and 10 ul were loaded on a Mini-PROTEAN TGX 5%–20% Tris-Glycine gel (Bio-Rad) and run in 1X Tris Glycine SDS buffer (KD biomedical). The gel was electro transferred into nitrocellulose membranes (Invitrogen) for

1 h at 100 V. Membranes were blocked in 1X PBS 0.1% of Tween 20 (PBS-T) with either 5% milk (Bio-Rad, for α-FLAG antibodies) or 5% BSA (for all other antibodies) for 1 h and subsequently probed with a 1:1,000 dilution of α-FLAG antibody (Sigma), 1:1,000 dilution of α-His antibody (Qiagen), 1:1,000 dilution of α-β-galactosidase antibody (Abcam), 1:1,000 dilution of α-GalE antibody (kind gift of S. Adhya) or 1:1,000 dilution of α-β-galactosidase antiserum (Abcam) in PBS-T with 5% milk or 5% BSA as above for 1 h or longer. Incubation with the rabbit α-GalE antibodies or the rabbit α-β-galactosidase antibodies was followed by incubation with a 1:4,000 dilution of HRP-conjugated α-rabbit antibodies. After four washes with PBS-T, the membranes were developed with an Amersham ECL Western Blotting Detection Kit (GE Healthcare).

RNA Isolation. Cells corresponding to the equivalent of 10 OD_{600} were collected by centrifugation, and snap frozen in liquid nitrogen. RNA was extracted according to the standard TRIzol (Thermo Fisher Scientific) protocol. Briefly, 1 ml of room temperature TRIzol was added to cell pellets, resuspended thoroughly to homogenization, and incubated for 5 min at room temperature. After the addition of 200 µl of chloroform and thorough mixing by inversion, samples were incubated for 10 min at room temperature. After samples were centrifuged for 10 min at 4˚C on maximal speed, the upper phase (~0.6 ml) was transferred into a new tube and 500 µl of isopropanol was added. Samples again were mixed thoroughly by inversion, incubated for 10 min at room temperature and centrifuged at maximal speed for 15 min at 4˚C. RNA pellets were washed twice with 75% ethanol. After the second wash, the ethanol was aspirated, and the RNA pellet was left to dry at room temperature. RNA was resuspended in 20-50 µl of DEPC water and quantified using a NanoDrop (Thermo Fisher Scientific).

Northern Analysis. Total RNA (5-10 μg per lane) was separated on denaturing 8% polyacrylamide gels, urea gels containing 6 M urea (1:4 mix of Ureagel Complete to Ureagel-8 (National Diagnostics) with 0.08% ammonium persulfate) in 1X TBE buffer at 300V for 90 min. The RNA was transferred to a Zeta-Probe GT membrane (Bio-Rad) at 20 V for 16 h in 0.5X TBE, UV-crosslinked, and probed with ^{32}P-labeled oligonucleotide (Listed in Appendix Table S3) in ULTRAhyb-Oligo buffer (Ambion Inc.) at 45˚C. Membranes were washed twice with 2X SSC/0.1% SDS at room temperature, once with 0.2X SSC/0.1% SDS at room temperature, washed for 25 min with 0.2 × SSC/0.1% SDS at 45˚C, followed by a final rinse with 0.2X SSC/0.1% SDS at room temperature before autoradiography was performed with HyBlot CL film (Denville Scientific Inc.).

β-galactosidase Assays. Cultures were grown in LB to OD_{600} ~1.0 with arabinose (0.2%). Cells (100 μl) were added to 700 μl of Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 50 mM β-mercaptoethanol). After the addition of 15 μl of 0.1% SDS and 30 μl of chloroform, each sample was vortexed for 30 s and then incubated at room temperature for 15 min to lyse the cells. The assay was initiated by the addition of 100 μl of ONPG (4 mg/ml). The samples were incubated at room temperature until the reaction was terminated by the addition of 500 μl of 1M Na_2CO_3. A_{420} and A_{550} values measured with a spectrophotometer were used to calculate Miller units.

Purification of Recoded SpfP. The Δ*spf::kan* strain (GSO433) transformed with pKK-SpfP-recoded$_{\text{N-terminal STOP}}$ or pKK-SpfP-recoded$_{\text{C-terminal STOP}}$ and pEVOL-*p*-AzF was grown in LB at

37˚C overnight. The overnight culture was diluted 1:100 into 1 l of M63 glucose minimal media and incubated at 37˚C. At OD_{600} ~0.5, 0.2% arabinose was added to induce the orthologous aaRS and tRNA pair along with 1 mM *p*-azido phenylalanine (Chem-Impex Int'l. Inc.) and incubated for another 3 h. Cells were collected by centrifugation (4,650 × g, 20 min) and the pellet was resuspended in 30 ml of TNG buffer [10 mM Tris (pH 7.5), 100 mM NaCl, 10% glycerol] supplemented with Protease Inhibitor Cocktail (Roche). The cells were lysed using a microfluidizer processor (Microfluidics) at 20,000 psi, and the insoluble cellular debris was removed by centrifugation (20,000 × g, 30 min). The cleared lysate was incubated with 5 μM $CuCl_2$, 1 mM ascorbic acid (pH adjusted to 7 by NaOH) and 5 μM of Biotin-PEG4-Alkyne (Click Chemistry Tools) dissolved in DMSO at room temperature by nutating for 30 min. The lysate was then supplemented with 25 mM DDM and incubated overnight at 4˚C with 500 ul of streptavidin agarose beads (Thermo Scientific). The lysate and beads were applied to a Bio-Spin disposable chromatography column (Bio-Rad Laboratories) and allowed to drain by gravity. The streptavidin column was washed with 15 ml TNG buffer with 2 mM DDM. Finally, proteins were eluted from the column with 1 ml of elution buffer (0.1 M glycine pH 3.5, 100 mM NaCl and 0.1% Triton X-100) and TCA precipitation was carried out to concentrate the protein samples. The pellet was resuspended in 30 μl of 2X Laemmli buffer. The samples were heated at 95˚C for 10 min, and aliquots were subjected to SDS/PAGE in a 10–20% Tris-glycine gel (Invitrogen) at 12 V/cm. Proteins were visualized with Coomassie Blue Stain. Bands of interest were excised from the gel and analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS).

Reciprocal Purification. Δ*spf* cells expressing CRP-HA-His_6 or GalK-HA-His_6 from the chromosome (GSO1061 and GSO1060, respectively) or Δ*spf::kan* Δ*crp::cm* cells GSO1063 were transformed with pKK-SpfP-$recoded_{N\text{-terminal FLAG}}$ and grown in LB at 37˚C for 16h, diluted 1:100 into 1L of LB and incubated at 37˚C. At ~OD=0.5 cells were collected by centrifugation (10,000 x g, 20min) and resuspended in 15 ml of TNG buffer supplemented with Protease Inhibitor cocktail (Roche). The cells were homogenized and incubated with Pierce α-HA magnetic beads (Thermo Scientific) in a 50 ml tube and incubated overnight at 4˚C. Beads and the bound protein were collected using a MagneSphere magnetic separation stand (Promega) and washed with 1 ml of TNG buffer (10X). The beads were resuspended in 1X PBS (25 ul) and 2X Laemmli buffer (25 ul) and heated at 95˚C for 5 min. Samples were analyzed on immunoblots using α-His and α-FLAG antibodies. Synthetic N-terminally biotin tagged SpfP (produced by Thermo Scientific) was added to lysates of Δ*spf* cells expressing CRP-HA-His_6 or Δ*spf* Δ*crp* cells prior to incubation with Pierce α-HA magnetic beads (Thermo Scientific).

CRP-SpfP Structure Prediction. CRP is known to form a homodimer (4). Although the CRP-SpfP stoichiometry has not been determined, it is reasonable to hypothesize that one SpfP subunit binds with one CRP subunit, forming a CRP-SpfP heterotetramer. The structure of CRP-SpfP complex was predicted with AlphaFold-Multimer (5) utilizing NIH's Biowulf cluster. The structure figures were prepared using PyMol.

1. Yu D*, et al.* (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci USA* 97:5978-5983.
2. Cherepanov PP, Wackernagel W (1995) Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* 158:9-14.

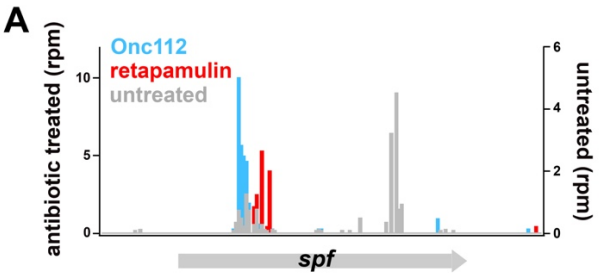
3. Opdyke JA, Kang J-G, Storz G (2004) GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J Bacteriol* 186:6698-6705.

4. Napoli AA, Lawson CL, Ebright RH, Berman HM (2006) Indirect readout of DNA sequence at the primary-kink site in the CAP-DNA complex: recognition of pyrimidine-purine and purine-purine steps. *J Mol Biol* 357:173-183.

5. Evans R*, et al.* (2021) Protein complex prediction with AlphaFold-Multimer. *bioRxiv*:10.1101/2021.1110.1104.463034.

6. Weaver J, Mohammad F, Buskirk AR, Storz G (2019) Identifying small proteins by ribosome profiling with stalled initiation complexes. *mBio* 10:e02819-02818.

7. Meydan S*, et al.* (2019) Retapamulin-assisted ribosome profiling reveals the alternative bacterial proteome. *Mol Cell* 74:481-493.



B

```
Escherichia coli                MFYLSDLLLHVIGFG  15
Erwinia billingiae              MFYLSNLLLHVIGFG  15
Salmonella enterica             MFYLSDLLLHVIGFG  15
Brenneria goodwinii             MFYLSDLLLHVIGFG  15
Hafnia alvei                    MFYLSDLLLHVIGFG  15
Buttiauxella agrestis           MFYLSDLLLHVIGFG  15
Pantoea ananatis                MFYLSDLLLHVIGFG  15
Raoultella ornithinolytica      MFYLSDLLLHVIGFG  15
Cedecea davisae                 MFYLSDLLLHVIGFG  15
Cronobacter sakazakii           MFYLSDLLLHVIGFG  15
Plesiomonas shigelloides        MFYLSDLLLHVIGFG  15
Pluralibacter gergoviae         MFYLSDLLLHVIGFG  15
Klebsiella pneumoniae           MFYLSDLLLHVIGFG  15
Serratia fonticola              MFYLSDLLLHVIGFG  15
Lonsdalea quercina              MFYLSDLLLHVIGFG  15
Tatumella morbirosei            MFYLSDLLLHVIGFG  15
Kluyvera ascorbata              MFYLSDLLLHVIGFG  15
Citrobacter koseri              MFYLSDLLLHVIGFG  15
Shigella dysenteriae            MFYLSDLLLHVIGFG  15
Enterobacter aerogenes          MFYLSDLLLHVIGFG  15
Candidatus sodalis              MFYLSDLLLHVIGFG  15
Leclercia adecarboxylata        MFYLSDLLLHVIGFG  15
Edwardsiella ictaluri           MFYLSDLLLHVIGFG  15
Sodalis praecaptivus            MFYLSDLLLHVIGFG  15
Trabulsiella quamensis          MFYLSDLLLHVIGFG  15
Yokenella regensburgei          MFYLSDLLLHVIGFG  15
Rahnella aquatilis              MFYLSDLLLHVIGFG  15
Chania multitudinisentens       MFYLSDLLLHVIGFG  15
Shimwellia blattae              MFYLSDLLLHVIGFG  15
Enterobacter cloacae            MFYLSDLLLHVIGFG  15
Yersinia mollaretii             MFYLSDLLLHVIGFG  15
Morganella morganii             MFYLSDLLLHVIGFS  15
Photorhabdus luminescens        MFYLSDLLLHVIGFS  15
Proteus penneri                 MFYLSDLLLHVIGFS  15
Moellerella wisconsensis        MFYLSDLLLHVIGFS  15
Xenorhabdus beddingii           MFYLSDLLLHVIGFS  15
Providencia burhodogranariea    MFYLSDLLLHVIGFS  15
Enterovibrio coralii            MFYLSDLLFHVIGFS  15
Grimontia hollisae              MFYLSDLLFHVIGFS  15
Vibrio cholerae                 MFYLSDLLFHVIGLG  15
Photobacterium aphoticum        MFYLSDLLFHVIGFG  15
Prevotella dentalis             MFY--DDLLHVIGFE  13
Hallella seregens               MFY--DDLLHVIGFE  13
Rhodopirellula maiorica         MFYL-GLLLQVIGLA  14
                                ***  . *::**
```

Fig. S1. Ribosome binding to and conservation of *spf*. (*A*) Ribosome density on the *spf* gene detected by ribosome profiling in the presence of the translation inhibitors Onc112-treated (blue) (6), retapamulin-treated (red) (7) and an untreated control sample (gray) (6). (*B*) The SpfP amino acid sequences from *E. coli* K12 and other bacterial species aligned with ClustalW. "*" indicates the residues are identical in all sequences, and ":" and "."indicate that conserved and semi-conserved substitutions, respectively, as defined by ClustalW.

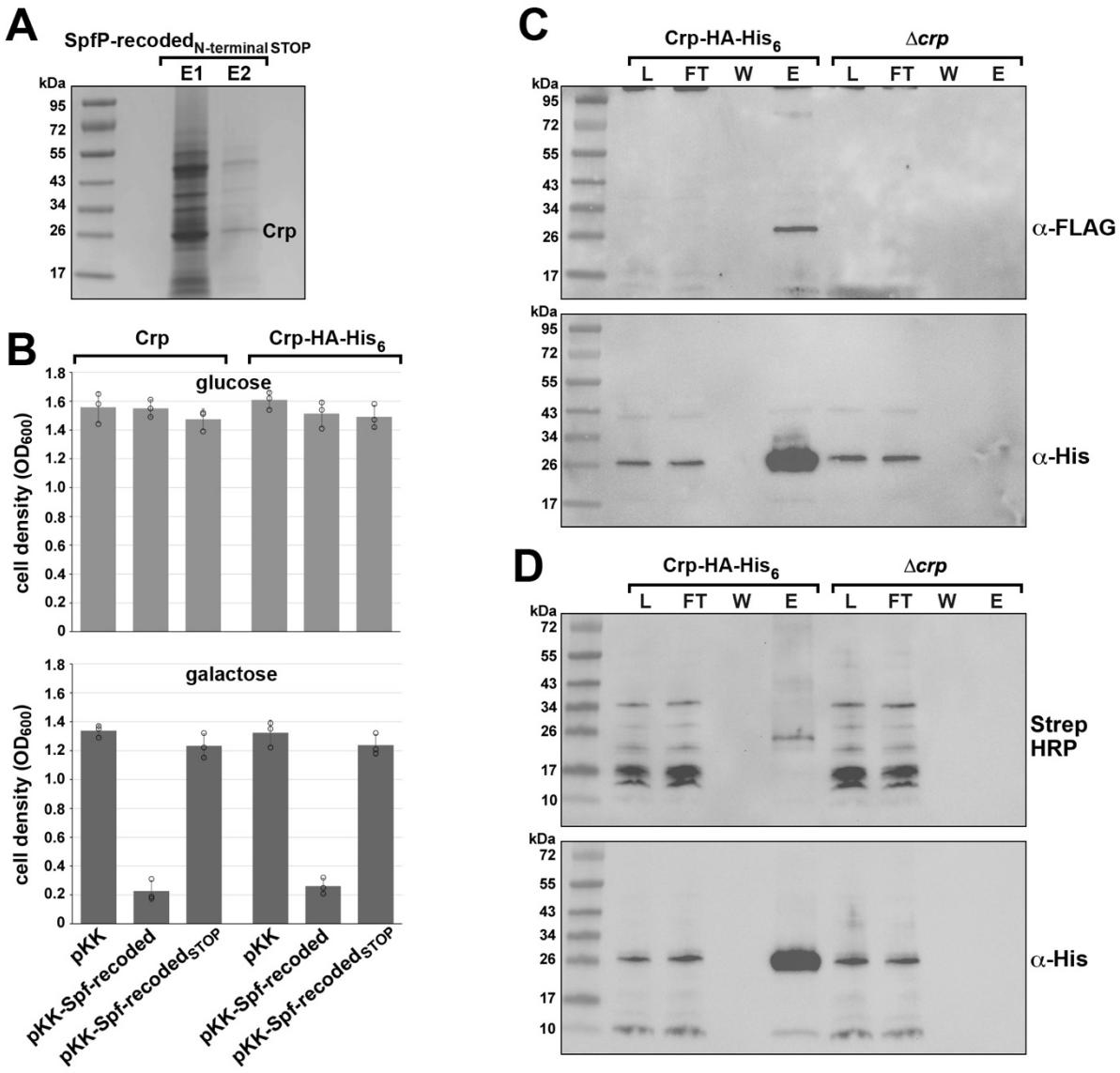


Fig. S2. Lack of SpfP-FLAG association in Δ*crp* mutant background. (*A*) CRP copurifies with biotin-tagged SpfP. WT cells (GSO433) carrying the SpfP-recoded$_{\text{N-terminal STOP}}$ (left) or YoaK$_{\text{N-terminal STOP}}$ (right) plasmids were both transformed with pEVOL-p-AzF, a plasmid that expresses an orthologous aminoacyl-tRNA and tRNA pair. Cells grown in LB to OD_{600}~ 0.5 were lysed and applied to streptavidin beads. After the beads were washed, the bound protein was eluted in two steps (E1 and E2). The samples were separated by SDS-PAGE and stained with Coomassie

blue. Unique bands in the eluate were sent for mass spectrometric analysis and CRP was identified in the indicated band. (*B*) Same repressive effect of SpfP is observed in *crp*$^{+}$ and *crp-HA-His*$_6$ backgrounds. Growth assays of Δ*spf::kan* (GSO433) and cells Δ*spf crp-HA-His*$_6$*::kan* (GSO1061) transformed with pKK, pKK-SpfP-recoded or pKK-SpfP-recoded$_{STOP}$ were grown in M63 minimal medium supplemented with either glucose (top panel) or galactose (bottom panel). Cells were grown overnight in LB with ampicillin were diluted to OD_{600}~0.05 in M63 minimal medium with the indicated carbon sources and grown for 16 h, at which point OD_{600} was measured. The average of three replicates is showed together with the standard deviation of the mean. (*C*) N-terminally FLAG tagged SpfP associates with ~25 kDa band in CRP-HA-His$_6$ but not Δ*crp* background. Δ*spf crp-HA-His*$_6$*::kan* (GSO1061) or Δ*spf::kan* Δ*crp::cm* (GSO1063) cells expressing N-terminally FLAG-tagged SpfP-recoded from the pKK plasmid were grown in LB to OD_{600}~ 0.5. Cells were lysed (L) and applied to anti-HA beads, and the flow through (FT) fraction was collected. Beads were washed (W) and the bound protein was eluted (E). The samples were examined on immunoblots using α-FLAG antibodies to detect SpfP-recoded$_{NFLAG}$ (top) or α-His antibodies to detect CRP-HA-His$_6$ (bottom panel). A cross-reacting band of high molecular weight is detected with the α-FLAG antibodies, and a cross-reacting band of the same size as CRP is detected with the α-His antibodies. (*D*) Synthetic SpfP associates with ~25 kDa band in CRP-HA-His$_6$ but not Δ*crp* background. Extracts prepared from Δ*spf crp-HA-His*$_6$*::kan* (GSO1061) or Δ*spf::kan* Δ*crp::cm* (GSO1063) cells grown LB to OD_{600}~ 0.5 were incubated with N-terminally biotinylated SpfP (biotin-MFYLSDLLLHVIGFG-COOH, ThermoFisher Scientific) for 2 h. The lysates (L) were applied to anti-HA beads, and the flow through (FT) fraction was collected. Beads were washed (W) and the bound protein was eluted (E). The samples were examined on immunoblots using streptavidin-HRP to detect biotin-SpfP (top

panel) or α-His antibodies to detect CRP-HA-His_6 (bottom panel). Some cross-reacting bands are seen with streptavidin-HRP and the α-His antibodies. One band that cross-reacts with the α-His antibodies migrates at the same position as CRP.

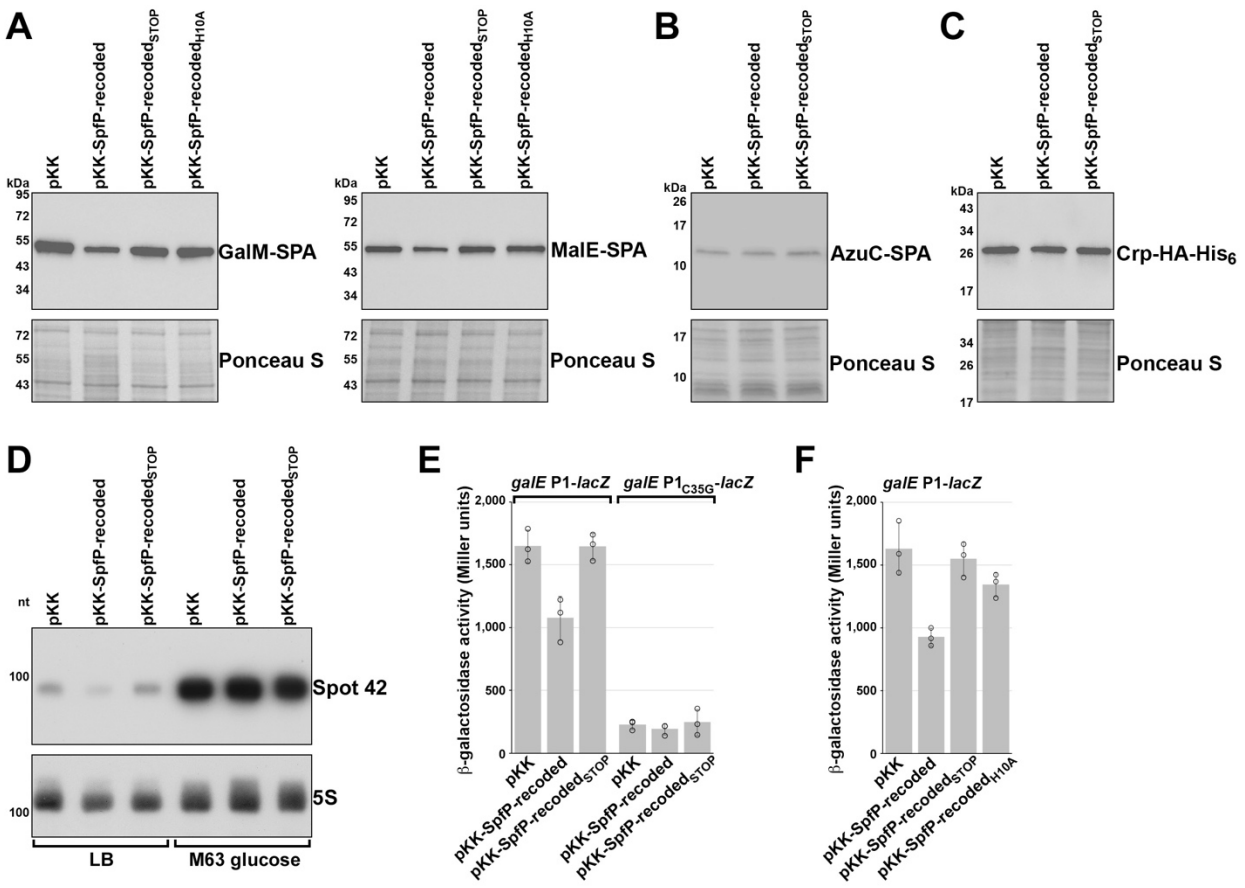


Fig. S3. Effects of SpfP and SpfP$_{H10A}$ overexpression (*A*) SpfP-recoded$_{H10A}$ is less effective at repressing GalM-SPA and MalE-SPA. Panels show same blots depicted in Fig. 4*A* but with an additional lane for cells transformed with pKK-SpfP-recoded$_{H10A}$. (*B*) Immunoblot blot analysis of expression of chromosomally-tagged AzuC-SPA (GSO351) from cells transformed with pKK, pKK-SpfP-recoded, or pKK-SpfP-recoded$_{STOP}$ and grown in LB to OD_{600}~0.5. α-FLAG antibody was used to detect the SPA tag. (*C*) Immunoblot blot analysis of expression of chromosomally-tagged CRP-HA-His$_6$ in Δ*spf crp-HA-His$_6$::kan* (GSO1061) cells transformed with pKK, pKK-SpfP-recoded, or pKK-SpfP-recoded$_{STOP}$. Cells were grown in M63 galactose to OD_{600}~0.5. α-His antibody was used to detect CRP. (*D*) Northern blot analysis of Spot 42 levels in WT (MG1655) cells transformed with pKK, pKK-SpfP-recoded, or pKK-SpfP-recoded$_{STOP}$. Cells

were grown in LB or M63 glucose to OD_{600}~0.5. The Spot 42 and 5S RNAs were detected with primers specific to each of these transcripts. The Spot 42 primer cannot detect the SpfP-recoded transcript. (*E*) β-galactosidase activity assay of transcriptional fusion of the P1 promoter of the galactose operon to *lacZ* (*galE* P1-*lacZ*) (GSO1069) and the same fusion with a C35G mutation (GSO1070). Both fusion strains carrying Δ*spf* (GSO1071 and GSO1072, respectively) were transformed with pKK, pKK-SpfP-recoded, and pKK-SpfP-recoded$_{STOP}$. (*F*) SpfP-recoded$_{H10A}$ is less effective at repressing *galE* P1-*lacZ*. β-galactosidase activity assay of *galE* P1-*lacZ* Δ*spf* (GSO1071) cells transformed with pKK, pKK-SpfP-recoded, pKK-SpfP-recoded$_{STOP}$, and pKK-SpfP-recoded$_{H10A}$. For (*D*) and (*E*), β-galactosidase assays were performed for cells grown to OD_{600}~0.5 in LB. Bars correspond to the average of three biological replicates with circles corresponding to individual data points and error bars representing 1 SD.

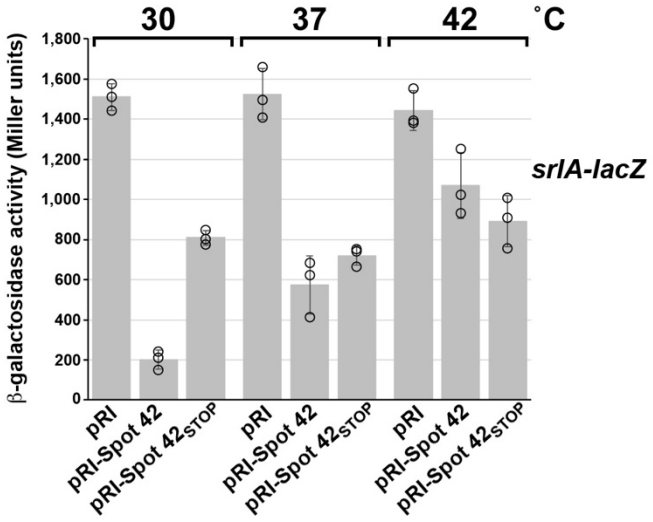


Fig. S4. Overexpression of Spot 42 and Spot 42_{STOP} leads to different effects at different temperatures. β-galactosidase assay of Δ*spf::kan srlA-lacZ* expressing cells (GSO441) transformed with pRI, pRI-Spot 42, and pRI-Spot 42_{STOP}. The cells were grown to OD_{600}~1.0 at 30˚C, 37˚C, and 42˚C in LB supplemented with 0.2% arabinose. Bars correspond to the average of three biological replicates with circles corresponding to individual data points and error bars representing 1 SD.

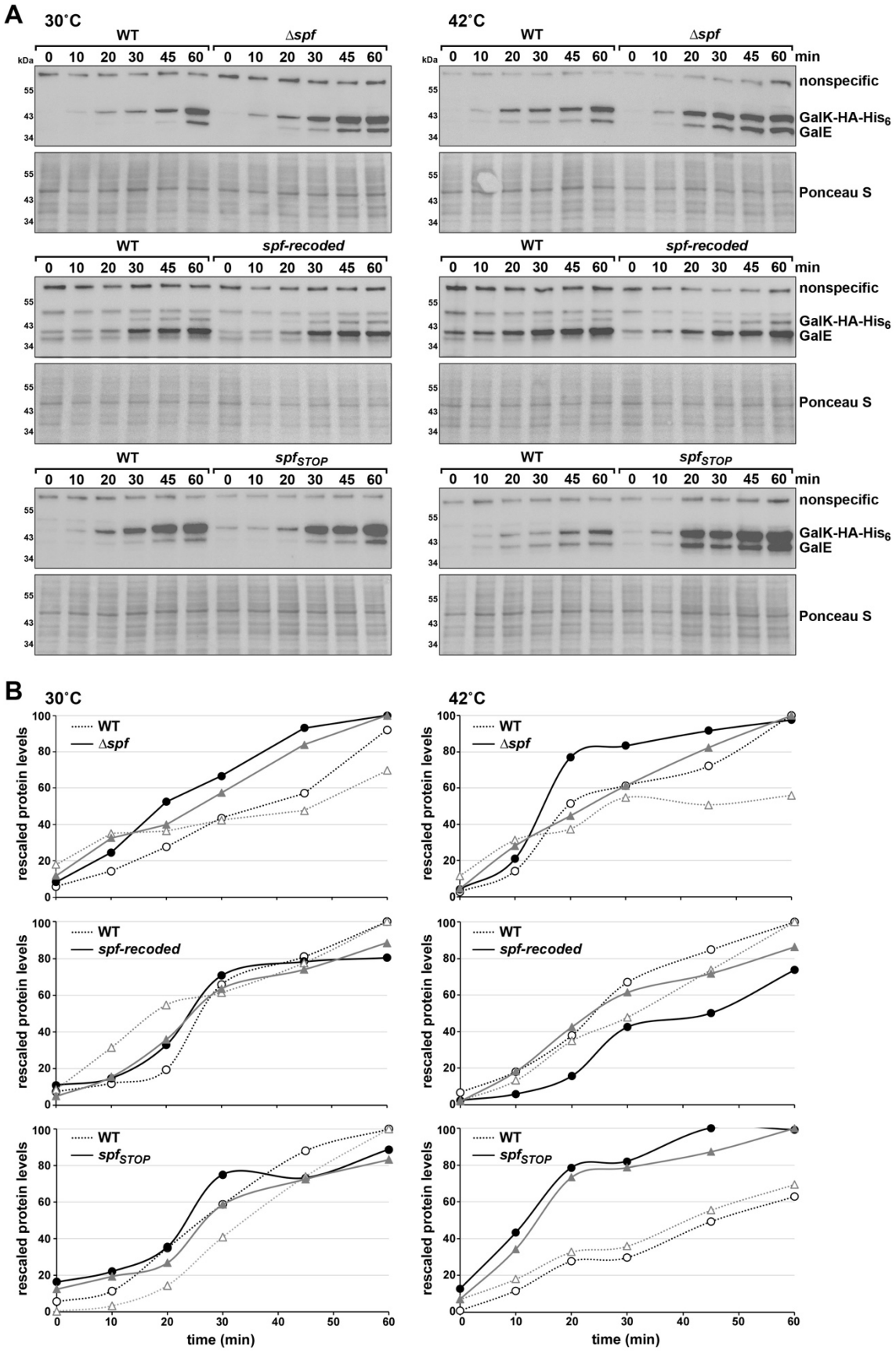

A
30˚C
42˚C
WT
Δspf
spf-recoded
spfSTOP
0 10 20 30 45 60
kDa
55
43
34
min
nonspecific
GalK-HA-His6
GalE
Ponceau S
B
rescaled protein levels
time (min)

Fig. S5. Chromosomally-encoded Spot 42 and SpfP have different effects at 30˚C and 42˚C (*A*) Immunoblot analysis of GalE levels in *galK-HA-His$_6$::kan* (GSO1057), Δ*spf galK-HA-His$_6$::kan* (GSO1060), *spf-recoded galK-HA-His$_6$::kan* (GSO1077), and *spf$_{STOP}$*::kan *galK-HA-His$_6$* (GSO1075). Blots in Fig. 6*C* were stripped and probed with a-GalE polyclonal antiserum. The stripping did not completely remove the GalK-HA-His6 signal, and some nonspecific bands are detected with the a-GalE antibodies. The Ponceau S stain documents approximately equal loading of the samples. (*B*) Quantitation of GalK-HA-His$_6$ levels in WT, Δ*spf*, *spf-recoded* and *spf$_{STOP}$* backgrounds at 30˚C and 42˚C for two independent experiments. ImageJ analysis was used to quantify relative intensities of GalK-HA-His$_6$ bands in Fig. 6*C* (black, circles) and a separate, independent experiment (gray, triangles). Quantification was rescaled to span 0-100.

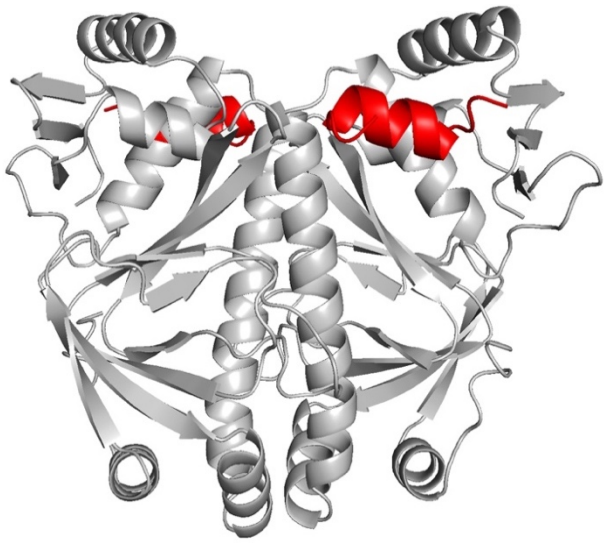


Fig. S6. Interaction between SpfP and CRP predicted by AlphaFold-Multimer (5). The prediction with the highest confidence is shown (pLDDT score of 40-60 for SpfP). The CRP dimer is in gray and SpfP is in red. CRP is oriented as in Fig. 5*C*.