Construction and validation of the Tn*5*-$P_{LtetO-1}$-*msfGFP* transposon as a tool to probe protein expression and localization

Ioannis Passaris[a,#], Wubishet M. Tadesse[a,#], Elisa Gayán[a,1,#], Abram Aertsen[a,*]

[a]Laboratory of Food Microbiology, Department of Microbial and Molecular Systems, KU Leuven. Faculty of Bioscience Engineering, Kasteelpark Arenberg 22, 3000 Leuven, Belgium.

[1]Present address: Laboratory of Food Technology, Department of Animal Production and Food Science, AgriFood Institute of Aragon (IA2), University of Zaragoza-CITA. Faculty of Veterinary, Miguel Servet 177, 50013 Zaragoza, Spain.

[#]These authors contributed equally to the manuscript

[*]Corresponding author: abram.aertsen@kuleuven.be

Abstract

In this study we report the design, construction and validation of a novel transposon aimed to systematically screen for protein localization and expression patterns in prokaryotes using fluorescence microscopy. Upon random insertion in an open reading frame in the proper frame and orientation, the transposon creates an N-terminal fluorescent protein fusion to the msfGFP reporter. Moreover, in order to examine the localization of fusion proteins whose native expression might be too low or absent, the transposon was fitted with a $P_{LtetO-1}$ promoter that makes the expression of the generated fluorescent protein fusions controllable by anhydrotetracycline. Importantly, upon flipping out the $P_{LtetO-1}$ promoter and neighboring antibiotic resistance marker, an in-frame "sandwich" msfGFP fusion is created in which the N- and C-terminal portions of the targeted protein are again controlled by its native promoter.

Keywords: transposon mutagenesis, fluorescent protein fusion, cellular localization, anhydrotetracycline inducible promoter.

1. Introduction

Insights into cellular organization drive a better understanding of life's design principles and enable synthetic biology approaches to construct purposeful cellular chassis (Long *et al.*, 2005; Murat *et al.*, 2010; Rudner and Losick, 2010; Nikel *et al.*, 2018). While the structure, activity and expression patterns of proteins are commonly studied, their potentially dynamic cellular localization is often neglected, even though this aspect often forms an integral part of a protein's role in cellular organization.

Among the approaches for studying dynamic protein localization is the use of live cell biology and fluorescent protein fusions (Lee *et al.*, 2013; Hashimoto *et al.*, 2016). Together with the advent of monomeric, fast-folding and bright fluorescent proteins with different spectral properties, this approach has allowed the accurate monitoring of dynamic expression and localization patterns of proteins and protein complexes in both prokaryotic and eukaryotic organisms (Breker *et al.*, 2013; Kuwada *et al.*, 2014; Chong *et al.*, 2015). However, while for a few well-studied model organisms ordered libraries have been made (Kitagawa *et al.*, 2006; Werner *et al.*, 2009; Taniguchi *et al.*, 2010; Breker *et al.*, 2013), the tools for genome-wide screening of potentially interesting protein localizations are still limited (Gregory *et al.*, 2010; Passaris *et al.*, 2014; Passaris *et al.*, 2018). Especially when trying to study proteins as they are (sometimes lowly) expressed from their native chromosomal locus.

In this report, we therefore designed, constructed and validated a new transposon for use in bacteria with the aim of creating and screening random fluorescent protein fusions that can be conditionally overexpressed with a $P_{LtetO-1}$ promoter in order to overrule the potentially low or absent native expression of the target gene.

2. Material and methods

2.1. Strains and growth conditions

Bacterial strains and plasmids used throughout this study are listed in Table 1, while all the relevant primers together with their purpose are listed in Table 2. For routine culturing of bacteria, Lysogeny Broth (LB) medium (Sambrook and Russell, 2001) was used either as broth or as agar plates after the addition of 1.5% bacteriological agar (LAB M, Lancashire, UK). For selecting auxotrophic conjugants, AB minimal medium (Clark and Maaløe, 1967) containing 0.2% D-glucose (AB-glucose; Sigma-Aldrich, St. Louis, MO, USA) and 10 μg/ml of thiamine (Acros Organics, Morris Plains, NJ, USA) was used. For fluorescence microscopy, cells were grown in AB medium supplemented with 0.2% D-glycerol (AB-glycerol; Acros Organics), 0.5% cas-amino acids (LAB M), 10 μg/ml of thiamine and 25 μg/ml of uridine (Sigma-Aldrich). When appropriate, the medium was supplemented with a final concentration of 100 μg/ml of ampicillin (Ap; Thermo Fisher Scientific, Waltham, MA, USA), 50 μg/ml of kanamycin (Km; Panreac-AppliChem, Darmstadt, Germany), 200 ng/ml of anhydrotetracycline (aTc to induce the $P_{LtetO-1}$; Sanbio B.V., Uden, The Netherlands) and/or 1 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG to induce the P_{lac}; Acros Organics).

Cultures were obtained by inoculating test tubes containing 4 ml of medium with a single colony, which were then incubated for 16–18 h at 37°C under well-aerated conditions (200 rpm on an orbital shaker) to reach stationary phase. Exponential phase cultures were in turn prepared by diluting stationary phase cultures 1/100 in pre-warmed tubes and allowing further incubation until an optical density at 630 nm (OD_{630}) of 0.4–0.6 was reached.

Table 1. Strains and plasmids used in this investigation.

Name	Relevant characteristics	Source or reference
Escherichia coli		
S17-1 *λpir*	F^- Tp^R Sm^R *recA1*, *thiE1*, *pro-82*, *hsdR17*-M+RP4-2 (*Tc:Mu: Km Tn7 λpir).* Donor strain used for Tn*5*-based transposon mutagenesis.	Provided by Prof. V. de Lorenzo (CNB-CSIC, Spain)
K-12 MG1655	F^- λ^- *ilvG*$^-$ *rfb*-50 *rph-1*.	Blattner *et al.*, 1997
*lacZ::*Tn*5*-$P_{LtetO-1}$-*msfGFP::lacZ*	MG1655 carrying the Tn*5*-$P_{LtetO-1}$-*msfGFP* transposon after the start codon of *lacZ*, creating an N-terminal msfGFP::LacZ fusion protein under aTc control.	This work
lacZ::msfGFP::lacZ	MG1655 carrying the *lacZ::msfGFP::lacZ* "sandwich" construct under native P_{lac} control, created after recombineering *frt* sites of *lacZ::*Tn*5*-$P_{LtetO-1}$-*msfGFP::lacZ*.	This work
Salmonella Typhimurium ATCC 14028s	Acceptor strain used for Tn*5*-based transposon mutagenesis.	Jarvik *et al.*, 2010
Plasmids		
pKD46	Expression of λ red genes under the control of *araBAD* promoter, temperature sensitive, Ap resistant.	Datsenko and Wanner, 2000
pKD4	Template for *frt-kan-frt* cassette, Ap and Km resistant.	Datsenko and Wanner, 2000
pCP20	Expression of Flp recombinase, temperature sensitive, Ap and chloramphenicol resistant.	Cherepanov and Wackernagel, 1995
pBAM1-GFP	Containing the Tn*5*-GFP transposon, Ap and Km resistant.	Provided by Prof. V. de Lorenzo (CNB-CSIC, Spain)
pBAM1-Tn*5*-$P_{LtetO-1}$-*msfGFP*	Containing the Tn*5*-$P_{LtetO-1}$-*msfGFP* transposon.	This work

Table 2. Primers used in this investigation.

Primer name	Sequence (5'➔3')[a]	Use
kan_BamHI_Rev	TACGGGATCC**GAAGAACTCC AGCATGAGAT**	Amplification of *kan* cassette from pKD4
kan_KpnI_Fw	TACGGGATCC**GAATAGGAAC TTCAAGATCC**	Amplification of *kan* cassette from pKD4
linker1	TTTCTGCTCGAATTCAAGCTT CTAACGATGTACGGGGACAC ATG	Y linker
phosphorylated_linker2	TGTCCCCGTACATCGTTAGAA CTACTCGTACCATCCACAT	Y linker
Y_linker_primer	CTGCTCGAATTCAAGCTTCT	Mapping of Tn5-P$_{LtetO-1}$-*msfGFP* insertions
Tn5-P$_{LtetO-1}$-msfGFP_dwn_out	CGGATAACCACTACCTGTCC	Mapping of Tn5-P$_{LtetO-1}$-*msfGFP* insertions
lacZ::Tn5-P$_{LtetO-1}$-msfGFP::lacZ_Fw	TGTTGTGTGGAATTGTGAGCG GATAACAATTTCACACAGGAA ACAGCTATG**CTGTCTCTTAT ACACATCTC**	Construction of MG1655 *lacZ::*Tn5-P$_{LtetO-1}$-*msfGFP::lacZ*
lacZ::Tn5-P$_{LtetO-1}$-msfGFP::lacZ_Rev	TTCCCAGTCACGACGTTGTAA AACGACGGCCAGTGAATCCGT AATCATGGT**CTGTCTCTTATA CACATCTC**	Construction of MG1655 *lacZ::*Tn5-P$_{LtetO-1}$-*msfGFP::lacZ*
lacZ_Fw	GCTTGCTGCAACTCTCTCAGG	Control of MG1655 *lacZ::*Tn*5*-P$_{LtetO-1}$-*msfGFP::lacZ* insertion and sequencing
lacZ_Rev	GCAGCCCGAGTTTGTCAGAA	Control of MG1655 *lacZ::*Tn*5*-P$_{LtetO-1}$-*msfGFP::lacZ* insertion and sequencing
tetR_int_Fw	AGACCCACTTTCACATTTAAG	Sequencing of Tn5-P$_{LtetO-1}$-*msfGFP*
kan_Fw	CTCTCAAATTTATGAATCTA	Sequencing of Tn5-P$_{LtetO-1}$-*msfGFP*
kan_int_Fw	CCCGCTTCAGTGACAACGTC	Sequencing of Tn5-P$_{LtetO-1}$-*msfGFP*
msfGFP_int_Rev	ACCTTCCGGCATTGCAGATT	Sequencing of Tn5-P$_{LtetO-1}$-*msfGFP*

[a]When relevant, primer attachment sites are shown in bold.

2.2. Construction of the pBAM1 plasmid harboring the Tn*5*-P$_{LtetO-1}$-*msfGFP* transposon

The Tn*5*-P$_{LtetO-1}$-*msfGFP* transposon was designed *in silico* and ordered from a DNA synthesis company (GenScript, Piscataway, NJ, USA). Several restriction enzyme sites were incorporated in the sequence for cloning the kanamycin resistance marker (encoded by the *npt* gene, but further referred to as *kan*; Datsenko and Wanner (2000)) afterwards. This construct was delivered in the pCC1-BAC vector, and the Tn*5*-P$_{LtetO-1}$-*msfGFP*-like transposon (lacking the *kan* gene) was cut out of this vector using PvuII (Thermo Fisher Scientific) and ligated (T4 DNA ligase; Thermo Fisher Scientific) in the backbone of a PvuII-digested pBAM1-GFP plasmid. The ligation mixture was electroporated into exponential phase cells of *Escherichia coli* S17-1 *λpir*, and several resulting clones were PCR verified and sequenced (Macrogen, Amsterdam, The Netherlands). Afterwards, the plasmid of a sequence verified clone was extracted (GeneJET Plasmid Miniprep Kit; Thermo Fisher Scientific) and then double digested with BamHI and KpnI (Thermo Fisher Scientific), while the *kan* cassette of pKD4 (Datsenko and Wanner, 2000) was PCR amplified using a forward primer with a 5' KpnI restriction site and a reverse primer with a 5' BamHI restriction site (Table 2). The latter fragment was directionally ligated in the pBAM1-Tn*5*-P$_{LtetO-1}$-*msfGFP*-like transposon and after electroporation of the ligation mixture into *E. coli* S17-1 *λpir*, cells were plated on LB agar containing Km. Resistant clones were subsequently verified by PCR and sequencing. The complete sequence of the Tn*5*-P$_{LtetO-1}$-*msfGFP* transposon can be found in the Supplementary Material. The *msfGFP* variant used was identical to the one used by Ke *et al.* (2016).

2.3. Transposon mutagenesis by conjugation

The suicide delivery of the Tn*5*-$P_{LtetO-1}$-*msfGFP* transposon was accomplished through mating of the donor and acceptor strain (Martínez-García *et al.*, 2011). More specifically, the *E. coli* S17-1 *λpir* pBAM1-Tn*5*-$P_{LtetO-1}$-*msfGFP* donor strain, which harbors both the π protein and the RP4 conjugative machinery in its chromosome (Miller and Mekalanos, 1988), and the *Salmonella* Typhimurium ATCC 14028s acceptor strain were grown overnight in LB with the appropriate antibiotics. Cells were then washed with 10 mM $MgSO_4$ (Sigma-Aldrich) and four times concentrated. Next, an aliquot of 100 µl of each donor and acceptor cell suspension were thoroughly mixed in 5 ml of 10 mM $MgSO_4$ and applied onto a filter disk (0.45 µm pore, 47 mm diameter; Pall Corporation, Port Washington, NY, USA). The filter was subsequently incubated on an LB agar plate for 2–4 h at 30°C, after which it was transferred to 5 ml of a 10 mM $MgSO_4$ solution and intensely vortexed to resuspend the cells. Finally, cells were plated out on AB-glucose supplemented with Km to select for Tn*5*-$P_{LtetO-1}$-*msfGFP* transposon mutants while counterselecting for the donor strain. The occurrence of false positive plasmid integrants was checked through streaking out on LB plates containing Ap (backbone marker of pBAM1) and was consistently found to be below 10% of the clones.

2.4. Mapping of transposon insertions

Mapping of transposon insertion sites was performed in analogy with the method used by Kwon and Ricke (2000). First, a mixture of equimolar concentration of linker1 and phosphorylated linker2 (Table 2) was heated at 95°C for 2 min and then slowly cooled down to obtain the annealed Y linker. Genomic DNA of transposon mutants was purified via phenol-chloroform extraction (Wilson, 2001) and completely digested with NlaIII (Thermo Fisher Scientific). The digested DNA was purified (GeneJET PCR

Purification Kit; Thermo Fisher Scientific), and approximately 40 μg of DNA were ligated to 1 μg of the Y linker. Subsequently, the flanking region of the transposon insertion site was amplified using a primer specific to the transposon and a primer specific to the Y linker (Table 2) and sequenced. The exact position of the transposon was determined by the NCBI Nucleotide Basic Local Alignment Search Tool (BLASTn; https:// blast.ncbi.nlm.nih.gov) using as a reference the *Salmonella* Typhimurium ATCC 14028s genome published at the GenBank database (accession number CP001363 for the bacterial chromosome and CP001362 for the pSLT plasmid (Jarvik *et al.*, 2010)).

2.5. Recombining (flipping) the *frt* flanked *tetR-kan*-P$_{LtetO-1}$ cassette

To recombine (flip) *frt* sites and thus remove the *tetR-kan*-P$_{LtetO-1}$ cassette from Tn*5*-P$_{LtetO-1}$-*msfGFP* transposon mutants, cells were first equipped with the temperature-sensitive pCP20 plasmid, which constitutively expresses the Flp recombinase (Cherepanov and Wackernagel, 1995). Afterwards, transformants were plated on LB agar plates at 37°C without antibiotic selection to cure the pCP20 plasmid. Resulting clones were examined for the loss of the *frt* flanked *tetR-kan*-P$_{LtetO-1}$ cassette and the pCP20 plasmid by PCR and Ap sensitivity, respectively.

2.6. Construction of the *E. coli lacZ::*Tn*5*-P$_{LtetO-1}$-*msfGFP::lacZ* reporter strain

For construction of the *E. coli* MG1655 *lacZ::*Tn*5*-P$_{LtetO-1}$-*msfGFP::lacZ* strain, the Tn*5*-P$_{LtetO-1}$-*msfGFP* transposon was recombineered in frame after the start codon of the *lacZ* gene in the chromosome of a pKD46 equipped *E. coli* MG1655 strain (Datsenko and Wanner, 2000), using a PCR product generated by the primers included in Table 2. Correct integration was further verified by PCR and sequencing. For creating the

190 *lacZ::msfGFP::lacZ* "sandwich" fusion, the *tetR-kan*-$P_{LtetO-1}$ cassette was flipped using
191 pCP20-borne Flp recombinase (Cherepanov and Wackernagel, 1995) as described
192 above.
193
194 **2.7. Fluorescence microscopy**
195 Fluorescence microscopy experiments were performed with a Ti-Eclipse inverted
196 microscope (Nikon, Champigny-sur-Marne, France) equipped with a 60× Plan Apo λ oil
197 objective, a TI-CT-E motorized condenser and a Nikon DS-Qi2 camera. A SpecraX
198 LED illuminator (Lumencor, Beaverton, USA) was used as a light source. GFP was
199 imaged using a triple excitation filter (Ex 473/30) and an emission filter (Em 520/35).
200 For imaging, cells were washed twice in an equal volume of a 0.85% KCl solution
201 (Sigma-Aldrich) and then placed in 0.85% KCl agarose pads and a cover glass,
202 essentially as described previously (Cenens *et al.*, 2013). Images were acquired using
203 NIS-Elements AR (Ver. 4.51; Nikon) and resulting pictures were further handled with
204 the open source software ImageJ (National Institutes of Health, USA;
205 http://rsbweb.nih.gov/ij/). For the images that need to be compared with each other,
206 identical acquisition parameters and processing steps were applied. Processing of the
207 images was limited to background subtraction and adjustment of brightness and contrast
208 values of the fluorescent channel. Figures show representative images from at least
209 three replicates performed in different working days.
210
211 **2.8. Evaluation of LacZ activity**
212 LacZ production in *E. coli* MG1655 and its derivatives in response to aTc or IPTG was
213 examined by blue color formation on X-gal plates. More specifically, a volume of 100
214 µl of a stationary phase culture grown in LB was plated into 15 ml of LB soft agar

(0.7%) containing 100 mM of X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside; VWR, Radnor, PA, USA). After soft agar solidification, a 5 µl drop of aTc (10 µg/ml) or IPTG (100 mM) was dropped on the center of the plate. Pictures were taken after overnight incubation at 37ºC. LacZ activity was obtained from three replicates performed in different working days.

3. Results and discussion

3.1 Design of the Tn*5*-P$_{LtetO-1}$-*msfGFP* transposon

The design of the Tn*5*-P$_{LtetO-1}$-*msfGFP* transposon is shown in Fig. 1. Insertion of this transposon in the correct reading frame of a random gene in the chromosome or plasmid of a target bacterium enables (*i*) the generation of a 5' translational fusion of a fluorescent reporter gene to the downstream part of the gene it is inserted in, and (*ii*) conditional P$_{LtetO-1}$ expression of the corresponding fluorescent fusion protein so that even lowly or sporadically expressed proteins can be probed for their cellular localization. Moreover, (*iii*) flipping out the conditional promoter together with the antibiotic resistance marker used to select for insertion of the transposon, results in a "sandwich" fusion construct in which the reporter gene ends up in-frame with the 5' and 3' ends of the gene it resides in.

In more detail, the fluorescent reporter gene used in this transposon was *msfGFP* due to the superior characteristics of the corresponding msfGFP protein with respect to brightness, maturation time, periplasmic folding and lack of oligomerization tendency (Ke *et al.*, 2016). In front of the *msfGFP* gene, the synthetic P$_{LtetO-1}$ promoter was placed because of its high dynamic range (up to 5,000-fold) and thus its capability to drive high expression of downstream genes (Lutz and Bujard, 1997). This promoter is tightly repressible by TetR and repression can be alleviated through the addition of aTc, which is a tetracycline analog that binds TetR with a *ca*. 30-fold higher binding constant compared to tetracycline and thus allows it to operate at very low concentrations (Lederer *et al.*, 1996; Berens and Hillen, 2003). This feature makes the presence of the TetA efflux pump unnecessary, as nanomolar concentrations of aTc do not inhibit growth, and only the presence of the TetR encoding gene on the transposon is required (Lutz and Bujard, 1997; Köstner *et al.*, 2006). The latter *tetR* gene was placed upstream

of the $P_{LtetO-1}$ promoter and is transcribed in the opposite orientation from the constitutive P_{N25} promoter (Lutz and Bujard, 1997) (Fig. 1), which normally controls expression of the early genes of phage T5 (Deuschle *et al.*, 1986). Transcription termination is accomplished through incorporation of an *rrnB* T1 terminator downstream of the *tetR* gene (Hartvig and Christiansen, 1996; Lutz and Bujard, 1997) (Supplementary Material). The kanamycin antibiotic resistance cassette (*i.e.* neomycin phosphotransferase encoded by *npt* and further referred to as *kan*; Datsenko and Wanner, 2000), which is located between the *tetR* gene and the $P_{LtetO-1}$ promoter and transcribed in the same direction as the *tetR* gene, allows selection of transposon insertion mutants.

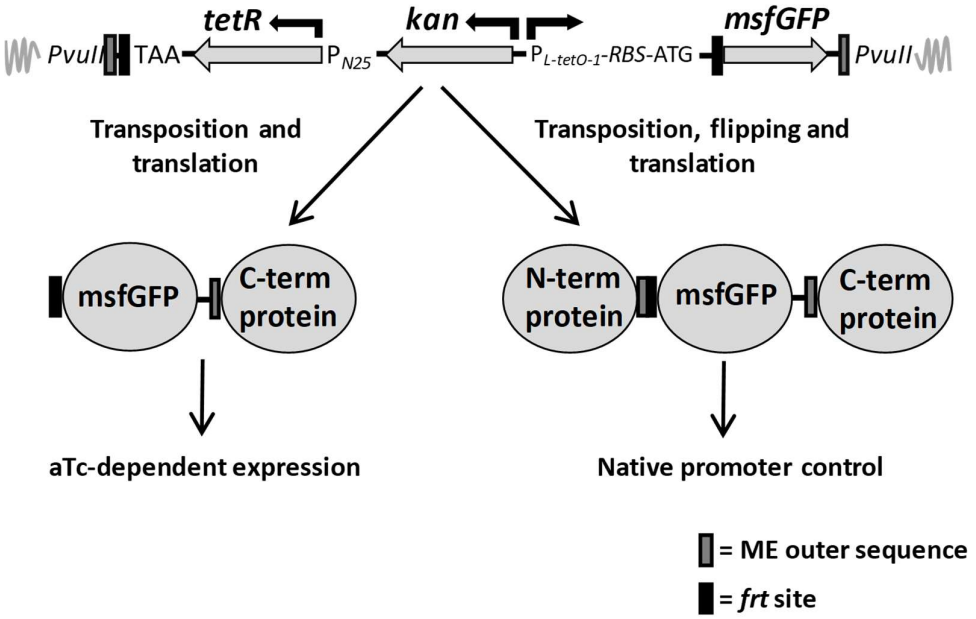

Figure 1: Scheme of the Tn*5*-$P_{LtetO-1}$-*msfGFP* transposon as integrated in the pBAM1 plasmid and its corresponding functionalities. Transposition of Tn*5*-$P_{LtetO-1}$-*msfGFP* in the correct reading frame results in N-terminal msfGFP fusions which can be induced by aTc. Flipping out the *frt*-flanked *tetR*-*kan*-$P_{LtetO-1}$ cassette yields "sandwich" fusions under native promoter control. Please note that the *msfGFP* gene lacks its natural start

and stop codon. The PvuII restriction enzyme sites were used for cloning the transposons in the pBAM1 plasmid. Images are not drawn to scale.

Importantly, the *tetR*-*kan*-$P_{LtetO-1}$ cassette is flanked by *frt* sites, in such a way that in case the original insertion generated an *msfGFP* fusion in-frame with the downstream open reading frame (ORF), flipping out this part yields a "sandwich" fusion transcribed from the native promoter of the ORF. The Tn*5*-$P_{LtetO-1}$-*msfGFP* transposon is thus capable of producing random aTc-inducible N-terminal msfGFP fusions in its unflipped configuration (with an 8-amino acid (AA) linker separating the C-terminus of the fluorescent protein from the N-terminus of the protein of interest). After recombining the *frt* sites "sandwich" fusions are created (with a 19-AA linker separating the C-terminus of the upstream part of the protein of interest from the N-terminus of msfGFP) under native promoter control. Technically, this was made feasible by the introduction of 1 or 2 bp at three different sites. First, 2 bp were added right in front of the second *frt* site to avoid a potential stop codon in that *frt* site and thus offering the possibility of N-terminal msfGFP fusions (Supplementary Material). Secondly, two additional modifications were made in order to obtain "sandwich" fusions: (*i*) 1 bp was added right in front of the first *frt* site, again to avoid a potential stop codon present in this site, and (*ii*) 2 bp were added right after the *msfGFP* gene to avoid a potential stop codon in the Mosaic End (ME) outer sequence and to align the ORFs of the 5' and 3' ends of the transposon with each other (Supplementary Material). It should be noted that when the transposon is not inserted in a gene or inserted in the wrong orientation or reading frame of a gene, the msfGFP can harbor a C-terminal nonsense peptide of which the exact length and sequence would depend on the actual insertion location.

306 In addition, the 5' end of the transposon contains stop codons in the three frames and
307 the T7 transcription terminator (Dunn *et al.*, 1983), in order to efficiently stop
308 translation and transcription originating from upstream genes and minimize possible
309 adverse effects of nonsense mRNA molecules and/or proteins (Supplementary
310 Material).
311 Finally, this transposon was cloned in the pBAM1 plasmid (Martínez-García *et al.*,
312 2011), which contains all the necessary features for transposition: the R6Kγ origin of
313 replication (dependent on presence of the π protein for its replication (Kolter *et al.*,
314 1978)), the origin of transfer region *oriT* and the *tnpA* transposase gene. The plasmid
315 can be transferred to the acceptor strain through electroporation or conjugation (bi- or
316 triparental mating). In the latter case, it can be mobilized into virtually any Gram-
317 negative or Gram-positive bacteria, if subjected to the transfer (Tra) proteins of RP4
318 (the RP4 plasmid encodes for its Tra proteins, which mediate conjugation via pilus
319 formation, DNA transfer and fusion of the outer membranes) (Trieu-Cuot *et al.*, 1987;
320 Lyras and Rood, 1998). While Tn5-based transposons do show an insertion site
321 preference for high GC content regions, they have been widely used in transposon
322 screening (Lodge et al., 1988; Green et al., 2012; Chao et al., 2016).
323

324 **3.2. Validation of the Tn*5*-P$_{LtetO-1}$-*msfGFP* and the flipped *msfGFP* fusion**
325 **functionality**

326 To validate the newly designed and constructed Tn*5*-P$_{LtetO-1}$-*msfGFP* transposon, we
327 first examined its aTc responsiveness and its constitution after flipping out the *tetR*-*kan*-
328 P$_{LtetO-1}$ cassette. For this purpose, the entire transposon was targetedly recombineered in-
329 frame immediately after the start codon of the *lacZ* gene in the *E. coli* MG1655
330 chromosome. This resulted in the *E. coli* MG1655 *lacZ::*Tn*5*-P$_{LtetO-1}$-*msfGFP::lacZ*

reporter strain in which the N-terminus of the virtually entire LacZ protein is fused to the C-terminus of msfGFP. As observed in Fig. 2, LacZ activity and fluorescence stemming from the *lacZ::*Tn*5*-P$_{LtetO-1}$-*msfGFP::lacZ* construct could indeed be induced with aTc but not with IPTG. In turn, after flipping out the the *tetR-kan*-P$_{LtetO-1}$ cassette by recombination of the *frt* sites, the resulting *lacZ::msfGFP::lacZ* in-frame "sandwhich" construct (similar to those obtained with other transposons (Gregory *et al.*, 2010; Passaris *et al.*, 2018)) could be confirmed with sequencing, and yielded LacZ activity and fluorescence after induction with IPTG instead of aTc.

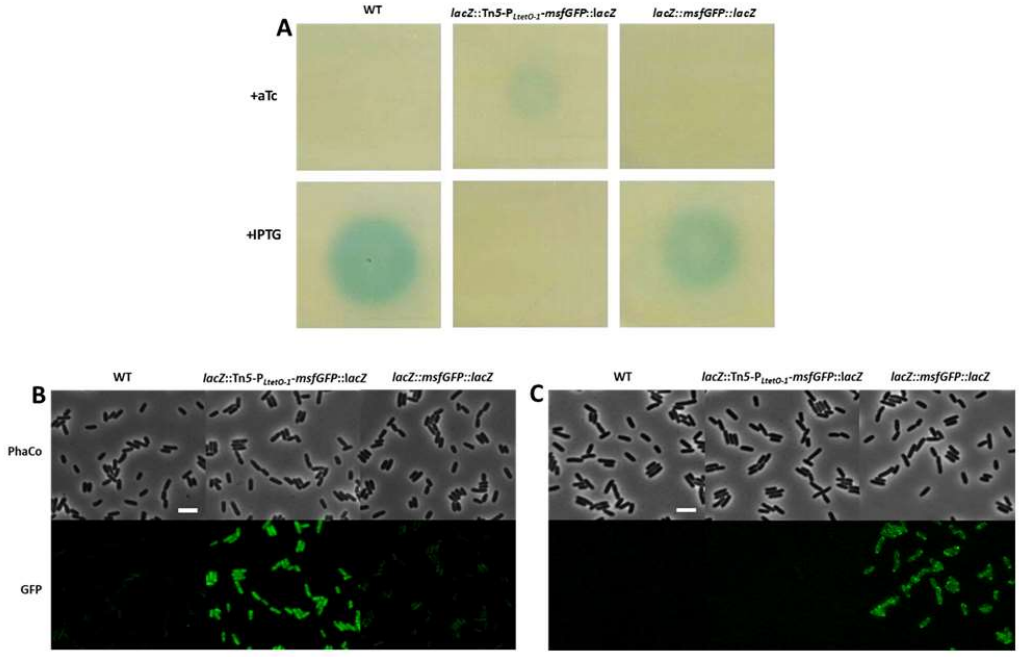

Figure 2: Representative images demonstrating (A) LacZ activity visualized on X-gal plates and cellular fluorescence signal under (B) aTc (200 ng/ml) or (C) IPTG (1 mM) induction stemming from *E. coli* MG1655 wild-type (WT; as negative control), its *lacZ::*Tn*5*-P$_{LtetO-1}$-*msfGFP::lacZ* mutant and its corresponding *lacZ::msfGFP::lacZ* "sandwich" fusion derivative. Images within a panel have been similarly acquired and adjusted, so that their brightness and contrast can be compared. PhaCo: phase-contrast channel; GFP: GFP channel. Scale bars correspond to 5 µm.

3.3. Validation of random Tn*5*-P$_{LtetO-1}$-*msfGFP* transposition

Subsequently, in order to further validate the modalities of the Tn*5*-P$_{LtetO-1}$-*msfGFP* transposon, transposon mutagenesis was carried out in *Salmonella* Typhimurium ATCC 14028s and the insertion location of the transposon was retrieved for 10 randomly picked transposon mutants. All ten insertions were found to be in different genes of the *Salmonella* genome or its associated virulence plasmid (pSLT) (Table 3), indicating that the Tn*5*-P$_{LtetO-1}$-*msfGFP* is indeed capable of random transposition. Moreover, two of the insertions (*RS02250::*Tn*5*-P$_{LtetO-1}$-*msfGFP::RS02250* and *RS21240::*Tn*5*-P$_{LtetO-1}$-*msfGFP::RS21240*) were found to be in the correct orientation and reading frame. The ATCC 14028s RS02250 protein is a putative permease, thought to reside in the inner membrane, while the *RS21240* gene codes for the putative alpha subunit of the formate dehydrogenase complex, of which its *E. coli* homolog (encoded by the *fdoG* gene) is thought to be translocated to the periplasm via the Tat-secretory pathway (Tullman-Ercek *et al.*, 2007). The two corresponding clones were retained to test whether aTc induction allowed transcription of the N-terminal msfGFP fusion proteins and whether proper "sandwich" fusions could be obtained after flipping.

As expected, a fluorescent signal for both reporters was only obtained in the presence of aTc (Fig. 3). Moreover, sequencing revealed that correct (*i.e*. both 5' and 3' in-frame) "sandwich" translational fusions were obtained after flipping of the transposon. The RS02250 "sandwich" fusion did not yield an observable fluorescence signal from its native promoter under the growth conditions we used, while the RS21240 "sandwich" fusion exhibited a constitutive diffuse fluorescent signal (Fig. 3).

370 **Table 3.** *Salmonella* Typhimurium transposon mutants with the gene and position of the

371 Tn*5*-P$_{LtetO-1}$-*msfGFP* insertion site, orientation, reading frame and, when relevant, the

372 resulting N-terminal and "sandwich" fusion protein.

Strain	Gene	Position[a]	Orientation[b]	In-frame	N-terminal / "Sandwich" fusion protein
Nr1	*STM14_RS02310*	+473	Opposite	NA	NA
Nr2	*STM14_RS00655*	+1141	Opposite	NA	NA
Nr3	*STM14_RS02250*	+978	Same	Yes	msfGFP::RS02250(AA327-AA410) / RS02250(AA1-AA326)::msfGFP::RS02250(AA327-AA410)
Nr4	*STM14_RS23550*	+2663	Same	No	NA
Nr5	Intergenic	NA	NA	NA	NA
Nr6	*STM14_RS19985*	+121	Opposite	NA	NA
Nr7	*STM14_RS15070*	+262	Opposite	NA	NA
Nr8	Intergenic	NA	NA	NA	NA
Nr9	*STM14_RS21240*	+546	Same	Yes	msfGFP::RS21240(AA183-AA804) / RS21240(AA1-AA182)::msfGFP::RS21240(AA183-AA804)
Nr10	*spvR*	+753	Opposite	NA	NA

[a]The position of the transposon gives the nucleotide after which the transposon was inserted, starting from the first base of the start codon.

[b]Orientation according to the gene it inserted in.

NA: not applicable.

373

374

375

376

377

378

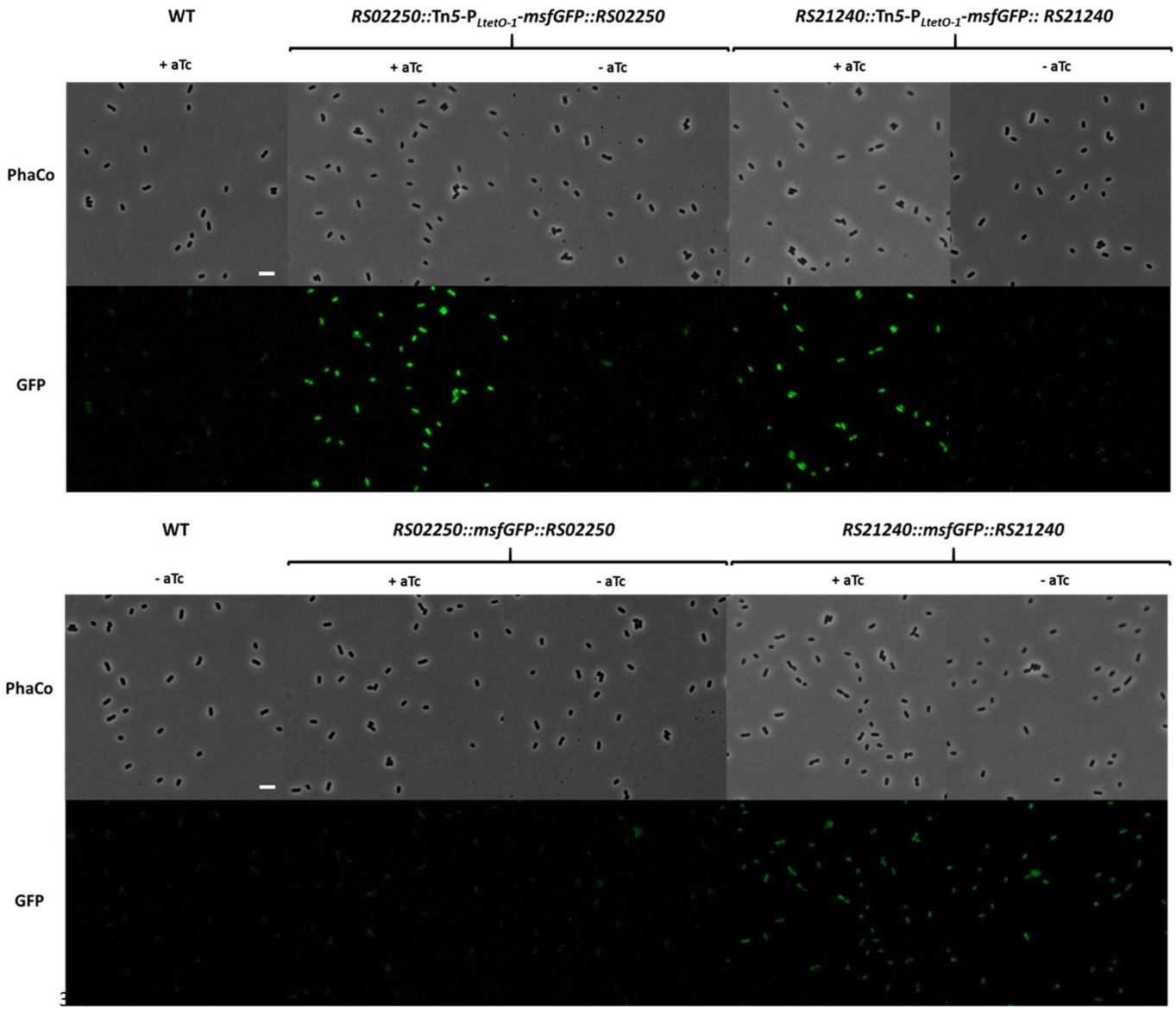

Figure 3: Representative images of fluorescence signal of *Salmonella* Typhimurium ATCC 14028s wild-type (WT; as negative control), its *RS02250::*Tn*5*-P$_{LtetO-1}$-*msfGFP::RS02250* and *RS21240::*Tn*5*-P$_{LtetO-1}$-*msfGFP::RS21240* mutants and their corresponding *RS02250::msfGFP::RS02250* and *RS21240::msfGFP::RS21240* "sandwich" fusion derivatives with and without aTc (200 ng/ml) induction. Images have been similarly acquired and adjusted, so that their brightness and contrast can be compared. PhaCo: phase-contrast channel; GFP: GFP channel. Scale bars correspond to 5 µm.

4. Conclusions

In conclusion, a new tool using Tn*5* transposon was designed, constructed and validated with the purpose of facilitating the screening for interesting protein localizations. More specifically, upon proper insertion into a random reading frame, the transposon enables the conditional (aTc-inducible) expression of an N-terminal msfGFP fusion protein that can thus be interrogated for its cellular localization through fluorescence microscopy without depending on the (perhaps too low) expression level of the native promoter. Moreover, upon flipping out the internal cassette with the conditional ($P_{LtetO-1}$) promoter, a both 5' and 3' in frame "sandwich" fusion is generated under native promoter control, allowing further fluorescence (microscopy) study of the native expression of the corresponding gene. While it cannot be excluded that the proper localization or transport of a protein suffers from the ("sandwich") fusion with a fluorescent protein and thus requires further validation, this tool nevertheless straightforwardly enables a genome-wide screen for localized proteins and their expression pattern in bacteria.

Acknowledgements

This work was supported by a doctoral fellowship from the Flemish Agency for Innovation by Science and Technology (IWT-Vlaanderen; to IP), a postdoctoral fellowship from the Research Foundation of Flanders (FWO-Vlaanderen; to EG), and grants from the KU Leuven Research Fund [IDO/10/012, DBOF/14/049] and the Research Foundation of Flanders (FWO-Vlaanderen [G058011N]).

Declarations of interest: none.

References

Berens, C., Hillen, W., 2003. Gene regulation by tetracyclines - Constraints of resistance regulation in bacteria shape TetR for application in eukaryotes. Eur. J. Biochem. 270, 3109-3121. https://doi.org/10.1046/j.1432-1033.2003.03694.x.

Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., ColladoVides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., Shao, Y., 1997. The complete genome sequence of *Escherichia coli* K-12. Science 277, 1453-1462. https://doi.org/10.1126/science.277.5331.1453.

Breker, M., Gymrek, M., Schuldiner, M., 2013. A novel single-cell screening platform reveals proteome plasticity during yeast stress responses. J. Cell Biol. 200, 839-850. 10.1083/jcb.201301120.

Cenens, W., Mebrhatu, M.T., Makumi, A., Ceyssens, P.-J., Lavigne, R., Van Houdt, R., Taddei, F., Aertsen, A., 2013. Expression of a novel P22 ORFan gene reveals the phage carrier state in *Salmonella* Typhimurium. Plos Genet. 9., e1003269. https://doi.org/10.1371/journal.pgen.1003269.

Chao, M. C., Abel, S., Davis, B. M., Waldor, M. K., 2016. The design and analysis of transposon insertion sequencing experiments. Nat. Rev. Microbiol. 14, 119-128. https://doi:10.1038/nrmicro.2015.7.

Cherepanov, P.P., 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. http://doi.org/10.1016/0378-1119(95)00193-A.

Chong, Y.T., Koh, J.L., Friesen, H., Duffy, S.K., Cox, M.J., Moses, A., Moffat, J., Boone, C., Andrews, B.J., 2015. Yeast proteome dynamics from single cell

463 imaging and automated analysis. Cell 161, 1413-1424.
464 http://doi.org/10.1016/j.cell.2015.04.051.
465 Clark, D.J., Maaløe, O., 1967. DNA replication and the division cycle in *Escherichia*
466 *coli*. J. Mol. Biol. 23, 99-112. https://doi.org/10.1016/S0022-2836(67)80070-6.
467 Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in
468 *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. U.S.A. 97,
469 6640-6645. http://doi.org/10.1073/pnas.120163297.
470 Deuschle, U., Kammerer, W., Gentz, R., Bujard, H., 1986. Promoters of *Escherichia*
471 *coli*: a hierarchy of in vivo strength indicates alternate structures. EMBO J. 5,
472 2987-2994. https://doi.org/10.1002/j.1460-2075.1986.tb04596.x.
473 Dunn, J.J., Studier, F.W., Gottesman, M., 1983. Complete nucleotide sequence of
474 bacteriophage T7 DNA and the locations of T7 genetic elements. J. Mol. Biol.
475 166, 477-535. https://doi.org/10.1016/S0022-2836(83)80282-4.
476 Green, B., Bouchier, C., Fairhead, C., Craig, N. L., Cormak, B. P., 2012. Insertion site
477 preference of Mu, Tn5, and Tn7 transposons. Mobile DNA 3, 3.
478 https://doi.org/10.1186/1759-8753-3-3.
479 Gregory, J.A., Becker, E.C., Jung, J., Tuwatananurak, I., Pogliano, K., 2010.
480 Transposon assisted gene insertion technology (TAGIT): A tool for generating
481 fluorescent fusion proteins. Plos One 5, e8731.
482 https://doi.org/10.1371/journal.pone.0008731.
483 Hartvig, L., Christiansen, J., 1996. Intrinsic termination of T7 RNA polymerase
484 mediated by either RNA or DNA. EMBO J. 15, 4767-4774.
485 https://doi.org/10.1002/j.1460-2075.1996.tb00854.x.
486 Hashimoto, M., Nozoe, T., Nakaoka, H., Okura, R., Akiyoshi, S., Kaneko, K., Kussell,
487 E., Wakamoto, Y., 2016. Noise-driven growth rate gain in clonal cellular

populations. Proc. Natl. Acad. Sci. U.S.A. 113, 3251-3256. https://doi.org/10.1073/pnas.1519412113.

Jarvik, T., Smillie, C., Groisman, E.A., Ochman, H., 2010. Short-term signatures of evolutionary change in the *Salmonella enterica* serovar Typhimurium 14028 genome. J. Bacteriol. 192, 560-567. https://doi.org/10.1128/jb.01233-09.

Ke, N., Landgraf, D., Paulsson, J., Berkmen, M., 2016. Visualization of periplasmic and cytoplasmic proteins with a self-labeling protein tag. J. Bacteriol. 198, 1035-1043. https://doi.org/10.1128/jb.00864-15.

Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., Mori, H., 2006. Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research. DNA Res. 12, 291-299. https://doi.org/10.1093/dnares/dsi012

Kolter, R., Inuzuka, M., Helinski, D.R., 1978. Trans-complementation-dependent replication of a low molecular weight origin fragment from plasmid R6K. Cell 15, 1199-1208. https://doi.org/10.1016/0092-8674(78)90046-6.

Köstner, M., Schmidt, B., Bertram, R., Hillen, W., 2006. Generating tetracycline-inducible auxotrophy in *Escherichia coli* and *Salmonella enterica* serovar Typhimurium by using an insertion element and a hyperactive transposase. Appl. Environ. Microbiol. 72, 717-4725. https://doi.org/10.1128/aem.00492-06.

Kuwada, N.J., Traxler, B., Wiggins, P.A., 2014. Genome-scale quantitative characterization of bacterial protein localization dynamics throughout the cell cycle. Mol. Microbiol. 95, 64-79. https://doi.org/10.1111/mmi.12841.

Kwon, Y.M., Ricke, S.C., 2000. Efficient amplification of multiple transposon-flanking sequences. J. Microbiol. Methods 41, 195-199. https://doi.org/10.1016/S0167-7012(00)00159-7.

Lederer, T., Kintrup, M., Takahashi, M., Sum, P.-E., Ellestad, G.A., Hillen, W., 1996. Tetracycline analogs affecting binding to Tn10-encoded Tet repressor trigger the same mechanism of induction. Biochemistry 35, 7439-7446. https://doi.org/10.1021/bi952683e.

Lee, S., Lim, W.A., Thorn, K.S., 2013. Improved blue, green, and red fluorescent protein tagging vectors for *S. cerevisiae*. Plos One 8, e67902. https://doi.org/10.1371/journal.pone.0067902.

Lodge, J. K., Weston-Hafer, K., Berg, D. E., 1988. Transposon Tn5 target specificity: preference for insertion at G/C pairs. Genetics 120, 645-650

Long, M.S., Jones, C.D., Helfrich, M.R., Mangeney-Slavin, L.K., Keating, C.D., 2005. Dynamic microcompartmentation in synthetic cells. Proc. Natl. Acad. Sci. U.S.A. 102, 5920-5925. https://doi.org/10.1073/pnas.0409333102.

Lutz, R., Bujard, H., 1997. 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. https://doi.org/10.1093/nar/25.6.1203.

Lyras, D., Rood, J.I., 1998. Conjugative transfer of RP4-oriT shuttle vectors from *Escherichia coli* to *Clostridium perfringens*. Plasmid 39, 160-164. https://doi.org/10.1006/plas.1997.1325.

Martínez-García, E., Calles, B., Arévalo-Rodríguez, M., de Lorenzo, V., 2011. pBAM1: an all-synthetic genetic tool for analysis and construction of complex bacterial phenotypes. BMC Microbiol. 11, 38. https://doi.org/10.1186/1471-2180-11-38.

536 Miller, V.L., Mekalanos, J.J., 1988. A novel suicide vector and its use in construction of
537 insertion mutations: osmoregulation of outer membrane proteins and virulence
538 determinants in *Vibrio cholerae* requires *toxR*. J. Bacteriol. 170, 2575-2583.
539 https://doi.org/10.1128/jb.170.6.2575-2583.1988.
540 Murat, D., Byrne, M., Komeili, A., 2010. Cell biology of prokaryotic organelles. Cold
541 Spring Harb. Perspect. Biol. 2, a000422.
542 https://doi.org/10.1101/cshperspect.a000422.
543 Nikel, P.I., de Lorenzo, V., 2018. *Pseudomonas putida* as a functional chassis for
544 industrial biocatalysis: From native biochemistry to trans-metabolism. Metab.
545 Eng. 50, 142-155. https://doi.org/10.1016/j.ymben.2018.05.005.
546 Passaris, I., Ghosh, A., Cenens, W., Michiels, C.W., Lammertyn, J., Aertsen, A., 2014.
547 Isolation and validation of an endogenous fluorescent nucleoid reporter in
548 *Salmonella* Typhimurium. Plos One 9, e93785.
549 https://doi.org/10.1371/journal.pone.0093785.
550 Passaris, I., Cambré, A., Govers, S.K., Aertsen, A., 2018. Bimodal expression of tnhe
551 *Salmonella* Typhimurium *spv* operon. Genetics 210, 621-635.
552 https://doi.org/10.1534/genetics.118.300822.
553 Rudner, D.Z., Losick, R., 2010. Protein subcellular localization in bacteria. Cold Spring
554 Harb. Perspect. Biol. 2, a000307. https://doi.org/10.1101/cshperspect.a000307.
555 Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual. Cold
556 Spring Harbor Lab Press, New York.
557 Taniguchi, Y., Choi, P.J., Li, G.-W., Chen, H., Babu, M., Hearn, J., Emili, A., Xie, X.S.,
558 2010. Quantifying *E. coli* proteome and transcriptome with single-molecule
559 sensitivity in single cells. Science 329, 533-538.
560 https://doi.org/10.1126/science.1188308.

Trieu-Cuot, P., Carlier, C., Martin, P., Courvalin, P., 1987. Plasmid transfer by conjugation from *Escherichia coli* to Gram-positive bacteria. FEMS Microbiol Lett. 48, 289-294. https://doi.org/10.1111/j.1574-6968.1987.tb02558.x.

Tullman-Ercek, D., De Lisa, M.P., Kawarasaki, Y., Iranpour, P., Ribnicky, B., Palmer, T., Georgiou, G., 2007. Export pathway selectivity of *Escherichia coli* twin arginine translocation signal peptides. J. Biol. Chem. 282, 8309-8316. https://doi.org/10.1074/jbc.M610507200.

Werner, J.N., Chen, E.Y., Guberman, J.M., Zippilli, A.R., Irgon, J.J., Gitai, Z., 2009. Quantitative genome-scale analysis of protein localization in an asymmetric bacterium. Proc. Natl. Acad. Sci. U.S.A. 106, 7858-7863. https://doi.org/10.1073/pnas.0901781106.

Wilson, K., 2001. Preparation of genomic DNA from bacteria. Curr. Protoc. Mol. Biol. 56, 2.4.1-2.4.5. https://doi.org/10.1002/0471142727.mb0204s56.