Flagellar region 3b supports strong expression of integrated DNA and the highest chromosomal integration efficiency of the Escherichia coli flagellar regions

Mario Juhas* and James W. Ajioka
Department of Pathology, University of Cambridge, Tennis Court Road, CB2 1QP Cambridge, UK.

Summary
The Gram-negative bacterium Escherichia coli is routinely used as the chassis for a variety of biotechnology and synthetic biology applications. Identification and analysis of reliable chromosomal integration and expression target loci is crucial for E. coli engineering. Chromosomal loci differ significantly in their ability to support integration and expression of the integrated genetic circuits. In this study, we investigate E. coli K12 MG1655 flagellar regions 2 and 3b. Integration of the genetic circuit into seven and nine highly conserved genes of the flagellar regions 2 (motA, motB, fliH, fliE, cheW, cheY and cheZ) and 3b (fliE, F, G, J, K, L, M, P, R), respectively, showed significant variation in their ability to support chromosomal integration and expression of the integrated genetic circuit. While not reducing the growth of the engineered strains, the integrations into all 16 target sites led to the loss of motility. In addition to high expression, the flagellar region 3b supports the highest efficiency of integration of all E. coli K12 MG1655 flagellar regions and is therefore potentially the most suitable for the integration of synthetic genetic circuits.

Introduction
The Gram-negative model bacterium Escherichia coli is capable of thriving in a wide variety of environments (Juhas et al., 2014a). Easily amenable to genetic manipulations, E. coli strain K-12 is also important for a number of industrial applications, biomanufacturing and metabolic engineering (Ajikumar et al., 2010; Zhang et al., 2010; Clomburg and Gonzalez, 2011; Yim et al., 2011; Zhou et al., 2012). With the advent of synthetic biology, E. coli K-12 has become one of the most frequently used synthetic biology host organisms (Juhas et al., 2013; 2014a; Juhas, 2015).

Introduction of the synthetic DNA fragments into the E. coli genome by chromosomal integration has many advantages over the plasmid-borne transformation (Cunningham et al., 2009; Marcellin et al., 2010). Furthermore, integration into the chromosome could be exploited for heterologous protein expression, particularly for expression of toxic proteins in E. coli. Work on plasmids has shown that regulation of expression is tighter when the copy number is low (Anthony et al., 2004; Guan et al., 2013). The frequently used methods of the E. coli chromosomal integration include the integrase-mediated recombination between the phage attachment sites (att) (St-Pierre et al., 2013) and the λ bacteriophage Red recombinase-mediated recombination employing knock-in/knock-out (KIKO) vectors (Sabri et al., 2013), plasmid pSB1K3(FRTK) (Juhas et al., 2014b) and the yeast mitochondrial homing endonuclease I-SceI (Ublinskaya et al., 2012). Chromosomal integration target sites differ significantly in their ability to support integration and expression of the integrated genetic circuits (Juhas et al., 2014b). As the traditionally used att sites are missing in a number of industrially important E. coli strains, identification and validation of the reliable chromosomal integration target sites is crucial for E. coli engineering. Ideally, integration target sites should be well-characterized, non-essential, conserved and highly expressed (Fraser et al., 1999; Baba et al., 2006; Vora et al., 2009; Kahramanoglu et al., 2011; Juhas et al., 2014b). Genes encoding flagellar functions meet all these prerequisites (Juhas et al., 2014b). Previous analyses of the E. coli K12 MG1655 flagellar regions 3a and 1 led to the identification of only three potential integration target sites (Juhas et al., 2014b; Juhas and Ajioka, 2015). The identification and validation of alternative integration sites is crucial for the development of a robust synthetic biology toolkit (Juhas et al., 2014b).
and Ajioka, 2015). This is critical particularly for applications that require integrations of multiple genetic circuits into the chromosome. Here, we investigate the *E. coli* K12 MG1655 flagellar regions 2 and 3b. Analysis of the seven and nine highly conserved genes of the flagellar regions 2 and 3b, respectively, revealed significant variability in their suitability for integration and expression of genetic circuits. Furthermore, we show that in addition to high expression, the *E. coli* K12 MG1655 flagellar region 3b supports highest efficiency of chromosomal integration of all *E. coli* flagellar regions.

Results and discussion

Integration target loci in the *E. coli* flagellar regions 2 and 3b

Identification of the reliable chromosomal integration target loci is crucial for engineering *E. coli* cells (Sabri et al., 2013; Juhas et al., 2014b). Chromosomal integration target sites should be well-characterized, conserved, non-essential and highly expressed (Fraser et al., 1999; Baba et al., 2006; Vora et al., 2009; Kahramanoglu et al., 2011; Juhas et al., 2014b; Juhas, 2015). Genes encoding flagellar functions are considered to be among the best targets for integration of genetic circuits into the *E. coli* chromosome (Juhas et al., 2014b). Previous studies investigating *E. coli* K12 MG1655 flagellar regions 3a (Juhas et al., 2014b) and 1 (Juhas and Ajioka, 2015) led to the identification of three putative chromosomal integration target sites. Identification and validation of the alternative loci is important particularly for those biotechnology and synthetic biology applications that require integrations of multiple genetic circuits into *E. coli* chromosome.

Here, we investigate *E. coli* K12 MG1655 flagellar regions 2 and 3b. *Escherichia coli* K12 MG1655 flagellar regions 2 (Fig. 1A) and 3b (Fig. 1B) show high probability of the RNA polymerase binding. This suggests that genetic circuits integrated into these regions will be strongly transcribed. *Escherichia coli* K12 MG1655 flagellar regions 2 and 3b harbour 28 open reading frames (flhA, flhB, flhC, flhD, flhE, motA, motB, cheA, cheB,

![Fig. 1. RNA polymerase binding to *E. coli* flagellar regions 2 and 3b. Figure depicts the probability of the RNA polymerase (RNA-Pol) binding (green peaks) to the *E. coli* K12 MG1655 genome regions 1957000-1980000 (A) and 2006000-2030000 (B). The investigated *E. coli* K12 MG1655 flagellar regions 2 (1962580-1978197) and 3b (2011038-2021702) show high probability of being occupied by RNA polymerase. Figure was created by uploading the ChIP-seq RNA-Pol data (Kahramanoglu et al., 2011) to the UCSC genome browser for *E. coli* K12 MG1655.](attachment:image.png)
Table 1. Integration target loci in the *E. coli* flagellar region 2.

| Gene | Function | References |
|------|----------|------------|
| motA | Flagellar motor component | Mohawk et al., 2014; Takahashi and Ito, 2014 |
| motB | Flagellar motor component | Reboul et al., 2011; Takahashi et al., 2014 |
| flhD | Master regulator of flagellar genes | Chatterjee et al., 2009; Mitra et al., 2013 |
| cheW | Proton influx regulator via T3SS | Lee and Harshey, 2012 |
| fliE | Chemotaxis signal transduction | Cashman et al., 2013 |
| cheY | Chemotaxis response regulator, clockwise flagellar rotation | Fraiberg et al., 2015 |
| cheZ | Phosphatase, cheY dephosphorylation | Freeman et al., 2011 |

Table 2. Integration target loci in the *E. coli* flagellar region 3b.

| Gene | Function | References |
|------|----------|------------|
| fliE | Flagellar basal body component | Dyszel et al., 2010 |
| fliF | Membrane and supramembrane (MS)-ring collar protein, flagellar basal body | Ogawa et al., 2015 |
| fliG | Flagellar motor switching | Lam et al., 2012 |
| fliJ | Flagellar protein export apparatus, rotor like function | Kishikawa et al., 2013 |
| fliK | Flagellar hook-length control | Aizawa, 2012 |
| fliL | Flagellar motor output control | Partridge et al., 2015 |
| fliM | Flagellar motor energizing | Delalez et al., 2014 |
| fliP | Flagellar export apparatus | Boyd and Gober, 2001 |
| fliR | Flagellar export apparatus | Minamino and Macnab, 1999 |
of the chromosomal integrations into the flagellar regions 2 and 3b by spotting 2 μl of the normalized overnight cultures of the engineered E. coli strains and E. coli K12 MG1655 wild type in the middle of the motility agar plates (Fig. 6). The motility of all strains harbouring integrations in the investigated genes of the flagellar regions 2 (Fig. 6A) and 3b (Fig. 6B) was completely abolished.

**Integrations into flagellar regions do not have negative impact on the growth**

As integrations of the synthetic genetic circuits into the E. coli chromosome should not negatively impact cell growth, target loci cannot be located within essential genes (Juhas et al., 2011; 2012a,b; 2014a). To assess the effect of chromosomal integrations into the seven investigated genes of the flagellar regions 2 (motA, motB, flhD, cheY) and 3b (fliE, fliK, fliL, fliM) was higher (twofold to fivefold) than the average expression of the housekeeping genes (Fig. 7A). The relative transcription of fliG was not significantly different, whereas the transcription of the remaining genes was lower than the mean expression of the housekeeping genes (Fig. 7A). The transcription of the genetic circuit integrated into motA (motAi), motB (motBi), flhD (flhDi), cheW (cheWi), cheY (cheYi) and cheZ (cheZi) of the flagellar region 2 measured by RT-PCR was high at all analysed loci (Fig. 7B). From the flagellar region 2, highest expressed (8- to 11-fold higher than the housekeeping genes) was the genetic circuit integrated into motA (motAi), motB (motBi) and flhD (flhDi) (Fig. 7B). The expression at flhE (flhEi), cheW (cheWi) and cheY (cheYi) was fourfold to sixfold higher than the mean expression of the housekeeping genes (Fig. 7B). From the flagellar region 3b, highest expressed (8- to 13-fold higher than the housekeeping genes) was the genetic circuit integrated into fliJ (fliJi), fliL (fliLi) and fliR (fliRi) (Fig. 7B). The expression at the remaining loci of the flagellar region 3b was sixfold to eightfold higher than the mean expression of the housekeeping genes (Fig. 7B). Such strong expression of the genetic circuit integrated into this flagellar region is interesting, particularly when considering that

**Transcription of the flagellar regions 2 and 3b**

The relative transcription of the investigated genes of the flagellar regions 2 and 3b was measured by real-time polymerase chain reaction (RT-PCR) using arcA and rpoD as the reference housekeeping genes (Jandu et al., 2009; Minty et al., 2011). Real-time polymerase chain reaction (RT-PCR) showed that the relative expression of four genes from both analysed flagellar regions 2 (motA, motB, flhD, cheY) and 3b (fliE, fliK, fliL, fliM) was higher (twofold to fivefold) than the average expression of the housekeeping genes (Fig. 7A). The relative transcription of fliG was not significantly different, whereas the transcription of the remaining genes was lower than the mean expression of the housekeeping genes (Fig. 7A). The transcription of the genetic circuit integrated into motA (motAi), motB (motBi), flhD (flhDi), cheW (cheWi), cheY (cheYi) and cheZ (cheZi) of the flagellar region 2 and fliE (fliEi), fliF (fliFi), fliG (fliGi), fliJ (fliJi), fliK (fliKi), fliL (fliLi), fliM (fliMi), fliP (fliPi) and fliR (fliRi) of the flagellar region 3b measured by RT-PCR was high at all analysed loci (Fig. 7B). From the flagellar region 2, highest expressed (8- to 11-fold higher than the housekeeping genes) was the genetic circuit integrated into motA (motAi), motB (motBi) and flhD (flhDi) (Fig. 7B). The expression at flhE (flhEi), cheW (cheWi) and cheY (cheYi) was fourfold to sixfold higher than the mean expression of the housekeeping genes (Fig. 7B). From the flagellar region 3b, highest expressed (8- to 13-fold higher than the housekeeping genes) was the genetic circuit integrated into fliJ (fliJi), fliL (fliLi) and fliR (fliRi) (Fig. 7B). The expression at the remaining loci of the flagellar region 3b was sixfold to eightfold higher than the mean expression of the housekeeping genes (Fig. 7B). Such strong expression of the genetic circuit integrated into this flagellar region is interesting, particularly when considering that
the flagellar region 3b shows lowest probability of being occupied by RNA polymerase (Fig. 1B). This suggests that other factors might be also important for the expression of the integrated synthetic DNA and shows that empirical characterization is necessary for engineering into integration sites. Expression of the integrated genetic circuit was determined by the quantitative measurement of the green fluorescent protein (GFP) and the red fluorescent protein (mCherry) fluorescence over time with the microplate reader (FLUOstar Omega). For this, we have used plasmids pSB1A1(GFP) and pSB1A1(mCh) harbouring GFP and mCherry, respectively, regulated by the pR promoter. Both GFP and mCherry were not expressed at permissive conditions for the repressor (30°C), while the temperature shift to 42°C set off GFP and mCherry expression (Figs S4–S7).

**Conclusions**

Flagellar regions are good targets for integration of genetic circuits into the *E. coli* chromosome. The identification of reliable target loci is crucial for building a robust synthetic biology toolkit and for *E. coli* bioengineering. Furthermore, it can be exploited for tighter regulation of expression of toxic proteins in *E. coli*. In this study, we have integrated genetic circuit into 16 well-conserved open reading frames of the *E. coli* K12 MG1655 flagellar regions 2 (motA (motAi), motB (motBi), flhD (flhDi), flhE (flhEi), cheW (cheWi), cheY (cheYi). The exact positions within the target genes where the integrations occurred are highlighted with stars.
of all *E. coli* flagellar regions. Notably, the genetic circuit integrated into flagellar region 3b was also highly expressed although the probability of the RNA polymerase binding into this region is significantly lower than into other flagellar regions. This suggests that other factors might also play a role in the expression of the integrated synthetic DNA. There appears to be a weak inverse correlation between the probability of RNA polymerase binding and the expression level.

**Fig. 4. Escherichia coli flagellar region 3b integrations.** Figure shows the sequences of the integration target sites in the *E. coli* K12 MG1655 flagellar region 3b (*fliE*(fliEi), *fliF*(fliFi), *fliG*(fliGi), *fliJ*(fliJi), *fliK*(fliKi), *fliL*(fliLi), *fliM*(fliMi), *fliP*(fliPi). The exact positions within the target genes where the integrations occurred are highlighted with stars.

© 2015 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology, *Microbial Biotechnology*, 8, 726–738
## Table 3. Primers used in this study.

| Primer (Sequence 5′→3′) | Description          |
|-------------------------|----------------------|
| motArepF: motA integration primer forward | CTCCCCAATACCCCAAAAGCAATGATATGATCTGCTGGCTCTGCTTTATCGGTTGATGGCGAAATCGTATCACGAGGCAGAATTTCAGAT |
| motArepR: motA integration primer reverse | TTTTGCAGCTAAAAAGCCTAAGGGAATAAAATCCGTAAGCAGAATTGCTCGGTTTTTTAAAAGAAAAAGGACG |
| motBrepF: motB integration primer forward | ACCATGGACCGGAGGATCTGCTGGTTTATTGACTGCGATGATGGCCTTTTTTCTGGTGATGTATCACGAGGCAGAATTTCAGAT |
| motBrepR: motB integration primer reverse | GAATTTCCGCAATCGACTTTGCTCCATGCGTTTTTTCAGCTCTTCGATGTTCGGCTGCTTATTCACTCGGTTTTAAAGAAAAAGGACG |
| flhDrepF: flhD integration primer forward | ACATGGATCGTGGAAGATTGCTTATGCCGACTTTATGACTGCGATGATGGCCTTTTTTCTGGTGATGTATCACGAGGCAGAATTTCAGAT |
| flhDrepR: flhD integration primer reverse | CTTCTTCAGGCTGATTAACATCATTCAGCAAGCGTGTTGAGAGCATGATGCCGGTATGAATTTGCTCGGTTTTAAAGAAAAAGGACG |
| flhErepF: flhE integration primer forward | CAATTGGCGGCAAATAATGAGAACCTTATTAGCAATATTATTGTTTCCGCTGCTGGTGCAAGCCGGTATCACGAGGCAGAATTTCAGAT |
| flhErepR: flhE integration primer reverse | TTGTCCTTCAGCGATAATTCACAATCACTTCATTACGTTGTACCTTTAGCGGTGGAATTAACCGCCCGGTTTTAAAGAAAAAGGACG |
| cheWrepF: cheW integration primer forward | TACCCTTGGTGATGAAGAGTACGGTATTGATATCCTGAAAGTGCAGGAGATCCGTGGCTACGATCAGTATCACGAGGCAGAATTTCAGAT |
| cheWrepR: cheW integration primer reverse | TATCTAACAGCGCCATCTCTTCGCTGTTCAGCAGTTTTTCGATGTTCACCAGAATCAACATCCGGTCGGTTTTAAAGAAAAAGGACG |
| cheYrepF: cheY integration primer forward | AAAGAGCTGGGATTCAATAATGTTGAGGAAGCGGAAGATGGCGTCGACGCTCTCAATAAGTTGCAGGTATCACGAGGCAGAATTTCAGAT |
| cheYrepR: cheY integration primer reverse | AGTTTCTCAAAGATTTTGTTGAGTTTTTCCTCCAGCGTCGCGGCGGTAAATGGCTTCACCACATAGCGGTTTTAAAGAAAAAGGACG |
| cheZrepF: cheZ integration primer forward | CAGGATTTTCAGGATCTCACCGGGCAGGTCATTAAGCGGAGATGAGTATGCATTACAGGAGATCAGGCGAGGACGGAGAATTTCAGAT |
| cheZrepR: cheZ integration primer reverse | TCAAAATCCAAGACTCCACAACAAATGTCCTACCTTCGACTCGCTACTCCACTACACCCGTTTTTCTCGGTTTTAAAAGAAAAAGGACG |
| fliEreprF: fliE integration primer forward | MVCACACATCCTTCGACATGGCAATATGGATATGGATCTGCTGGCTCTGCTTTATCGGTTGATGGCGAAATCGTATCACGAGGCAGAATTTCAGAT |
| fliEreprR: fliE integration primer reverse | TTTTGCAGCTAAAAAGCCTAAGGGAATAAAATCCGTAAGCAGAATTGCTCGGTTTTTTAAAAGAAAAAGGACG |
| fliFreprF: fliF integration primer forward | ACCATGGACCGGAGGATCTGCTGGTTTATTGACTGCGATGATGGCCTTTTTTCTGGTGATGTATCACGAGGCAGAATTTCAGAT |
| fliFreprR: fliF integration primer reverse | GAATTTCCGCAATCGACTTTGCTCCATGCGTTTTTTCAGCTCTTCGATGTTCGGCTGCTTATTCACTCGGTTTTAAAGAAAAAGGACG |
| fliGreprF: fliG integration primer forward | ACCATGGACCGGAGGATCTGCTGGTTTATTGACTGCGATGATGGCCTTTTTTCTGGTGATGTATCACGAGGCAGAATTTCAGAT |
| fliGreprR: fliG integration primer reverse | GAATTTCCGCAATCGACTTTGCTCCATGCGTTTTTTCAGCTCTTCGATGTTCGGCTGCTTATTCACTCGGTTTTAAAGAAAAAGGACG |
| fliJreprF: fliJ integration primer forward | ACCATGGACCGGAGGATCTGCTGGTTTATTGACTGCGATGATGGCCTTTTTTCTGGTGATGTATCACGAGGCAGAATTTCAGAT |
| fliJreprR: fliJ integration primer reverse | GAATTTCCGCAATCGACTTTGCTCCATGCGTTTTTTCAGCTCTTCGATGTTCGGCTGCTTATTCACTCGGTTTTAAAGAAAAAGGACG |
binding to the target loci and their ability to support integration of the genetic circuit; however, this will require further investigation. Furthermore, as flagellar genes are closer to the terminal (TER) region of the *E. coli* chromosome than oriC, their copy number is approximately sixfold lower than those genes close to oriC during exponential growth. Therefore, genes nearer to oriC are also potentially interesting target loci for integration and expression of genetic circuits. Besides the modified lambda Red recombinase method used in our analysis, clustered regularly interspaced short palindromic repeats (CRISPR) and integrases could be exploited for *E. coli*

© 2015 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology, *Microbial Biotechnology*, 8, 726–738
engineering. A variety of high complexity integrase sites, such as phiC31, R4 and Bxb1, could be moved to the hotspot integration regions in the *E. coli* chromosome employing CRISPR for appending new functionalities.

Overall, the *E. coli* K12 MG1655 flagellar region 3b is the most suitable of all *E. coli* flagellar regions for integration and expression of genetic circuits. However, there is a significant variation between individual target loci. For instance, *motA* of the *E. coli* K12 MG1655 flagellar region 2 supports the second highest integration and expression efficiency of all investigated target sites in this study (Figs 5 and 7). Therefore, when considered individually, *fliJ* and *motA* appear to be the most suitable integration target loci of the analysed flagellar regions 2 and 3b.

**Experimental procedures**

**Bacterial strains, plasmids and growth conditions**

All strains and plasmids used in this study are recorded in Table 4. *Escherichia coli* was routinely grown in Luria–Bertani (LB) medium supplemented with ampicillin (100 μg ml⁻¹) or kanamycin (50 μg ml⁻¹) when required. Liquid *E. coli* cultures were cultivated on a rotatory shaker at 200 r.p.m. at 30°C, 37°C or 42°C. Plate cultures were supplemented with 1% agar (w/v) and grown for about 24 h at 30°C, 37°C or 42°C.

**DNA amplification and modification**

DNA was amplified by PCR in 50 μl of reaction volumes employing Phusion DNA polymerase (Thermo Scientific) or Dream Taq master mix kit (Thermo Scientific) according to the supplier’s instructions. Oligonucleotide primers for PCR amplifications were synthesized by Integrated DNA Technologies (IDT) and Sigma-Aldrich. DNA fragments were purified by gel electrophoresis, followed by gel extraction employing Qiaquick Gel Extraction kit (Qiagen), according to the manufacturer’s instructions. Plasmid DNA was performed with the Qiaprep Spin Miniprep kit (Qiagen), according to the supplier’s recommendations. Sequencing was performed by Source Bioscience (Cambridge, UK). A Gibson Isothermal Assembly method (Gibson et al., 2009; Merryman and Gibson, 2012) was employed to assemble DNA fragments. The original Gibson Isothermal Assembly method protocol was modified as described previously (Juhas et al., 2014b).
Fig. 7. RT-PCR analysis.
A. Relative transcription of the analysed target genes of the E. coli K12 MG1655 flagellar regions 2 (motA, motB, flhD, flhE, cheW, cheY and cheZ) and 3b (fliE, fliF, fliG, fliJ, fliK, fliM, fliP, fliR) compared with the housekeeping genes (H).
B. Transcription of the genetic circuit integrated in the investigated integration target loci of the E. coli flagellar region 2 [motA (motAi), motB (motBi), flhD (flhDi), flhE (flhEi), cheW (cheWi), cheY (cheYi) and cheZ (cheZi)] and flagellar region 3b [(fliE (fliEi), fliF (fliFi), fliG (fliGi), fliJ (fliJi), fliK (fliKi), fliL (fliLi), fliM (fliMi), fliP (fliPi) and fliR (fliRi)] relative to the transcription of the housekeeping genes (H). Bars and errors represent averages and standard errors from three experiments. H (mean transcription of the reference housekeeping genes arcA and rpoD).

Relative transcription was quantified with REST9 Software (Qiagen) employing Pfaffl method (Pfaffl et al., 2002).

Table 4. Bacterial strains and plasmids.

| Characteristics                      | Reference                        |
|--------------------------------------|----------------------------------|
| **Strains**                          |                                  |
| K12 MG1655                           | E. coli wild type                | Hayashi et al., 2006               |
| Ec:motAi                             | E. coli K12 MG1655, motA integration | This study                           |
| Ec:motBi                             | E. coli K12 MG1655, motB integration | This study                           |
| Ec:flhDi                             | E. coli K12 MG1655, flhD integration | This study                           |
| Ec:cheWi                             | E. coli K12 MG1655, cheW integration | This study                           |
| Ec:cheYi                             | E. coli K12 MG1655, cheY integration | This study                           |
| Ec:cheZi                             | E. coli K12 MG1655, cheZ integration | This study                           |
| Ec:fliEi                             | E. coli K12 MG1655 integration into fliE | This study                           |
| Ec:fliFi                             | E. coli K12 MG1655 integration into fliF | This study                           |
| Ec:fliGi                             | E. coli K12 MG1655 integration into fliG | This study                           |
| Ec:fliJi                             | E. coli K12 MG1655 integration into fliJ | This study                           |
| Ec:fliK                              | E. coli K12 MG1655 integration into fliK | This study                           |
| Ec:fliLi                             | E. coli K12 MG1655 integration into fliL | This study                           |
| Ec:fliMi                             | E. coli K12 MG1655 integration into fliM | This study                           |
| Ec:fliPi                             | E. coli K12 MG1655 integration into fliP | This study                           |
| Ec:fliRi                             | E. coli K12 MG1655 integration into fliR | This study                           |
| **Plasmids**                         |                                  |
| pCP20                                | Plasmid encoding FLP recombinase | Datsenko and Wanner, 2000            |
| pKM208                                | IPTG-induced Red recombinase system | Murphy and Campellone, 2003       |
| pSB1A1(GFP)                           | λ, promoter-controlled GFP, Amp<sup>n</sup> | Juhas et al., 2014b               |
| pSB1A1(mCh)                           | λ, promoter-controlled mCherry, Amp<sup>n</sup> | This study                           |
| pSB1K3(FRTKr)                        | λ, repressor, Kan<sup>n</sup> | Juhas et al., 2014b               |
Integration of the genetic circuit into the chromosome

Altered Hannah (Hanahan et al., 1991) and Miller and Nickoloff (1995) protocols were used to prepare the chemically competent and electro-competent E. coli cells respectively. Integrations of the genetic circuit into target open reading frames of the analysed E. coli flagellar region were carried out using method described previously (Juhas et al., 2014b). Briefly, plasmid pKM208 was transformed into the wild-type E. coli K12 MG1655 and selected on plates with ampicillin at 30°C. Escherichia coli K12 MG1655 harbouring pKM208 was inoculated into LB with ampicillin and grown at 30°C. After reaching OD_{600} of 0.2, 1 mM IPTG was added and the bacterial culture was cultured to the final OD_{600} of 0.4–0.6. Bacteria were subsequently washed and resuspended in 10% glycerol and transformed with the genetic circuit harbouring the flanking sequences of the target genes. Bacteria with chromosomal integrations were selected on plates with kanamycin at 37°C and subsequently grown at 42°C to cure out the temperature-sensitive plasmid pKM208. Chromosomal integrations were proved by PCR with flanking primers and sequencing.

GFP and mCherry fluorescence quantitation with the microplate reader

200 μl of the E. coli cultures (grown overnight and diluted to OD_{600} of 0.05) were transferred into flat-bottomed black 96 well plates (Greiner BioOne, UK). The plates with the E. coli cultures were placed into Fluostar Omega fluorimeter (BMG Labtech, UK) and incubated first at 30°C for 3 h and then at 42°C for 17 h or 7 h for GFP and mCherry fluorescence measurement respectively. GFP fluorescence was quantified with an automatically repeated protocol each 30 min using emission filter EM520, excitation filter 485-12, double orbital shaking at 200 r.p.m. and gain 1400. mCherry fluorescence was measured with an automatically repeated protocol each 30 min using emission filter EM620, excitation filter 584, double orbital shaking at 200 r.p.m. and gain 2800.

Absorbance measurement with the microplate reader

The diluted overnight E. coli cultures (OD_{600} of 0.05) were transferred into flat-bottomed clear 96 well plates (Sterlin Sero-Well, UK). The plates were then incubated in the microplate reader (Fluostar Omega, BMG Labtech, UK) at 37°C and 30°C for 24 h. Absorbance was measured each 30 min using 600 nm absorbance filter and double orbital shaking at 500 r.p.m.

RNA isolation and purification

Total RNA was isolated from 10^9 E. coli cells at mid-exponential phase with Isolate II RNA Mini Kit (Bioline) according to manufacturer’s instructions. To elute RNA from the Isolate II RNA columns, 60 μl of RNase-free H2O was used. To avoid contamination with genomic DNA, the isolated RNA was purified with TURBO DNA-free Kit (Applied Biosystems) according to supplier’s instructions.

RT-PCR

Isolated and purified RNA (1 μg) was used to synthesize cDNA using SuperScript III Reverse Transcriptase (Invitrogen) according to supplier’s instructions. Primers for RT-PCR designed with PRIMER3 Software were prepared to generate 100–150 bp long DNA sequences. Expression levels were measured using QuantiTect SYBR Green PCR Kit (Qiagen). MicroAmp Fast Optical 96-Well Reaction Plates (Applied Biosystems) with RT-PCR reactions were incubated in the 7500 Fast Real-Time PCR System (Applied Biosystems) according to manufacturer’s instructions. The relative expression was computed employing REST9 Software (Qiagen) with Pfaffl method (Pfaffl et al., 2002). The RT-PCR was performed in triplicate, and the means and standard errors were calculated.

Evaluation of motility

Motility agar plates for motility assay were made by transferring 100 ml of motility agar [composed of 0.25% Bacto-Agar (Difco), 5 g NaCl and 10 g tryptone] in the 13 cm plates and let to set overnight. Plates were then pre-warmed to 37°C and inoculated with the 2 μl of the overnight bacterial cultures normalized to OD_{600} of 1.0. Pictures were taken after incubation for 4–6 h at 37°C.

Sequence analyses

The annotated E. coli K-12 MG1655 genome from the E. coli K-12 project website (http://www.xbase.ac.uk-genome/escherichia-coli-str-k-12-substr-mg1655) was used to retrieve DNA sequences of the target loci. DNA sequencing was carried out by Source Bioscience (Cambridge, UK). The BLASTN (Altschul et al., 1990) and TBLASTX algorithms from the National Center for Biotechnology Information (NCBI) website (http://ncbi.nlm.nih.gov) and the position-specific iterated BLAST (PSI-BLAST) (Altschul et al., 1997) were used to compare DNA sequences.

Acknowledgements

We would like to thank to Dr. Kerstin Ewen for plasmid pKM208 and Dr. Gillian Fraser for constructive discussions.

Conflict of interest

There are no conflicts of interest associated with this manuscript.

References

Aizawa, S. (2012) Mystery of flIK in length control of the flagellar hook. J Bacteriol 194: 4798–4800.

Ajikumar, P.K., Xiao, W.H., Tyo, K.E., Wang, Y., Simeon, F., Leonard, E., et al. (2010) Isopenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. Science 330: 70–74.
Marcellin, E., Chen, W.Y., and Nielsen, L.K. (2010) Understanding plasmid effect on hyaluronic acid molecular weight produced by Streptococcus equi subsp. zooepidemicus. Metab Eng 12: 62–69.

Merryman, C., and Gibson, D.G. (2012) Methods and applications for assembling large DNA constructs. Metab Eng 14: 196–204.

Miller, E.M., and Nickoloff, J.A. (1995) Components of the Salmonella flagellar export apparatus and classification of export substrates. J Bacteriol 181: 1388–1394.

Minty, J.J., Lesnfsky, A.A., Lin, F., Chen, Y., Zaroff, T.A., Veloso, A.B., et al. (2011) Evolution combined with genomic study elucidates genetic bases of isobutanol tolerance in Escherichia coli. Microb Cell Fact 10: 18.

Mitra, A., Palaniyandi, S., Herren, C.D., Zhu, X., and Mukhopadhyay, S. (2013) Pleiotropic roles of uvrY on biofilm formation, motility and virulence in uropathogenic Escherichia coli CFT073. PLoS ONE 8: e55492.

Mohawk, K.L., Poly, F., Wahl, J.W., Rasko, D.A., and Guerry, P. (2014) High frequency, spontaneous motA mutations in Campylobacter jejuni strain 81–176. PLoS ONE 9: e88043.

Murphy, K.C., and Campellone, K.G. (2003) Lambda Red-mediated recombinogetic engineering of enterohemorrhagic and enteropathogenic E. coli. BMC Mol Biol 4: 11.

Ogawa, R., Abe-Yoshizumi, R., Kishi, T., Homma, M., and Kojima, S. (2015) Interaction of the C-terminal tail of Flil with Flig from the Na+-driven flagellar motor of Vibrio alginolyticus. J Bacteriol 197: 63–72.

Partridge, J.D., Nieto, V., and Harshey, R.M. (2015) A new player at the flagellar motor: Flil controls both motor output and bias. MBio 6: e02367–02314.

Pfaffl, M.W., Horgan, G.W., and Dempfle, L. (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res 30: e36.

Reboul, C.F., Andrews, D.A., Nahar, M.F., Buckle, A.M., and Roujeinikova, A. (2011) Crystallographic and molecular dynamics analysis of loop motions unmasking the peptidoglycan-binding site in stator protein MotB of flagellar motor. PLoS ONE 6: e18981.

Sabri, S., Steen, J.A., Bongers, M., Nielsen, L.K., and Vickers, C.E. (2013) Knock-in/knock-out (KIKO) vectors for rapid integration of large DNA sequences, including whole metabolic pathways, onto the Escherichia coli chromosome at well-characterized loci. Microb Cell Fact 12: 60.

St-Pierre, F., Cui, L., Priest, D.G., Endy, D., Dodd, I.B., and Shearwin, K.E. (2013) One-step cloning and chromosomal integration of DNA. ACS Synth Biol 2: 537–541.

Takahashi, Y., and Ito, M. (2014) Mutational analysis of charged residues in the cytoplasmic loops of MotA and MotP in the Bacillus subtilis flagellar motor. J Biochem 156: 211–220.

Takahashi, Y., Koyama, K., and Ito, M. (2014) Suppressor mutants from MotB-D24E and MotS-D30E in the flagellar stator complex of Bacillus subtilis. J Gen Appl Microbiol 60: 131–139.

Ubinskaya, A.A., Samsonov, V.V., Mashko, S.V., and Stoyanova, N.V. (2012) A PCR-free cloning method for the targeted φ80 Int-mediated integration of any long DNA fragment, bracketed with meganuclease recognition sites, into the Escherichia coli chromosome. J Microbiol Methods 89: 167–173.

Vora, T., Hottes, A.K., and Tavazoie, S. (2009) Protein occupancy landscape of a bacterial genome. Mol Cell 35: 247–253.

Yim, H., Haselbeck, R., Niu, W., Pujol-Baxley, C., Burgard, A., Boldt, J., et al. (2011) Metabolic engineering of Escherichia coli for direct production of 1,4-butanediol. Nat Chem Biol 7: 445–452.

Zhang, X., Shanmugam, K.T., and Ingram, L.O. (2010) Fermentation of glycerol to succinate by metabolically engineered strains of Escherichia coli. Appl Environ Microbiol 76: 2397–2401.

Zhou, L., Niu, D.D., Tian, K.M., Chen, X.Z., Prior, B.A., Shen, W., et al. (2012) Genetically switched D-lactate production in Escherichia coli. Metab Eng 14: 560–568.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site: