An RNA-dependent mechanism for transient expression of bacterial translocation filaments

Citation for published version:
Wang, D, McAteer, SP, Wawszczyk, AB, Russell, CD, Tahoun, A, Elmi, A, Cockroft, SL, Tollervey, D, Granneman, S, Tree, JJ & Gally, DL 2018, ‘An RNA-dependent mechanism for transient expression of bacterial translocation filaments’ Nucleic Acids Research, vol 46, no. 7, pp. 3366–3381. DOI: 10.1093/nar/gky096

Digital Object Identifier (DOI):
10.1093/nar/gky096

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Nucleic Acids Research

Publisher Rights Statement:
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
An RNA-dependent mechanism for transient expression of bacterial translocation filaments

Dai Wang
Sean P. McAteer
Agata B. Wawszczyk
Clark D. Russell
Amin Tahoun
Alex Elmi
Scott L. Cockroft
David Tollervey
Sander Granneman
Jai J. Tree
David L. Gally

1 Division of Infection and Immunity, The Roslin Institute and R(D)SVS, University of Edinburgh, Edinburgh EH25 9RG, UK, 2 State Key Laboratory of Molecular Vaccinology and Molecular Diagnostics, School of Public Health, Xiamen University, South Xiangan Rd., Xiangan District, Xiamen, Fujian Province 361102, China, 3 Centre for Synthetic and Systems Biology, University of Edinburgh, Edinburgh EH9 3BF, UK, 4 Faculty of Veterinary Medicine, Kafrelsheikh University, 33516 Kafrel-Sheikh, Egypt, 5 EaStCHEM School of Chemistry, University of Edinburgh, Joseph Black Building, David Brewster Road, Edinburgh EH9 3FJ, UK, 6 Wellcome Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3BF, UK and 7 School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney 2052, NSW, Australia

Received November 03, 2016; Revised January 28, 2018; Editorial Decision January 30, 2018; Accepted February 06, 2018

ABSTRACT

The prokaryotic RNA chaperone Hfq mediates sRNA–mRNA interactions and plays a significant role in post-transcriptional regulation of the type III secretion (T3S) system produced by a range of Escherichia coli pathotypes. UV-crosslinking was used to map Hfq-binding under conditions that promote T3S and multiple interactions were identified within polycistronic transcripts produced from the locus of enterocyte effacement (LEE) that encodes the T3S system. The majority of Hfq binding was within the LEE5 and LEE4 operons, the latter encoding the translocon apparatus (SepL-EspADB) that is positively regulated by the RNA binding protein, CsrA. Using the identified Hfq-binding sites and a series of sRNA deletions, the sRNA Spot42 was shown to directly repress translation of LEE4 at the sepL 5′ UTR.

In silico and in vivo analyses of the sepL mRNA secondary structure combined with expression studies of truncates indicated that the unbound sepL mRNA is translationally inactive. Based on expression studies with site-directed mutants, an OFF-ON-OFF toggle model is proposed that results in transient translation of SepL and EspA filament assembly. Under this model, the nascent mRNA is translationally off, before being activated by CsrA, and then repressed by Hfq and Spot42.

INTRODUCTION

Enterohaemorrhagic Escherichia coli (EHEC) strains can cause life-threatening infections in humans and the main serotype associated with disease in parts of Europe, North America and Japan is O157:H7. Cattle are the primary reservoir for EHEC O157 strains (1) and colonization of cattle or humans is dependent on a type III secretion (T3S) system expressed from the LEE pathogenicity island (2). The T3S system injects multiple effector proteins into host epithelial cells with various functions including cytoskeletal manipulation for intimate bacterial adherence and suppression of host cell inflammatory responses (2,3). Assembly of the T3S system is a staged process requiring the expression of a basal apparatus spanning both bacterial membranes (LEE1–3 operons) before production of the hollow filaments (LEE4 operon) through which effector proteins will be transferred (4), including the translocated intimin receptor (Tir) expressed from LEE5. T3S translocation filaments are produced by both EHEC and enteropathogenic E. coli (EPEC) and are composed of EspA (5). These filaments allow secretion of bacterial effector proteins through a pore in the host cell membrane composed of EspD and EspB, all expressed from LEE4 (6,7). Together this structure is known as the translocon and its secretion is dependent on two interacting proteins, SepL and SepD, that govern the switch from translocon assembly to effector protein secretion (8–10).

The assembly and function of complex multi-component organelles, such as T3S systems, requires hierarchical control at the transcriptional, translational and post transla-
tional levels (4,11,12). Translational control by sRNAs facilitates rapid regulatory responses in bacteria (13–16) and there is increasing evidence for the role of sRNAs in the regulation of virulence factors including T3S systems (17–21). Interactions between mRNAs by sRNAs are often catalysed by Hfq which forms doughnut shaped hexamers that guide sRNA–mRNA interactions. Hfq controls virulence factor expression in a number of different pathogens via sRNA interactions (22). In particular, Hfq is implicated in the regulation of T3S in *E. coli* via expression of the regulatory proteins Ler (LEE1) and GrlA (located between LEE1 and LEE2) (23,24). We previously reported separate post-transcriptional regulation of the LEE4 and LEE5 operons (25,26) and proposed a ‘checkpoint’ in secretion system assembly between the basal apparatus and the translocon filament (4,25). Bhatt et al. (27) demonstrated that the regulatory protein CsrA (Carbon Storage regulator A) which acts primarily on mRNA transcripts (28), activates T3S since CsrA protein (CarbonStorageregulator A) which acts as a repressor on T3S (4,25). Butte et al. (27) demonstrated that the regulatory protein CsrA (Carbon Storage regulator A) which acts primarily on mRNA transcripts (28), activates T3S since a *csrA* mutant exhibits reduced secretion of EspADB and Tir. CsrA was shown to bind to the LEE4 transcript, acting at two predicted sites in the *sepL* 5′ UTR (27). CsrA also has a repressive effect when over-expressed due to negative regulation of *grrA*, indicating that physiological levels of CsrA are required for normal control. Once transcribed, the LEE4 transcript is then processed by RNase E towards the 3′ end of *sepL*, which presumably prevents further SepL production and may have consequences for espADB transcript translation (29). While *sepL* and *espADB* are initially transcribed on the same transcript (Figure 1F), there is processing of this transcript (29), and *sepL* mutations can be complemented in *trans* (10,30). In the absence of *sepL*, EHEC has a hypersecretion phenotype consistent with the SepL–SepD complex acting as a gate to control translocon filament production. EspA is still produced inside the bacterium in the absence of *sepL* indicating its main role is to govern assembly of the filaments and not regulate their expression (8,10,30).

We previously employed UV cross-linking and high throughput sequencing of cDNAs (CRAC) to define Hfq–RNA interactions (31). Here, we investigated the role of Hfq in the regulation of translocon expression in *E. coli* O157. CRAC identified LEE mRNA sequences that interact with Hfq to control the expression of T3S components. Using our Hfq-binding data to reduce the sequence space for *in silico* predictions of sRNA–mRNA interactions, we constructed a series of sRNA deletions to determine which sRNAs control translation of *sepL* and T3S secretion more broadly. We show that *sepL* is repressed by Spot42, a small RNA that is highly expressed under the T3S-permissive conditions used. The primary Hfq binding site overlaps a sequence identified for CsrA activation of translation and we provide evidence that the nascent folded *sepL* transcript would be translationally inactive. The interactions identified in our study provide a model for transient translation of protein expression, an OFF–ON–OFF toggle switch, to produce a short burst of SepL expression leading to EspA filament production.

**MATERIALS AND METHODS**

**Bacterial strains and plasmids**

The strains and plasmids used in this study are described in Tables 1 and 2 respectively. Two strains were studied as they differ markedly in their level of T3S. *Escherichia coli* O157 strain Sakai is a well annotated strain that has a relatively low level of secretion under laboratory conditions and for which a Shiga toxin negative variant is available (Table 1). *Escherichia coli* O157 ZAP193 (NCTC12900, Table 1) is a Shiga toxin negative strain that has a high secretion level compared with strain Sakai. In order to further investigate the role of sRNAs in EHEC O157:H7, sRNA mutants were constructed using allelic exchange with pIB307 as described previously (32). In brief, primer pairs (Table 3) were used to amplify flanking regions of each sRNA and the PCR products were cloned into pIB307 (Table 2). A tetracycline resistance gene from pTOF61 (Table 2) was then cloned using NotI restriction sites into the constructs containing sRNA flanking regions, except for pDW-l3Hf where the sacB/kan cassette from pDG028 was used (Table 2). These constructs were then transformed into both *E. coli* O157 strain ZAP193 and *E. coli* O157 strain Sakai *stx*-and allelic exchanges carried out as previously described (32) to generate the sRNA deletion strains (Table 1). The final plasmid constructs (Table 2) were sequenced prior to the deletion exchange and each deletion confirmed by PCR analysis.

Full-length and 51 bp translational fusions of GFP to *sepL* (pDW6 & 26) were as published previously (Table 2). Further translational fusions to the first genes of LEE1 (*ler*) and LEE5 (*tir*) were as also as published previously (Table 2). For the current study an extensive number of site-specific changes and truncations were engineered in the *sepL* transcript, based on pDW6, in order to examine the regions required for translational control (Table 2). A series of SepL–GFP fusions (Table 2) were also made for analysis of SepL translation. The primers used for the various constructs are listed in Table 3. In addition, specific changes were made in the chromosomal copy of Spot42 in the Sakai strain by allelic exchange (Table 1). The C1–C5 and H1 changes in *stx* of strain Sakai were generated with the Q5 method (New England Biolabs) whilst the other mutants were generated with the Quickchange method (Agilent Technologies). All constructs were confirmed by Sanger sequencing.

**Preparation of secreted proteins and bacterial fractions for protein analyses**

Bacteria were cultured in 50 ml of MEM-HEPES at 37°C (200 rpm) to an OD$_{600}$ of 0.7–1.0. Bacterial cells were pelleted by centrifugation at 4000 g for 20 min, and supernatants were passed through low protein binding filters (0.45 μm). 10% TCA was used to precipitate proteins overnight, which were separated by centrifugation at 4000 g for 30 min at 4°C. The proteins were suspended in 150 μl of 1.5 M Tris (pH 8.8). For bacterial lysates, bacterial pellets were suspended directly in SDS PAGE loading buffer. Proteins were separated by SDS-PAGE using standard methods and Western blotting performed as described previously for EspD, RecA and EscJ (33,34). Bacteria were stained for EspA filaments following fixation with 4% paraformalde-
Table 1. Bacterial strains, media, plasmids and primers used in this study

| Strains          | Details/reference                       | Plasmids used for construction (if required) |
|------------------|----------------------------------------|---------------------------------------------|
| ZAP193           | E. coli O157:H7 stx-, NCTC 12900       |                                             |
| Sakai stx-       | Dahan et al. (60)                       |                                             |
| Sakai stx- Hfq-HTF | Tree et al. (31)                        |                                             |
| ZAP1419          | ZAP193: micF-                           | pDW-mFf then pDW-mFf2                       |
| ZAP1420          | ZAP193: micA-                           | pDW-mAf then pDW-mAf2                       |
| ZAP1421          | ZAP193: rybB-                           | pDW-hBf then pDW-hBf2                       |
| ZAP1422          | ZAP193: ryhB-                           |                                             |
| ZAP1423          | ZAP193: omrA-                           |                                             |
| ZAP1424          | ZAP193: omrB-                           |                                             |
| ZAP1425          | ZAP193: spf-                            | pDW-sfft then pDW-sfl2                      |
| ZAP1426          | ZAP193: furS-                           | pDW-fSf then pDW-fSf2                      |
| ZAP1427          | ZAP193: mcaS-                           |                                             |
| ZAP1428          | ZAP193: rprA-                           | pDW-rAf then pDW-rAf2                       |
| ZAP1429          | Sakai stx-, micF-                       | pDW-mFf then pDW-mFf2                       |
| ZAP1430          | Sakai stx-, micA-                       | pDW-mAf then pDW-mAf2                       |
| ZAP1431          | Sakai stx-, rybB-                       | pDW-bBf then pDW-bBf2                       |
| ZAP1432          | Sakai stx-, ryhB-                       | pDW-hBf then pDW-hBf2                       |
| ZAP1433          | Sakai stx-, omrA-                       | pDW-oAf then pDW-oAf2                       |
| ZAP1434          | Sakai stx-, omrB-                       | pDW-oBf then pDW-oBf2                       |
| ZAP1435          | Sakai stx-, spf-                        | pDW-sfft then pDW-sfl2                      |
| ZAP1436          | Sakai stx-, furS-                       | pDW-fSf then pDW-fSf2                      |
| ZAP1437          | Sakai stx-, mcaS-                       |                                             |
| ZAP1438          | Sakai stx-, rprA-                       | pDW-rAf then pDW-rAf2                       |
| ZAP1439          | Sakai spf tet                           |                                             |
| ZAP1442          | Sakai spf (WT)                          | pIB-spf                                    |
| ZAP1477          | Sakai spf (S1)                          | pIB-spf-S1                                 |
| ZAP1477          | Sakai spf (S2)                          | pIB-spf-S2                                 |

hydro for 5 min. His-EspA was purified and anti-EspA serum raised in rabbits as described (35). 1 ml of each culture was centrifuged and washed twice in PBS and the pellets were suspended in 4% PFA for 15 min. The fixed bacteria were washed twice with PBS and the pellet suspended in PBS. 20 μl was added to 5 ml aliquots of the strain and mixed thoroughly. These were incubated with a 1/100 dilution of anti-EspA antibody (laboratory stocks as described above) in PBS with 0.1% bovine serum albumin (PBS-0.1% BSA) and incubated in a moist box overnight at 4°C. After three washes with PBS-0.1% BSA sample was incubated with fluorescein isothiocyanate-conjugated donkey anti-rabbit immunoglobulin antibody in PBS-0.1% BSA (1/500; Abcam) for 2 h in a moist box, and the slides washed three times with PBS-0.1% BSA. The slides were mounted with ProLong Gold mount medium (Thermo Fisher). The slides were then examined by confocal microscopy using a Zeiss Plan Apochromat 63×/1.4 oil immersion lens.

**In vivo probing of mRNA structure**

The structure of sepL mRNA around the RBS was probed using 2-methylnicotinic acid imidazolide (NAI-N3) based on modifications of published protocols (36,37). Synthesis of NAI-N3 was performed according to (36). *Escherichia coli* were grown in LB overnight, diluted in fresh media and cultured until OD600 = 0.5. 25 mM NAI was added to 5 ml aliquots of the bacteria and mixed thoroughly. The mixture was incubated for 10 min at 37°C with shaking and the cells harvested by centrifugation at 4°C and flash-frozen in liquid nitrogen before storage at −80°C if not used immediately. Control and NAI-modified bacterial pellets were resuspended in 650 μl of GTC phenol and vortexed for 3 min with 50 μl of glass beads. The suspension was incubated at 65°C for 10 min and placed on ice for another 10 min. 300 μl of chloroform/IAA was added to the suspension and then 80 μl of 100 mM NaOAc pH 5.2. The sample was vortexed and micro-centrifuged at 14 000 rpm for 5 min. The upper phase was added to 500 μl phenol:chloroform:isoamyl alcohol (IAA) and centrifuged for 2 min at 14 000 rpm. The upper phase was removed and added to 500 μl of chloroform:IAA. The upper phase was taken and re-suspended and centrifuged and added to 1 ml ethanol and 20 μg glycogen. The RNA was precipitated overnight at −20°C and pelleted by centrifugation for 30 min at 14 000 rpm and 4°C; the ethanol was removed and the pellet washed with 70% ethanol, air-dried and suspended in 50 μl MilliQ water. The RNA was measured with a Nanodrop. The primer extension was carried out using a 32P end-labeled primer pDW6(1)—GAATTTTTTTGCGTTACGGCATTGCATTGGCTGAG based on a published method (38). RNA was used at a concentration of 3.84 μg/μl in 1.75 μl. The extension products and sequencing ladder were run on a 6% acrylamide gel and products visualized using a FUJI-FLA 5000 imager.

**Flow cytometry**

When the OD600 of the cultures used for population fluorescence analysis reached 0.7, bacteria were fixed by diluting a 500 μl aliquot 1:2 in 4% paraformaldehyde. Single colour fluorescence was measured using a FACScalibur flow cytometer (Becton Dickinson). Excitation was at 488 nm and emission captured at 530 nm. CELLQuest software was used to acquire and analyse flow cytometry data. A wild type culture of the strain for analysis was stained by indirect immunofluorescence for O157, a FITC conjugated secondary antibody. This fluorescence signal was used for quantitative analysis as described (34).
### Table 2. Plasmids used in this study

| Plasmids     | Details                                                                 |
|--------------|-------------------------------------------------------------------------|
| pIB307       | pMAK705-based vector for allelic exchange; temperature-sensitive replicon (61) |
| pDW-mAI2     | pIB307 containing upstream and downstream flanking regions of micA       |
| pDW-mFI2     | pIB307 containing upstream and downstream flanking regions of micF       |
| pDW-hBI2     | pIB307 containing upstream and downstream flanking regions of ryhB       |
| pDW-bBI2     | pIB307 containing upstream and downstream flanking regions of rybB       |
| pDW-oAI2     | pIB307 containing upstream and downstream flanking regions of omrA       |
| pDW-oBI2     | pIB307 containing upstream and downstream flanking regions of omrB       |
| pDW-rAI2     | pIB307 containing upstream and downstream flanking regions of rprA       |
| pDW-rSI2     | pIB307 containing upstream and downstream flanking regions of fnrS       |
| pDW-nSI2     | pIB307 containing upstream and downstream flanking regions of mcaS       |
| pDW-mAIft    | pDW-mA2 + tet cassette from pTOF61                                      |
| pDW-mFIft    | pDW-mFI2 + tet cassette from pTOF61                                     |
| pDW-hBIft    | pDW-hBI2 + tet cassette from pTOF61                                     |
| pDW-bBIft    | pDW-bBI2 + tet cassette from pTOF61                                     |
| pDW-oAIft    | pDW-oAI2 + tet cassette from pTOF61                                     |
| pDW-oBIft    | pDW-oBI2 + tet cassette from pTOF61                                     |
| pDW-rAIft    | pDW-rAI2 + tet cassette from pTOF61                                     |
| pDW-rSIft    | pDW-rSI2 + sacB/kan cassette from pDG028                                 |
| pDW-nSIft    | pDW-nSI2 + tet cassette from pTOF61                                     |
| pTOF61       | A plasmid derived from pTOF1 (62) has TcR cloned into the Smαl site     |
| pDG028       | Low-copy-number vector containing sacB/kan cassette (25)               |
| pDW6         | pAJR70 digested with BamHI/KpnI; whole sepL gene with its own promoter amplified from ZAP193 and inserted (10) |
| pDW-LEE1     | Encodes full length Ler fusion to GFP expressed from WT promoter (33)  |
| pDW-LEE5     | Encodes full length Tir fusion to GFP expressed from WT promoter (33)  |
| pDW26        | First 51bp of sepL cloned in frame to eGFP in pAJR70 (10)              |
| pDW6-S1      | Mutation of sepL 5’UTR with S1 primers                                  |
| pDW6-S2      | Mutation of sepL 5’UTR with S2 primers                                  |
| pDW6-S3      | Mutation of sepL 5’UTR with S3 primers                                  |
| pDW6-C1      | Mutation of sepL 5’UTR with C1 primers                                  |
| pDW6-C2      | Mutation of sepL 5’UTR with C2 primers                                  |
| pDW6-C3      | Mutation of sepL 5’UTR with C3 primers                                  |
| pDW6-C4      | Mutation of sepL 5’UTR with C4 primers                                  |
| pDW6-C5      | Mutation of sepL 5’UTR with C5 primers                                  |
| pDW6-H1      | Mutation of sepL 5’UTR with H1 primers                                  |
| pIB-spf      | pIB307 containing spf plus upstream and downstream flanking regions    |
| pIB-spf-S1   | pIB-spf mutated with S1 primers                                        |

**to set the R1 gate, to exclude small particles and debris from being counted by the cytometer.**

### Analysis of fluorescence levels

The total fluorescence produced by the population was determined by analyzing 150 µl aliquots of culture with a fluorescent plate reader (Fluostar Optima; BMG). Each expression experiment was carried out at least three times with separate bacterial cultures and fluorescence values were only used for analysis when the bacterial cultures were between 0.5 and 1.1 OD₆₀₀ and these were normalized to the optical density. A student t-test was used to analyse the level of significance between strain backgrounds. For microscopy fluorescence measurements, a 50 µl aliquot was removed from the culture and diluted 1:1 in 4% PFA, and 20 µl was dried on a glass slide (37°C, 15 min) which was then washed three times with PBS and a coverslip applied using DAKO fluorescent mounting medium. Fluorescence imaging was carried out using a Leica DM LB2 microscope and a 100× objective lens. Narrow-band width filters to excite and detect eGFP/FITC were used (4107 Endow GFP, CHROMA). Images were captured using a Hamamatsu ORCA-ER black and white CCD digital camera.

### Identification of Hfq-binding sites by UV-crosslinking (CRAC analysis)

All Hfq UV-crosslinking datasets were previously described in Tree et al. (31) and deposited at NCBI GEO under the accession number GSE46118. Raw sequence data was aligned to the EHEC str. Sakai genome (NC_002695.1) using novoalign software and analysed using pyCRAC software (39).

### RESULTS

**Hfq interactions with the LEE determined by UV-crosslinking (CRAC)**

RNAs that interact with Hfq were identified under conditions known to promote T3S expression in *E. coli* O157:H7 strains (growth in MEM-HEPES medium). Hfq UV-crosslinking (CRAC) was previously applied to *E. coli* O157:H7 strain Sakai containing chromosomally-inserted 6× His and FLAG-tagged Hfq as described (31). Under these conditions, LEE transcripts were a significant target for Hfq interactions and accounted for between 1.8–8.8% of total mapped reads in replicate Hfq-CRAC datasets (Figure 1A). In line with published research (40,41), Hfq-binding...
### Table 3. Primers used in this study

| Primer | Sequence | Application |
|--------|----------|-------------|
| micA55 | CCCAAGCTTccggcggcatcgagcctc | pDW-mAf2 |
| micA53 | GCCTTGACGCGGCGGC | atatactcagactcgcctgg |
| micA35 | GCCCTGACGCGGCGGC | pDW-mAf2 |
| micA33 | GGTTACCTcgcgaactactatcgcttcg | pDW-mAf2 |
| micF55 | CCCAAGCTTcatcagtacgtttggagctg | pDW-mFf2 |
| micF53 | GCCCTGACGCGGCGGC | ggtagttattctagttgcgag |
| micF35 | GCCCTGACGCGGCGGC | pDW-mFf2 |
| micF33 | GCCCTGACGCGGCGGC | pDW-mFf2 |
| mcaS55 | GCCGGATCCatcagtacgtttggagctg | pDW-mSf2 |
| mcaS53 | GCCGGATCCatcagtacgtttggagctg | pDW-mSf2 |
| mcaS35 | GCCGGATCCatcagtacgtttggagctg | pDW-mSf2 |
| mcaS33 | GCCGGATCCatcagtacgtttggagctg | pDW-mSf2 |
| ryhB55 | GCCGGATCCatcagtacgtttggagctg | pDW-hBf2 |
| ryhB53 | GCCGGATCCatcagtacgtttggagctg | pDW-hBf2 |
| ryhB35 | GCCGGATCCatcagtacgtttggagctg | pDW-hBf2 |
| ryhB33 | GCCGGATCCatcagtacgtttggagctg | pDW-hBf2 |
| rprA55 | GCCGGATCCatcagtacgtttggagctg | pDW-rAf2 |
| rprA53 | GCCGGATCCatcagtacgtttggagctg | pDW-rAf2 |
| rprA35 | GCCGGATCCatcagtacgtttggagctg | pDW-rAf2 |
| rprA33 | GCCGGATCCatcagtacgtttggagctg | pDW-rAf2 |
| fnrS55 | GCCGGATCCatcagtacgtttggagctg | pDW-fSf2 |
| fnrS53 | GCCGGATCCatcagtacgtttggagctg | pDW-fSf2 |
| fnrS35 | GCCGGATCCatcagtacgtttggagctg | pDW-fSf2 |
| fnrS33 | GCCGGATCCatcagtacgtttggagctg | pDW-fSf2 |
| omrA55 | GCCGGATCCatcagtacgtttggagctg | pDW-oAf2 |
| omrA53 | GCCGGATCCatcagtacgtttggagctg | pDW-oAf2 |
| omrA35 | GCCGGATCCatcagtacgtttggagctg | pDW-oAf2 |
| omrA33 | GCCGGATCCatcagtacgtttggagctg | pDW-oAf2 |
| spf55 | GCCGGATCCatcagtacgtttggagctg | pDW-sff2 |
| spf53 | GCCGGATCCatcagtacgtttggagctg | pDW-sff2 |
| spf35 | GCCGGATCCatcagtacgtttggagctg | pDW-sff2 |
| C1 SDM FOR | tagatgtaaatattacgcttgta | pDW6-C1 |
| C2 SDM FOR | tagatgtaaatattacgcttgta | pDW6-C2 |
| C3 SDM FOR | tagatgtaaatattacgcttgta | pDW6-C3 |
| C1, C2, C3 SDM REV | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| C4 SDM FOR | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| C4 SDM REV | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| C5 SDM FOR | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| C5 SDM REV | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| H1 SDM REV | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| H1 SDM REV | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| S1 SDM FOR | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| S1 SDM REV | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| S2 SDM FOR | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| S2 SDM REV | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| S3 SDM FOR | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| S3 SDM REV | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| spf-S1 SDM FOR | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| spf-S1 SDM REV | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| spf-S2 SDM FOR | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| spf-S2 SDM REV | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| sepL-full used with below ATG | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| sepL-ATG | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| sepL+6 | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| sepL+9 | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| sepL+12 | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| sepL+15 | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| sepL+18 | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| sepL+27 | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
| sepL+30 | tagatgtaaatattacgcttgta | pDW6-C1, pDW6-C2, pDW6-C3 |
sites were recovered at both LEE1 (Ler) and within the grlRA di-cistronic transcript (Figure 1B,D,E). Extensive Hfq-binding was identified throughout the polycistronic LEE operons, suggesting that T3S is post-transcriptionally regulated by Hfq at many sites. Within the LEE, the LEE4 and LEE5 transcripts were the most highly recovered accounting for 27.9 ± 5.2% and 52.5 ± 9.7% of reads mapping to the LEE respectively. LEE4 encodes SepL and the EspADB translocon and LEE5 encodes intimin, Tir and CesT. The LEE1 and LEE5 respectively. LEE4 encodes SepL and the EspADB translocon and LEE5 encodes intimin, Tir and CesT.

Table 3. Continued

| Primer | Sequence | Application |
|--------|----------|-------------|
| sepL+33 | gggtaacgcgaggaacgtcaggtcaggggg |             |
| sepL+36 | gggtaacgcgaggaacgtcaggtcagggttg |             |
| sepL+39 | gggtaacgcgaggaacgtcaggtcagggttg |             |
| sepL+42 | gggtaacgcgaggaacgtcaggtcagggtgc |             |
| sepL+45 | gggtaacgcgaggaacgtcaggtcagggtgc |             |
| sepL+48 | gggtaacgcgaggaacgtcaggtcagggtgc |             |
| sepL+75 | aaaggtatacaatattatcattataaatcag |             |
| sepL+78 | aaaggtatacaatattatcattataaatcag |             |
| sepL+81 | ggggtaacgacccaccggggtggtgcctg |             |
| sepL+84 | ggggtaacgacccaccggggtggtgcctg |             |
| sepL+87 | ggggtaacgacccaccggggtggtgcctg |             |
| sepL+90 | ggggtaacgacccaccggggtggtgcctg |             |
| sepL+102 | aaaaaggaagcaggttattttgtcattatg |             |
| sepL+123 | aaaggtatacaatattatcattataaatcag |             |
| sepL+210 | ggggtaacacatttcttcttcatccaggg |             |

Support for a sequestered RBS from structural probing of the sepL transcript in vivo

The 5′ UTR and early coding region of the sepL mRNA are predicted to form the structure shown in Figure 2A which includes four stem loops. According to this RNAfold structure, the sepL RBS is occluded in the second stem loop (SL2) which may limit translational activation unless there are further interactions to disrupt the structure allowing ribosome access to the RBS. In order to gain insights into the structure of this region of the sepL mRNA transcript in vivo, selective 2′-hydroxyl acylation (SHAPE analysis) was applied on the basis that the more flexible and often single-stranded RNA regions show more 2′-hydroxyl reactivity than RNA regions that are less flexible and usually base-paired. In our case 2-methyl nicotinic acid imidazolide (NAI) was used as the electrophile. Following in vivo treatment of the bacteria and RNA extraction, primer extension analysis was then used which truncates at modified bases therefore indicating where the mRNA has been processed. DMSO was used as a control treatment. This in vivo analysis provides support for the predicted structure, in particular the presence of stem loop 2 (SL2, Figure 2A and B), for which the two sides of the stem, including the RBS are protected from processing compared to the control and the predicted loop bases are more heavily processed (Figure 2B). The analysis also supports the prediction of the AUG codon being more accessible between two more protected regions (Figure 2B).

An activating CsrA binding site overlaps with an Hfq interaction site in the 5′ UTR of sepL

CsrA is an activator of T3S in EPEC/EHEC and directly interacts with the 5′ UTR of the LEE4/sepL transcript (27). Bhatt et al identified two potential CsrA binding sequences that closely match the consensus AUGGA (Figure 3A, motifs 1 and 2) and demonstrated that CsrA binds the sepL 5′ UTR using an electrophoresis mobility shift assay (EMSA). To examine the relative contributions of each motif to sepL...
Figure 1. The RNA chaperone, Hfq, binds multiple sites within the polycistronic mRNAs of the LEE. (A) The circular plot defines the location of Hfq-associated RNA sequences within the E. coli O157 Sakai genome; shown for both strands based on our previous Hfq UV-crosslinking study (Materials and Methods, (31)). The location of integrated and cryptic prophage elements (Sp), and the Locus of Enterocyte Effacement (LEE) are on the inner ring indicated (grey boxes). (B) Hfq binding across the LEE. Genomic features within the LEE are shown (centre, grey boxes), including the five polycistronic LEE operons (LEE1–5, green boxes) and grlRA. Positive strand features are shown above the genomic features and negative strand below. Hfq binding sites identified by UV-crosslinking (31) are shown in blue for triplicated datasets. From this data, non-genomically encoded oligo(A) tails (indicative of RNA cleavage sites) were also extracted and plotted in red. Replicate control (wild type untagged Hfq) are also shown in blue and represent the background recovery of mRNAs under our experimental conditions. Deletions in sequencing reads are indicative of direct Hfq-RNA contact and are plotted in green for each Hfq UV-crosslinking experiment. (C) Hfq binding sites with the 5′ UTR and 5′ end of sepL are shown from two rounds of Hfq-CRAC analysis. The mapping indicates two predominant sites of Hfq interaction at peaks 702 and 703 (indicated by arrowheads). (D) Hfq binding at the grlRA dicistronic mRNA. Replicate Hfq UV-crosslinking experiments are presented. (E) Hfq-binding to the 5′ UTR and first 100nt of the master regulator of the LEE, ler. Replicate Hfq UV-crosslinking experiments are shown. (F) Diagram of the LEE4 operon. The operons in the Locus of Enterocyte Effacement (LEE) are ordered LEE1, 2, 3, 5, 4 (opposite direction to that shown in panel B) and LEE4 is positioned after LEE5 and the escD promoter (shown). The operon extends from sepL, through the translocon-encoding genes, espADB, and three genes including escF that encodes a needle structure over which the EspA filament is assembled. The polycistronic mRNA is expressed from a promoter (P_{LEE4}) in front of sepL. The LEE4 mRNA is processed by an RNaseE site in the 3′ end of sepL (29).
regulation, we introduced point mutations predicted to disrupt CsrA binding (Figure 3A, mutations C1–4). All point mutations were introduced into a translational reporter for sepL (pDW6) that includes the native promoter, the 83 nt 5′ UTR and the entire sepL open reading frame fused to C-terminal GFP. The WT and mutant reporters were transformed into E. coli O157 strain ZAP193 (NCTC12900; Table 1) that exhibits a high level of T3S expression, facilitating the measurement of any repressive effects of the mutations. A consensus motif for CsrA binding has been described (43) and so three mutations (C1–3) in motif 1 were constructed based on this consensus. C1 was predicted to have a greater effect than C2 and this in turn was predicted to have a greater impact than C3. In fact, all three markedly reduced SepL-GFP expression (P<0.0022) indicating a strong requirement for this consensus CsrA binding sequence (Figure 3Bi and C). While the relative levels of repression did decrease from C1 to C2 to C3, in line with their expected impact, the values were not significantly different from each other (Figure 3Bi and C).

The C4 mutation in CsrA motif 2 also destabilizes the repressive stem loop SL2 (Figure 3A) and so we also introduced a compensatory mutation C5 into SL2 to distinguish CsrA-dependent and SL2-dependent regulation. The mutation of motif 2 (C4) had little effect on SepL-GFP translation, while the C5 mutation, that changes the opposite strand to maintain predicted base-pairing, mildly increased translation (Figure 3Bii). Motif 2 is therefore unlikely to be required for CsrA-mediated translation activation. Taken together, the site-directed mutagenesis indicates that CsrA recruitment to motif 1 in SL1 is important for CsrA-dependent activation of sepL translation.

Since the Hfq interaction peak 703 overlaps with the motif 1 CsrA binding site, we determined whether mutation of the Hfq consensus sequence ARN5m2 altered SepL translation. The H1 mutation (Figure 3A) is predicted to remove the ARN5m2 motif without disrupting base-pairing within the pitch-fork RNA structure. Introduction of mutation H1 into the translational SepL-GFP fusion increased GFP expression 1.5 fold in strain ZAP193 and 23 fold in Sakai (Figure 3Biii-iv). We conclude that the ARN5m2 motif is
Figure 3. Mutagenesis of CsrA and Hfq interacting sequences in the sepL 5′ UTR. (A) the predicted secondary structure of sepL 5′ UTR and 5′ CDS on the left-hand side of the panel shows the two AGGnA sequences predicted to interact with CsrA. The Hfq binding peaks, deletions, and ARN5m2 motif are as indicated in Figure 1. A mutation (H1) was introduced into the ARN5m2 motif to disrupt Hfq interactions at this site (peak 703, Figure 1). Four site-specific mutations (C1–4) were introduced to test their impact on CsrA activation of sepL translation while C5 was introduced to restore base pairing in SL4 that may be destabilised by the C4 mutations. (B) Measurement of the wild type and mutated SepL-GFP fusions in either the E. coli O157:H7 strain ZAP193 (high T3 secretor background) or strain Sakai (low T3 secretor). Fluorescence was measured for cultures over a range of optical densities and then adjusted for optical density as described in Materials and Methods. (C) Flow cytometry analysis of SepL-GFP and C1–4 mutations in E. coli O157:H7 ZAP193 (high T3 secretor). Expression of the wild type fusion in the ΔcsrA background indicates minimal levels of secretion in the absence of translational activation. Significance was calculated using a paired t-test for replicate data collected over multiple experiments (panel B, ZAP193 sepL-C4, -C5 and -C4&5). An unpaired t-test was used for all other samples.
repressive, consistent with our Hfq-CRAC data and other
Hfq binding data.

Identification of sRNAs controlling T3S and SepL expression

We, and others, have previously shown that the sRNA binding site is closely associated with the Hfq binding site defined by CRAC or CLIP-Seq (31,44). We therefore restricted our search for complementary sRNAs to the Hfq binding site defined by Hfq-CRAC reads in sepL (bases −82 to +50; Figure 1C) using the IntaRNA program (45). We focused our search to sRNAs that were abundant in our Hfq-CRAC data, which were: MicA, MicF, RybB, RyhB, OmrA, OmrB, FnrS, Spot42, McaS and RprA. Of these sRNAs, MicF, FnrS, Spot42, and OmrA had some complementarity to the sepL Hfq-binding site. Each of the ten sRNAs were deleted in ZAP193 (high T3S) and Sakai (low T3S), to allow the assessment of negative and positive regulatory effects, respectively (Figure 4A). These strains were analysed for secreted EspD in bacterial supernatants as an indicator of either EspD and total levels of EscJ.

bindingsite defined by Hfq-CRAC readsin
stricted our search for complementary sRNAs to the Hfq

Spot42 directly represses SepL translation

The polycistrionic LEE4 operon encodes the filament protein EspA and pore forming proteins, EspBD. Repression of sepL translation by Spot42 binding in SL2 (Figure 5A) is predicted to block filament expression and transport to the cell surface. As further confirmation of regulation by Spot42, we used immunofluorescence to monitor EspA filament expression in the low-secretor strain Sakai, an iso-
genric Δspf variant, and its chromosomally-repaired comple-
ment (Δspf::spf). Deletion of spf increased the frequency of EspA filamentation on bacterial cells, whereas filamentation was reduced to wild type levels in the chromosomally repaired strain (Figure 5B). This result was supported by western blotting for EspD in the Δspf and complemented backgrounds (Figure 5C).

Our data indicated that Spot42 acts on several operons contributing to T3S, including LEE1, and we therefore determined whether Spot42 directly regulates sepL translation. Spot42 was predicted to interact with the LEE4 transcript adjacent to the sepL ribosomal binding site, at peak 702 (Figure 5A and D). Two mutations (S1 and S2, Figure 5A) were constructed in the full length SepL translational fusion and transformed into Sakai to monitor potential de-
pression of T3S. As SL2 is predicted to form a repressive stem loop at the RBS, mutation S2 was designed to retain base pairing of this structure with reciprocal changes introduced on the opposite side of the stem loop (Figure 5A). The S1 mutation did not result in a significant increase in expression in the wild type background, while the S2 mutation increased expression about five-fold ($P < 0.05$), which is ~50% of the increase in expression observed for the spf deletion under these conditions (Figure 5E).

Two compensatory mutations were then introduced into the chromosomal copy of spf to match the changes made in the proposed Spot42 binding site (sepL-S1 and -S2, Figure 5D). Both of the chromosomal spf-S1 and spf-S2 point mutations increased T3S to levels comparable to the Δspf strain (Figure 5C), indicating that these chromosomal spf mutations act in an equivalent way to spf deletion. Similarly, when the strains expressing the spf-S1 or spf-S2 chromoso-
mal alleles were transformed with the wild type SepL-GFP fusion, GFP expression was de-repressed to levels compara-
ble to the Δspf mutant (Figure 5E). When the reciprocally modified SepL-GFP fusions were introduced into the Δspf strain it was noted that sepL-S2 trended towards higher expression ($P = 0.27$) suggesting that part of this increased activity may be spf independent. When both fusions were then matched with their cognate spf-S1 or -S2 backgrounds, both showed reduced expression (Figure 5E). This was statistic-
ally significant for the sepL-S1 mutant ($P = 0.05$) relative to the wild type sepL fusion in this background, and tend-
Figure 4. Analysis of sRNA mutations on type 3 secretion and specific LEE transcripts. (A) Western blot analysis for EspD (supernatant), EscJ (whole cell) and RecA (whole cell control) from WT and sRNA deletion mutants of \textit{E. coli} O157:H7 strains Sakai (low T3 secretor) and ZAP193 (high T3 secretor). (B) Bar graphs showing relative expression levels of the SepL+51-GFP fusion (see text for details) in the WT and sRNA deletion backgrounds of \textit{E. coli} O157:H7 strain Sakai. The asterisk indicates that the mutant did not grow to an OD\textsubscript{600} > 0.6 under our experimental conditions. (C) Translation of the LEE encoded proteins Ler, SepL and Tir was monitored in WT, Δspf, and Δmic\textsubscript{A} backgrounds for both low T3 secretor strain Sakai, and high T3 secretor strain ZAP193. Asterisks indicate \( P < 0.05 \).

Sequence requirements for SepL translation
To identify additional sequence requirements for translational control of SepL, a series of GFP fusions were constructed at different positions in the \textit{sepL} CDS (Figure 6A-B). All included the 83 base 5′ UTR and native promoter of \textit{sepL} (Table 2). GFP expression was measured in strain Sakai (Figure 6B-D). There was little detectable translation until the first 27 nucleotides of the open reading frame were included in the construct, indicating the presence of 5′ elements essential for translation, potentially through formation of predicted stem loop SL3 (Figure 6B). Expression increased for the fusion at nt +27 and remained relatively high until nt +75, for which expression was significantly lower. GFP levels increased again for fusions at nt +81 and nt +84, after which expression was reduced and remained relatively low and equivalent to the full-length \textit{sepL} truncated fusions to eGFP were plotted (Figure 6E) and this predicts key transitions in structure at \( \sim 27\text{nt} \) and \( \sim 75\text{nt} \) (Figure 6E). While these are only estimations, the latter transition may correlate with variability in expression measured for constructs from bases +75 to +87. (Figure 6A-D). Specifically, the +81 nt construct exhibited a biphasic expression pattern representing instability between high- and low-level expression (Figure 6E). The transitions in GFP expression measured in Figure 6B correlated well with the positions of stem loops SL3 and SL4 in the predicted structure (Figure 6A), although complete deletion of the sequence within the predicted stem loop SL4 had no ef-
Figure 5. Analysis of Hfq and Spot42 interactions with the sepL transcript. (A) Two site specific mutations in predicted stem loop (SL) 2 are shown (S1 and S2) along with other features described in Figures 1 and 2. (B) EspA filament staining on the surface of E. coli O157:H7 strain Sakai and the isogenic Δspf and complemented strains. EspA filament staining (green) was carried out as described in Materials and Methods with the bacteria surface stained using anti-O157 (red). (C) Western blot for EspD in bacterial supernatants and RecA from whole cell preparations from WT strain Sakai, the Δspf deletion mutant, and chromosomal knock-ins using spf, the spf-S1, or spf-S2, to replace the deleted region. (D) Predicted base pairing between sepL and Spot42, and compensatory mutations in the sepL 5′ UTR predicted to restore base pairing with the spf-S1 and spf-S2 site-directed mutants. (E) Plasmid-based SepL-GFP fusions carrying the S1 and S2 mutations were assayed in wild type E. coli O157:H7 strain Sakai, the isogenic spf deletion, and chromosomal knock-ins of the spf-S1 and spf-S2 mutations (indicated below). Significance was calculated relative to WT spf and pSepL-GFP fluorescence unless indicated. A single asterisk indicates $P \leq 0.1$ and a double asterisk indicates $P \leq 0.05$. 
is proposed to result in transient SepL translation and assembly of the EspA translocon.

sepL initially prevents translation due to poor access to the RBS. CsrA is required for translational activation along with specific sequence in the early part of the sepL CDS is shown for orientation of possible secondary structure absent in sepL truncates shown in panel B. (B) A systematic series of fusions to GFP were constructed into the sepL open reading frame to investigate sequence requirements for expression. All the fusions included the 5′ UTR and natural promoter of sepL (Table 2). The shortest fusion encompassed the native sepL promoter to the initiation codon (ATG). Successive three nucleotide extensions are numbered by how far they extend (in bases) into sepL. The final construct is the full length (FL) fusion used in previous studies (pDW6, Table 2). Expression from this series of constructs was measured in the low secretor strain Sakai. (C) Fluorescence microscopy of selected sepL-GFP constructs to illustrate transitions in sepL translation. The 81 nt fusion exhibits a clear biphasic expression pattern, with approximately half of the cells expressing GFP. (D) Flow cytometry of sepL truncates presented in panel C. (E) Graph showing the maximum ΔG values for the different truncates fused to GFP. This provides an indication of when major transitions in structure may occur between the different length fusions. (F) Model of the post-transcriptional regulation leading to transient expression of the sepL transcript. The model is described within the Discussion section of the manuscript. It is proposed that the transcript adopts a structure that initially prevents translation due to poor access to the RBS. CsrA is required for translational activation along with specific sequence in the early part of sepL (1–18 bases). SepL translation is silenced by Hfq-Spot42 binding to occlude the RBS and potentially by direct competition with CsrA. This dynamic is proposed to result in transient SepL translation and assembly of the EspA translocon.

Figure 6. SepL translation and summary model. (A) The predicted RNAfold structure of the 5′ UTR and first 89 nt of sepL CDS is shown for orientation of possible secondary structure absent in sepL truncates shown in panel B. (B) A systematic series of fusions to GFP were constructed into the sepL open reading frame to investigate sequence requirements for expression. All the fusions included the 5′ UTR and natural promoter of sepL (Table 2). The shortest fusion encompassed the native sepL promoter to the initiation codon (ATG). Successive three nucleotide extensions are numbered by how far they extend (in bases) into sepL. The final construct is the full length (FL) fusion used in previous studies (pDW6, Table 2). Expression from this series of constructs was measured in the low secretor strain Sakai. (C) Fluorescence microscopy of selected sepL-GFP constructs to illustrate transitions in sepL translation. The 81 nt fusion exhibits a clear biphasic expression pattern, with approximately half of the cells expressing GFP. (D) Flow cytometry of sepL truncates presented in panel C. (E) Graph showing the maximum ΔG values for the different truncates fused to GFP. This provides an indication of when major transitions in structure may occur between the different length fusions. (F) Model of the post-transcriptional regulation leading to transient expression of the sepL transcript. The model is described within the Discussion section of the manuscript. It is proposed that the transcript adopts a structure that initially prevents translation due to poor access to the RBS. CsrA is required for translational activation along with specific sequence in the early part of sepL (1–18 bases). SepL translation is silenced by Hfq-Spot42 binding to occlude the RBS and potentially by direct competition with CsrA. This dynamic is proposed to result in transient SepL translation and assembly of the EspA translocon.
DISCUSSION

T3S in *E. coli* O157 is subject to complex regulation as confirmed by the significant number of regulators that act on the system and marked variation in the secretion phenotype between isolates (4,46). This variability reflects multiple regulatory inputs at both transcriptional and post-transcriptional levels. We previously reported that EspA filaments, expressed from the LEE4 operon, are controlled post-transcriptionally but the basis for this was not known (25,26). Subsequently, the sRNA chaperone, Hfq, was shown to repress T3S in two studies (40,41) although a third showed activation (47). The main proposed sites for this control were Ler and GrlA/R, both key early regulators of T3S. Further post-transcriptional regulation of the LEE was identified by Bhatt et al. (27), demonstrating that CsrA negatively regulated GrlRA expression and positively regulated LEE4 at the *sepL* 5′ UTR.

The present study extends our genome-wide approach (31) to analyse the sequences interacting with Hfq under culture conditions in which the T3S system is expressed, allowing the mapping of Hfq binding across the locus of enterocyte effacement operons (LEE1–5). CRAC UV-crosslinking of Hfq confirmed previously proposed targets for regulation by Hfq adjacent to ler (LEE1) and grlRA. In addition, we identified extensive Hfq binding to LEE4 and LEE5 mRNAs, consistent with our previous report of coupled post-transcriptional regulation of these two operons (26). The CRAC data define two peaks of Hfq binding in the *sepL* 5′ UTR (peaks 703 and 702, Figure 1), with peak 703 containing a canonical ARN motif recognized by the distal face of Hfq. This recruits Spot42 to the *sepL* seed motif 43nt downstream to occlude translation initiation. Mutational analysis also confirmed the primary AUGGA motif required for CsrA activation, which overlaps with the second mapped Hfq interaction sequence (peak 703) suggesting Hfq and CsrA binding is mutually exclusive.

We propose that the combination of activation by CsrA followed by Hfq-mediated Spot42 repression are components of an autonomous OFF-ON-OFF switch or ‘toggle’ that results in transient SepL translation (Figure 6F). Under the model it is assumed that SepL translation is a prerequisite for EspA filament production, as demonstrated by others (10,30). Our *in vivo* structural probing of the *sepL* transcript indicates that the default structure formed correlates well with that based on free energy modelling (shown in Figures 2-6). In this structure, the Shine Dalgarno (SD) is sequestered in a stem and the start codon is protected at the base of this stem loop. We conclude that translation would initially be repressed following transcription based on structural constraints imposed by the mRNA structure as indicated in this study and depicted in our model (Figure 6F).

Multiple regulatory factors may contribute to translation of *sepL* and here we demonstrate that CsrA binding to SL1 is required for translation activation, most likely by opening the repressive SL2 structure to allow ribosome access to the SD. Additional factors may be required for this activation which would help explain heterogeneity in SepL and EspA expression at the single cell level observed in this and other studies (Figures 5B and 6C–D) (25,26).

Our analysis then demonstrates that translation of SepL is repressed by Spot42 chaperoned by Hfq. Spot42 is present at high levels under conditions in which cAMP concentrations are low, as *spf* (encoding Spot42) is repressed by CRP when complexed with cAMP (48–50). Therefore, Spot42 is also present in bacterial cells with an overall positive energetic state, along with CsrA. A regulatory circuit that juxtaposes CsrA and Spot42 function would then be expected to rapidly transition between ON and OFF states. We therefore propose they are part of an ON-OFF regulatory circuit acting at the translational level and this is likely to result in transient expression of SepL at the single cell level (Figure 6F). We note that a tipping point between repression and translation was identified in our 81nt translational fusion which we presume represents the mRNA oscillating between closed and open structures producing a heterogeneous population in terms of GFP expression (Figure 6C–E). This transition point matched closely with RNAfold predictions of conformations for the different modelled fusion transcripts and their respective free energies (Figure 6E). Future work will address the full structure of the mRNA and any additional factors needed to alter the stability of the stem loops demonstrated in the 5′ UTR and early part of the reading frame.

Our study, building on the previous work of Bhatt et al. (27), shows that CsrA is essential for translation and CsrA is likely to work in a manner similar to MoaA activation in *E. coli* (51) and Phz2 activation in *Pseudomonas* (52) by opening or maintaining a transcriptionally active conformation of the transcript (53). The fact that Spot42 repression of translation of both the full length SepL-GFP fusion and the shorter (51 base/17aa) fusion (Figure 3B-C), which shows high levels of expression and is predicted to represent the open structure, supports the idea that Hfq-Spot42 act on a transcriptionally ‘ON’ *sepL* mRNA. Biophysical characterization of Hfq-mRNA interactions indicate a dissociation constant (*K*ₐ) of 1–4 nM, and CsrA has been shown to bind the *sepL* 5′ UTR with a *K*ₐ of 23 nM (27) suggesting that Hfq repression would be favoured once access to the single-stranded site is available.

We had previously found that Hfq- and sRNA-binding sites were closely associated in mRNAs, potentially to facilitate dissociation of mRNA–sRNA duplexes, but also to position the mRNA seed for annealing at the lateral surface of Hfq (31,54,55). The distal face of Hfq can accommodate 18nt of RNA, and a linear spacer of 43nt before the Spot42 binding site would position the mRNA seed far from the lateral edge of Hfq. For *sepL*, it appears that extensive secondary structure may allow longer range contacts with Hfq and a second ‘ARN-less’ Hfq binding peak (702) at the Spot42 binding site indicating that this distant site is drawn in close enough to make contact with Hfq.

Hfq was also found to bind at multiple sites throughout the polycistronic LEE5 operon (that encodes *tir-cesT-eae*), with the most prominent binding sites occurring within the upstream *map* 5′ UTR, within the *tir* CDS, and at the 3′ end...
of eae. We found that Spot42 regulates the expression of Tir in the low secretor Sakai background, indicating multiple points of Spot42-dependent regulation. Recent work in Vibrio parahaemolyticus has demonstrated Spot42 regulation of the CsrF-family protein VP1680 (56) and it appears that post-transcriptional Spot42 regulation of this effector chaperone is likely conserved in EHEC. The LEE4 transcript also contained extensive sites of Hfq-binding with prominent peaks within the CDS of espD, espB, and the 3′ of orf29. The latter is in good agreement with the binding site identified for GlmZ (57) that is on the 5′ edge of this Hfq peak. Interestingly, RNA-RNA interactions 5′ of the Hfq binding site is more consistent with an sRNA than an mRNA (44). The arrangement of Hfq and sRNA binding sites at the 3′ of orf29 suggests that this site may represent a 3′ UTR sRNA and is reminiscent of the sRNA-sponge, ScrC, that is generated from the 3′ of ghlI to sponge the sRNA GcvB (58). The precise mechanism of GlmZ-orf29 regulation of LEE4 is currently unclear, but our Hfq-binding data confirms that this is a site of Hfq-dependent regulation and suggests a 3′ UTR sRNA may be involved.

In the present study we focused on strain Sakai, a low secretor, which reflects the secretion state of many wild type strains we have examined (26). Our proposed model is based on Hfq interactions measured in this strain and mutations studied in different Sakai genetic backgrounds. However, the high secretor strain (ZAP193) was also included in the current study, and while SepL translation is repressed by Spot42 in this background, there was no evidence that this led to increased EspADB translocon production in the Δsfp background. We consider that this indicates export control of the translocon apparatus is saturated in this high secretor background and so de-repressing SepL levels has no impact on filament production under these conditions.

In summary, CsrA and Spot42 are used to create a post-transcriptional toggle that we predict will lead to a burst of translation that is rapidly silenced. The ON-OFF kinetics would be tuned by the concentrations of free CsrA, and Hfq-Spot42. Spot42 is an early bacterial sRNA in evolutionarily terms and is expressed at high levels and so we anticipate that this simple post-transcriptional toggle would be a useful regulatory motif for a range of proteins required transiently in limited quantities.

FUNDING

Wellcome Trust funding [WT090231MA to D.L.G., J.J.T., D.L.T., S.P.M.]; BBSRC Institute programme grant funding to the Roslin Institute [BB/J004227/1, BB/P013740/1 to D.L.G., S.P.M.]; EASTBIO Doctoral Training Programme from the BBSRC (to A.B.W.); Wellcome Trust [091549 to S.G.]; NSFC funding [31370166, 81301474 to D.W.]; Australian National Health and Medical Research Council [APP1067241 to J.J.T.]. Funding for open access charge: University of Edinburgh/Wellcome Trust. Conflict of interest statement. None declared.

REFERENCES

1. Naylor, S.W., Gally, D.L. and Christopher Low, J. (2005) Enterohaemorrhagic E. coli in veterinary medicine. Int. J. Med. Microbiol., 295, 419–441.
2. Jerse, A.E., Yu, J., Tall, B.D. and Kaper, J.B. (1990) A genetic locus of enteropathogenic Escherichia coli necessary for the production of attaching and effacing lesions on tissue culture cells. Proc. Natl. Acad. Sci. U.S.A., 87, 7839–7843.
3. Stevens, M.P. and Frankel, G.M. (2014) The locus of enterocyte effacement and associated virulence factors of enterohemorrhagic Escherichia coli. Microbiol. Spectr., 2. doi:10.1128/microbiolspec.EHEC-0007-2013.
4. Tree, J.J., Wolfson, E.B., Wang, D., Roe, A.J. and Gally, D.L. (2009) Controlling injection: regulation of type III secretion in enterohaemorrhagic Escherichia coli. Trends Microbiol., 17, 361–370.
5. Knutson, S., Rosenshine, J., Pallen, M.J., Nisan, J., Neves, B.C., Bain, C., Wolff, C., Dougan, G. and Frankel, G. (1998) A novel EspA-associated surface organelle of enteropathogenic Escherichia coli involved in protein translocation into epithelial cells. EMBO J., 17, 2166–2176.
6. Ide, T., Laarmann, S., Greune, I., Schiller, H., Oberleithner, H. and Schmidt, M.A. (2001) Characterization of translocation pores inserted into plasma membranes by type III-secreted Esp proteins of enteropathogenic Escherichia coli. Cell. Microbiol., 3, 669–679.
7. Wuchter, C., Beinke, C., Mattes, M. and Schmidt, M.A. (1999) Insertion of EspD into epithelial target cell membranes by infecting enteropathogenic Escherichia coli. Mol. Microbiol., 31, 1695–1707.
8. O’Connell, C.B., Creasey, E.A., Knutton, S., Elliott, S., Crowther, L.J., Luo, W., Albert, M.J., Kaper, J.B., Frankel, G. and Donnenberg, M.S. (2004) SepL, a protein required for enteropathogenic Escherichia coli type III translocation, interacts with secretion component SepD. Mol. Microbiol., 52, 1613–1625.
9. Deng, W., Li, Y., Hardwidge, P.R., Frey, E.A., Pfuetzner, R.A., Lee, S., Gruenheid, S., Strynak, N.C.J., Puente, J.L. and Finlay, B.B. (2005) Regulation of type III secretion hierarchy of translocators and effectors in attaching and effacing bacterial pathogens. Infect. Immun., 73, 2135–2146.
10. Wang, D., Roe, A.J., McAteer, S., Shipston, M.J. and Gally, D.L. (2008) Hierarchical type III secretion of translocators and effectors from Escherichia coli O157:H7 requires the carboxy terminus of SepL that binds to Tir. Mol. Microbiol., 69, 1499–1512.
11. Apel, D. and Surette, M.G. (2008) Bringing order to a complex molecular machine: the assembly of the bacterial flagella. Biochem. Biophys. Acta, 1778, 1851–1858.
12. Portaliao, A.G., Tsolis, K.C., Loos, M.S., Zorzinii, V. and Economou, A. (2016) Type III secretion: building and operating a remarkable nanomachine. Trends Biochem. Sci., 41, 175–189.
13. De Lay, N. and Gottesman, S. (2009) The Crp-activated small noncoding regulatory RNA CyaR (RyeR) links nutritional status to group behavior. J. Bacteriol., 191, 461–476.
14. Rice, J.B. and Vanderpool, C.K. (2011) The small RNA SgrS controls sugar-phosphate accumulation by regulating multiple PTS genes. Nucleic Acids Res., 39, 3806–3819.
15. Klein, G. and Raina, S. (2015) Regulated control of the assembly and diversity of LPS by noncoding sRNAs. Biomed. Res. Int., 2015, 153561.
16. Vogel, J. and Papenfort, K. (2006) Small non-coding RNAs and the bacterial outer membrane. Curr. Opin. Microbiol., 9, 605–611.
17. Schiano, C.A., Koo, J.T., Schipma, M.J., Caulfield, A.J., Jafari, N. and Latham, W.W. (2014) Genome-wide analysis of small RNAs expressed by Yersinia pestis identifies a regulator of the Yap-Ysc type III secretion system. J. Bacteriol., 196, 1659–1670.
18. Colgan, A.M., Kröger, C., Diard, M., Hardt, W.-D., Puente, J.L., Sivasankaran, S.K., Hokamp, K. and Hinton, J.C.D. (2016) The impact of 18 ancestral and horizontally-acquired regulatory proteins on filament production under these conditions.
19. Wolfson, E.B., Wang, D., Roe, A.J. and Gally, D.L. (2009) SepL, a protein required for enteropathogenic Escherichia coli type III translocation, interacts with secretion component SepD. Mol. Microbiol., 31, 1695–1707.
20. Bhatt, S., Egan, M., Jenkins, V., Muche, S. and El-Fenej, J. (2016) The tip of the iceberg: on the roles of regulatory small RNAs in the virulence of enterohaemorrhagic and enteropathogenic Escherichia coli. Front. Cell Infect. Microbiol., 6, 105.
21. Zeng, Q., McNally, R.R. and Sundin, G.W. (2013) Global small RNA chaperone Hfq and regulatory small RNAs are important virulence regulators in Erwinia amylovora. J. Bacteriol., 195, 1706–1717.

Nucleic Acids Research, 2018 15

Downloaded from https://academic.oup.com/nar/advance-article-abstract/doi/10.1093/nar/gky096/4844049 by Edinburgh University user on 16 February 2018
31. Tree,J.J., Granneman,S., McAteer,S.P., Tollervey,D. and Gally,D.L. (2010) Characterization of SepL of enterohemorrhagic Escherichia coli. Mol. Microbiol., 77, 359–365.  
32. Fernandez-Brando,R.J., Yamaguchi,N., Tahoun,A., McAteer,S.P., Gillespie,T., Wang,D., Argyle,S.A., Palermo,M.S. and Gally,D.L. (2014) Generation of gene deletions and gene replacements in Escherichia coli O157:H7: using a temperature sensitive allelic exchange system. Biol. Proced. Online, 8, 153–162.  
33. Funahashi,T. (2015) The small RNA Spot 42 regulates the expression of the LEE4 operon in enterohaemorrhagic Escherichia coli. Mol. Microbiol., 97, 10655–10659.  
34. Holmqvist,E., Wright,P.R., Li,L., Bischler,T., Barquist,L., Kendall,M.M., Gruber,C.C., Rasko,D.A., Hughes,D.T. and Mellies,J.L. and Lorenzen,E. (2014) Enterohemorrhagic Escherichia coli Virulence Gene Regulation. Microbiol. Spectr., 2, doi:10.1128/microbiolspec.EHEC-00004-2013.  
35. Lang,J.C., Babitzke,P. and Romeo,T. (2005) CsrA post-transcriptionally represses pgaABCD, a carbon catabolite regulator of the locus of enterocyte effacement pathogenicity island. Mol. Microbiol., 55, 153–162.  
36. Spike,L., Wang,J. and Ibarra,J.A. (2012) Lysogeny with Shiga toxin 2-encoding Pseudomonas aeruginosa. Nat. Biotechnol., 30, 5452–5459.  
37. Blomfield,I.C., Vaughn,V., Rest,R.F. and Eisenstein,B.I. (1991) Transcriptome of enterohemorrhagic Escherichia coli. Mol. Microbiol., 7, 1447–1457.  
38. Holmqvist,E., Wright,P.R., Li,L., Bischler,T., Barquist,L., Reinhardt,R., Backofen,R. and Vogel,J. (2014) Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo. EMBO J., 33, 991–1011.  
39. Busch,A., Richter,A.S. and Backofen,R. (2008) IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. Bioinformatics, 24, 2849–2856.  
40. Dahan,S., Knutton,S., Shaw,R.K., Crepin,V.F., Dougan,G. and Romeo,T. (2011) Hfq virulence regulation in enterohemorrhagic Escherichia coli O157:H7 strain 86-24. J. Bacteriol., 193, 6843–6851.  
41. Urban,J.H. and Vogel,J. (2007) Translational control and target recognition by Escherichia coli small RNAs in vivo. Nucleic Acids Res., 35, 1018–1037.  
42. Ren,B., Shen,H., Lu,Z.J., Liu,H. and Xu,Y. (2014) The phzA2-G2 transcript exhibits direct RmA-mediated activation in Pseudomonas aeruginosa M18. PLoS One, 9, e89653.  
43. Dutta,R.K., Finn,W. and Woodson,S.A. (2014) Structural model of an mRNA in complex with the bacterial chaperone Hfq. Proc. Natl. Acad. Sci. U.S.A., 111, 17134–17139.  
44. Kamal,A., Vazquez,A., Barba,J., Ibarra,J.A., O'Donnell,P., Metalnikov,P. and Sperandio,V. (2014) Hfq affects the expression of the LEE4 operon in enterohaemorrhagic Escherichia coli. Mol. Microbiol., 97, 10639–10650.  
45. Ren,B., Shen,H., Lu,Z.J., Liu,H. and Xu,Y. (2014) The phzA2-G2 transcript exhibits direct RmA-mediated activation in Pseudomonas aeruginosa M18. PLoS One, 9, e89653.  
46. Dahan,S., Knutton,S., Shaw,R.K., Crepin,V.F., Dougan,G. and Romeo,T. (2011) Hfq virulence regulation in enterohemorrhagic Escherichia coli O157:H7 strain 86-24. J. Bacteriol., 193, 6843–6851.  
47. Urban,J.H. and Vogel,J. (2007) Translational control and target recognition by Escherichia coli small RNAs in vivo. Nucleic Acids Res., 35, 1018–1037.  
48. Ren,B., Shen,H., Lu,Z.J., Liu,H. and Xu,Y. (2014) The phzA2-G2 transcript exhibits direct RmA-mediated activation in Pseudomonas aeruginosa M18. PLoS One, 9, e89653.  
49. Vazquez,A., Barba,J., Ibarra,J.A., O'Donnell,P., Metalnikov,P. and Sperandio,V. (2014) Hfq affects the expression of the LEE4 operon in enterohaemorrhagic Escherichia coli. Mol. Microbiol., 97, 10639–10650.  
50. Dahan,S., Knutton,S., Shaw,R.K., Crepin,V.F., Dougan,G. and Romeo,T. (2011) Hfq virulence regulation in enterohemorrhagic Escherichia coli O157:H7 strain 86-24. J. Bacteriol., 193, 6843–6851.  
51. Urban,J.H. and Vogel,J. (2007) Translational control and target recognition by Escherichia coli small RNAs in vivo. Nucleic Acids Res., 35, 1018–1037.  
52. Ren,B., Shen,H., Lu,Z.J., Liu,H. and Xu,Y. (2014) The phzA2-G2 transcript exhibits direct RmA-mediated activation in Pseudomonas aeruginosa M18. PLoS One, 9, e89653.  
53. Vazquez,A., Barba,J., Ibarra,J.A., O'Donnell,P., Metalnikov,P. and Sperandio,V. (2014) Hfq affects the expression of the LEE4 operon in enterohaemorrhagic Escherichia coli. Mol. Microbiol., 97, 10639–10650.  
54. Dahan,S., Knutton,S., Shaw,R.K., Crepin,V.F., Dougan,G. and Romeo,T. (2011) Hfq virulence regulation in enterohemorrhagic Escherichia coli O157:H7 strain 86-24. J. Bacteriol., 193, 6843–6851.  
55. Urban,J.H. and Vogel,J. (2007) Translational control and target recognition by Escherichia coli small RNAs in vivo. Nucleic Acids Res., 35, 1018–1037.  
56. Ren,B., Shen,H., Lu,Z.J., Liu,H. and Xu,Y. (2014) The phzA2-G2 transcript exhibits direct RmA-mediated activation in Pseudomonas aeruginosa M18. PLoS One, 9, e89653.  
57. Vazquez,A., Barba,J., Ibarra,J.A., O'Donnell,P., Metalnikov,P. and Sperandio,V. (2014) Hfq affects the expression of the LEE4 operon in enterohaemorrhagic Escherichia coli. Mol. Microbiol., 97, 10639–10650.  
58. Dahan,S., Knutton,S., Shaw,R.K., Crepin,V.F., Dougan,G. and Romeo,T. (2011) Hfq virulence regulation in enterohemorrhagic Escherichia coli O157:H7 strain 86-24. J. Bacteriol., 193, 6843–6851.  
59. Urban,J.H. and Vogel,J. (2007) Translational control and target recognition by Escherichia coli small RNAs in vivo. Nucleic Acids Res., 35, 1018–1037.  
60. Dahan,S., Knutton,S., Shaw,R.K., Crepin,V.F., Dougan,G. and Romeo,T. (2011) Hfq virulence regulation in enterohemorrhagic Escherichia coli O157:H7 strain 86-24. J. Bacteriol., 193, 6843–6851.  
61. Blomfield,I.C., Vaughn,V., Rest,R.F. and Eisenstein,B.I. (1991) Transcriptome of enterohemorrhagic Escherichia coli. Mol. Microbiol., 7, 1447–1457.  
62. Merlin,C., McAteer,S. and Masters,M. (2002) Tools for characterization of Enterococci of the lists of function. J. Bacteriol., 184, 4573–4581.