Biochimie 117 (2015) 129–137

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Research paper

Deletion of the Sm1 encoding motif in the *lsm* gene results in distinct changes in the transcriptome and enhanced swarming activity of *Haloferax* cells

Lisa-Katharina Maier [a], Juliane Benz [a], Susan Fischer [a], Martina Alstetter [a], Katharina Jaschinski [b], Rolf Hilker [c], Anke Becker [d], Thorsten Allers [e], Jörg Soppa [b], Anita Marchfelder [a, *]

[a] Biology II, Ulm University, 89069 Ulm, Germany
[b] Institute of Molecular Biosciences, Goethe-University, 60438 Frankfurt, Germany
[c] Institute of Medical Microbiology, Justus-Liebig University, 35392 Giessen, Germany
[d] LOEWE Center for Synthetic Microbiology, Faculty of Biology, Philipps-University Marburg, 35032 Marburg, Germany
[e] School of Life Sciences, University of Nottingham, Queen's Medical Centre, Nottingham NG7 2UH, United Kingdom

ARTICLE INFO

Article history:
Received 5 December 2014
Accepted 26 February 2015
Available online 6 March 2015

Keywords:
Lsm
Hfq
Haloferax volcanii
Archaea
L37.eR

ABSTRACT

Members of the Sm protein family are important for the cellular RNA metabolism in all three domains of life. The family includes archaeal and eukaryotic Lsm proteins, eukaryotic Sm proteins and archaeal and bacterial Hfq proteins. While several studies concerning the bacterial and eukaryotic family members have been published, little is known about the archaeal Lsm proteins. Although structures for several archaeal Lsm proteins have been solved already more than ten years ago, we still do not know much about their biological function, however one can confidently propose that the archaeal Lsm proteins will also be involved in RNA metabolism. Therefore, we investigated this protein in the halophilic archaeon *Haloferax volcanii*. The *Haloferax* genome encodes a single Lsm protein, the *lsm* gene overlaps and is co-transcribed with the gene for the ribosomal L37.eR protein. Here, we show that the reading frame of the *lsm* gene contains a promoter which regulates expression of the overlapping *rpl37R* gene. This *rpl37R* specific promoter ensures high expression of the *rpl37R* gene in exponential growth phase. To investigate the biological function of the Lsm protein we generated a *lsm* deletion mutant that had the coding sequence for the Sm1 motif removed but still contained the internal promoter for the downstream *rpl37R* gene. The transcriptome of this deletion mutant was compared to the wild type transcriptome, revealing that several genes are down-regulated and many genes are up-regulated in the deletion strain. Northern blot analyses confirmed down-regulation of two genes. In addition, the deletion strain showed a gain of function in swarming, in congruence with the up-regulation of transcripts encoding proteins required for motility.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Sm and like-Sm (Lsm) proteins belong together with the Hfq proteins to the Sm superfamily of proteins and are involved in several pathways of the cellular RNA metabolism [1–4]. Hfq proteins can be found in bacteria and archaea, whereas Sm and Lsm proteins are present in eukaryotes and archaea [1,3]. On the amino acid sequence level Hfq and Lsm proteins are not very similar, but they show striking similarities concerning the tertiary and quaternary structure [3,5,6]. Characteristic for this protein family is a bipartite signature sequence, the Sm domain, which consists of two segments, the Sm1 and Sm2 motif (Fig. 1) [6]. In addition these proteins form homomeric or heteromeric rings consisting of six or seven proteins [3]. The Hfq protein was initially identified as host factor required for the replication of Qβ RNA bacteriophage [7–9] but it is nowadays known for having a plethora of functions in the cell [5]. Because of these observations it was suggested that the Hfq protein is an ancient representative of the Sm superfamily of

* Corresponding author.
E-mail address: anita.marchfelder@uni-ulm.de (A. Marchfelder).

http://dx.doi.org/10.1016/j.biochi.2015.02.023
0300-9084/© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

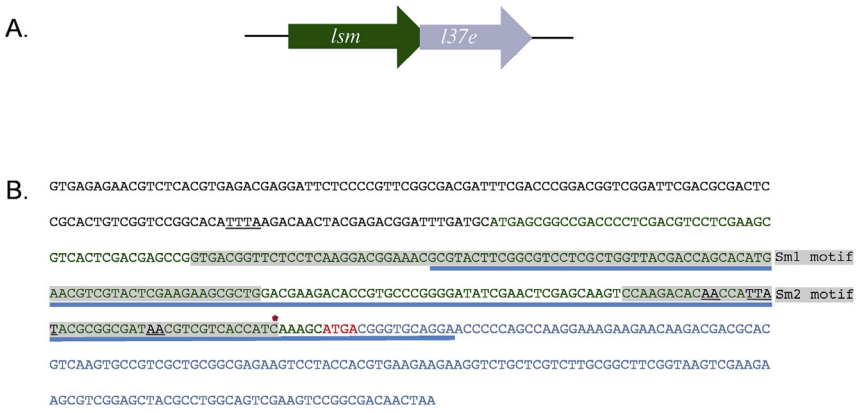

Fig. 1. The *lsm* and *rpl37R* genes overlap. A. The genomic organisation of *lsm* (ORF is shown in green) and *rpl37R* (ORF is shown in blue) is shown. **B**. Both genes overlap by four nucleotides (in red). The potential transcription start site of the *rpl37R* gene as identified by high throughput sequencing is marked with a red asterisk. The two conserved motifs Sm1 and Sm2 are shaded in grey. The Sm2 motif contains a promoter motif for *rpl37R* (underlined and in black: AA-TTAT-AA). To leave the internal *rpl37R* promoter intact, only the coding region for the Sm1 domain was deleted, generating deletion strain ΔSm1. The sequence used for the promoter-reporter gene fusion is underlined in blue.

proteins being less specialised [10]. Approximately half of the bacterial organisms including many pathogens encode an Hfq protein. The Hfq protein has been shown to be a key player in sRNA regulated gene expression [5,6,11]. It interacts as chaperone protein with small regulatory RNAs as well as with their target mRNAs assisting in their interaction [12]. In addition, it is involved in mRNA decay by stimulating polyadenylation [13,14] and it can also repress mRNA translation [15,16]. The Hfq protein forms highly stable hexamers that bind preferentially to A-U rich sequences [17]. It is highly conserved in bacteria [18–20] but also found in a few archaea, like *Methanocaldoccus jannaschii* [10,21].

Pleiotropic phenotypes with increased sensitivity to stress conditions, increased cell sizes and decreased growth rates were seen in an *hfq* insertion mutant in *Escherichia coli* [22]. At least in part these phenotypes might be due to the fact that Hfq is required for the function of regulatory sRNAs, like RhyB, OxyS, Spot42, RprA and DsrA [23–26]. Whereas much is known about the function of the Hfq protein in the gram negative *E. coli*, comparatively little is known about the function of Hfq in gram positive bacteria. In *Bacillus subtilis*, deletion of the *hfq* gene had no global effects on the transcriptome and affected the expression of only six out of more than 100 sRNAs, suggesting only a minor influence of the Hfq protein on sRNA and mRNAs in *Bacillus* [27]. The absence of the Hfq protein had only an impact on a few specific regulons and a toxin encoding gene [27]. In pathogenic bacteria Hfq was shown to play an important role being required for fitness and virulence [28,29]. In eukaryotes at least 18 different Sm and Lsm proteins are present, that are involved in mRNA degradation, mRNA decapping, RNA stabilisation, mRNA splicing, telomere maintenance and histone maturation [6,30]. The proteins form different hetero-heptameric complexes in the cell. For instance the Lsm proteins can already form five different complexes [6]. Currently we know most about the structure and function of the Sm proteins involved in splicing, since this was the first characterized function of Sm proteins [31].

Although structures of the archaeal Lsm proteins have been solved already more than ten years ago [32–34], not much is known about the biological function of the archaeal Lsm proteins. Since Lsm proteins are found in archaea and eukaryotes they might have been present in the common ancestor shared by archaea and eukaryotes. The presence of snoRNAs in archaea [35] and the observation that several eukaryotic proteins evolved from archaea-related precursors [36] support this hypothesis. Still the question remains, whether archaeal Lsm proteins act as scaffolds for the assembly of complex snRNPs like in eukaryotes or whether they act as chaperones for small regulatory RNAs like in bacteria [31]. Archaea can encode up to three different types of Lsm proteins (Lsm 1–3) [37–39]. The Lsm1 proteins bind RNA [34,40,41] and form heptamers [32–34,40], while the Lsm2 proteins from *Archaeoglobus fulgidus* have been shown to form either hexameric [42] or heptameric [40] complexes. Complexes consisting of 14 Lsm3 proteins have been reported for the Lsm3 protein from *Pyrobaculum aerophilum* [38], a similar 14mer complex has also been observed for Lsm1 proteins [43].

In the halophilic archaeon *Haloferax volcanii*, a single Lsm protein is encoded, which allows easy genetic analysis of the biological function of the Lsm protein. Recently, the sRNAs of *Haloferax* were investigated [44,45] and it was shown that the *Haloferax* Lsm protein binds to sRNAs, poly(U)-RNA and tRNAs *in vitro* [46]. In addition co-immunoprecipitation identified a plethora of proteins as binding partners for the *Haloferax* Lsm protein [46]. Furthermore, the deletion strain, that had the complete *lsm* gene removed, was viable, showing that the Lsm protein is not essential. The Lsm protein is very often encoded immediately upstream or even overlapping with the gene for the 50S ribosomal protein L37.eR [31]. Here, we show that the gene for the *Haloferax* Lsm protein contains a promoter for the downstream encoded L37.eR protein gene. Deletion of the Sm1 motif encoding sequence in the *lsm* gene results in clear changes in the transcriptome and in enhanced swarming activity of *Haloferax* cells.

2. Materials and methods

2.1. Strains and culture conditions

H. volcanii strain H119 (Δ*pyrE2*, Δ*trpA*, Δ*leuB*) [47], H555 [48] and ΔSm1 (Δ*pyrE2*, Δ*trpA*, Δ*leuB*, ΔSm1) (construction of mutant strains is described below) were grown aerobically at 45 °C in Hv-YPC or Hv-Ca medium [47].

2.2. Generation of the Sm1 deletion strain

The coding region for the Sm1 domain was removed in strain Δ*Sm1* using the pop-in/pop-out method [47,49]. Using chromosomal DNA from *H. volcanii* as template, the upstream and downstream region of the Sm1 coding region were amplified by PCR using primers Lsm37.u1 and Lsm37.u2 and Lsm37.d1 and L3sm7.d2, respectively (for primer sequences see Supplementary Table 4). The resulting fragments Lsm37up and Lsm37down are each 1 kb long. Fragment Lsm37up was digested with *Kpn*I and *Bam*HI and subsequently cloned into the integrative pTA131 (digested with *Kpn*I

and *Bam*HI), yielding plasmid pTA131-Lsm37up. pTA131-Lsm37up was digested with *Not*I and *Bam*HI to ligate fragment Lsm37down (digested with *Not*I and *Bam*HI) into it, yielding plasmid pTA131-Lsm37up + down. *Haloferax* strain H119 was transformed with this plasmid, and cells, that had the plasmid integrated (pop-in clones), which contains the *pyrE2* marker, were selected by growth on uracil free medium. Pop-out was induced by plating cells on a medium containing 5-fluoro-orotic acid (5-FOA) and uracil. Resulting pop-out clones were analysed with Southern blots as described [50], with the following modifications. Chromosomal DNA isolated from wild type and deletion strain was digested with *Not*I and *Bam*HI and 10 μg of the resulting DNA fragments were separated on an 0.8% agarose gel and transferred to a nylon membrane (Hybond™-N, GE-healthcare). Hybridisation probe Smwt12 was generated by PCR with primers Smwt1 and Smwt2 and genomic DNA from *Haloferax* as template, which was radioactively labelled using the random prime kit Readiprime™ II (GE Healthcare).

2.3. High throughput sequencing of RNA from wild type and deletion strains ΔSm1

Haloferax wild type cells (H119) and deletion strain ΔSm1 were grown to an OD of 0.5. RNA was isolated, fragmented to approx. 200 nucleotide long molecules using the Ambion® RNA fragmentation reagents (AM8740, Applied Biosystems) and sent to LGC (LGC Limited, UK) for cDNA preparation and high throughput sequencing. Library generation for the Illumina HiSeq 2000 sequencing was carried out according to the Illumina mRNA sample preparation guide (version Sept. 2009, Illumina, Inc., San Diego, USA). In short, fragmented RNA of an average size of 75–200 nucleotides was applied to first strand cDNA synthesis using reverse transcriptase and random primers followed by second strand synthesis to remove the RNA template and to synthesize a replacement strand generating double-stranded cDNA. The cDNA sample was end polished, the 3′ ends were adenylated and the adaptors added to the ends of the fragments by ligation. Fragments ranging in size from 175 bp to 300 bp were isolated from an agarose gel, purified using the MinElute Gel Extraction Kit (Qiagen, Hilden, Germany) and finally enriched by PCR. Library concentration was measured using the Qubit 2.0 fluorometer and the Agilent Bioanalyzer (Thermo/Life Technologies GmbH, Darmstadt, Germany; Agilent Technologies Sales and Service GmbH, Waldbronn, Germany). One library per sample was constructed. The HiSeq PE Cluster Generation Kit was used to load and hybridize each library to 1 channel of a flow cell using the cBot (Illumina, Inc., San Diego, USA). Sequencing of 2× 100 bp was performed using the TruSeq SBS Kit – HS chemistry (200 cycles) on a HiSeq 2000 resulting in 160 Mio. raw single reads (Illumina, Inc., San Diego, USA). For expression analysis of wild type and mutant samples the reads (2 × 100 bp paired-end) from Illumina HiSeq 2000 sequencing were mapped with SOAPaligner (Release 2.20, 08-13-2009, http://soap.genomics.org.cn/soapaligner.html) against the reference genome using standard settings. Unmapped reads were aligned using Blat (http://genome.cshlp.org/content/12/4/656). Taking all aligned reads together for expression profiling, the RPKM values were calculated as described by Mortazavi et al. (http://www.nature.com/nmeth/journal/v5/n7/full/nmeth.1226.html), which allows comparing expression values of gene with different length and sequencing depth.

2.4. Northern analyses

Total RNA was isolated from *H. volcanii* cells as described [51]. After separation of 10 μg RNA (total RNA) on 0.8% denaturing agarose gels, RNA molecules were transferred to nylon membranes (Hybond-N+, GE Healthcare) and incubated with probes against the mRNAs for the following proteins: Lsm, L37.eR, β-galactosidase, glucan 1,4-α-glucosidase and ABC-type sugar transport protein. For the detection of the Lsm, L37.eR and β-galactosidase mRNAs, PCR fragments were generated from *Haloferax* genomic DNA with the following primers: Lsm#11/Lsm#12 (Lsm), L37#11/L37#12 (L37.eR) and bgah#1/bgah#2 (β-galactosidase). For detection of the ABC-type sugar transport protein (permease) mRNA (HVO_A0146) and the Glucan 1,4-α-glucosidase (HVO_A0149) mRNA DNA probes were generated using PCR with oligos ABC#1 and ABC#2 and Glucan#1 and Glucan#2, respectively. The resulting PCR products were labelled using α-^{32}P-dCTP and random prime kit Readiprime™II (GE Healthcare). For the analyses of sRNAs in the ΔSm1 strain and wild type *Haloferax* cells RNAs were isolated from cells grown to an OD_{650} of 0.5 and 1.5. RNA was isolated as described [51] and separated using 8% PAGE. After transfer of the RNA to nylon membranes, the membranes were hybridized with probes against sRNAs 30, 34, 45 and 132 [45].

2.5. Promoter reporter gene fusion experiments

To allow the use of a plasmid carrying the β-galactosidase gene of *Haloferax alicantei* (bgaHa), a *H. volcanii* strain (H555) was generated with a deletion of the *H. volcanii* β-galactosidase gene (*bgaH*). This chromosomal deletion prevents homologous recombination between the two similar β-galactosidase genes. H555 is a derivative of H26 [47], it was generated using the *bgaH* deletion plasmid pTA617. To construct pTA617, a HindIII-BsrGI fragment of pTA128 [52] containing the *bgaH* gene was cloned in pTA131 [47] at Asp718 and NotI sites (blunt-ended) to generate pTA506, then an EcoRV-BamHI fragment containing the *bgaH* gene was excised. pTA919 is shuttle vector for *E. coli* and *H. volcanii*, carrying the *pyrE2* marker for selection and a promoter-less *bgaHa* gene for transcription analysis [48]; promoter regions are inserted into pTA919 at the NdeI restriction site located at the start codon of *bgaHa*. To construct pTA919, the *bgaHa* gene [53] was amplified from pTA128 [52] using PCR primers bgaNde2 and bgaR, and cloned in pTA230 [47] at ClaI and XbaI sites. A KpnI fragment containing the L11e transcription terminator of pTA425 [54] was inserted upstream of the promoter-less *bgaHa* gene. The promoter region for the *rpl37R* gene was amplified from *Haloferax* genomic DNA with PCR using the following primers: ProL37P#2 and ProL37P#3. After digestion with ApaI and NdeI the resulting DNA fragments were cloned into pTA919 (digested with ApaI and NdeI) yielding the *rpl37R* gene promoter fragment *bgaHa* fusions. For verification the resulting plasmid pTA919-L37#2 was sequenced. As control plasmid, a synthetic *Haloferax* promoter (Anice Sabag-Daigle and Charles J. Daniels, in preparation) was cloned upstream of the β-galactosidase gene, yielding pTA919-psyn. As additional control the vector pTA919 without the addition of a promoter was used. Subsequently, strain H555 was transformed with these plasmids. To analyse the mRNA levels, RNA was isolated as described above. For activity tests with β-galactosidase soluble protein extracts were isolated and subjected to the o-Nitrophenyl-β-D-galactopyranosid (ONPG) assay [53]. Shortly, cells were harvested by centrifugation and after resuspension in bgaHa buffer (2.75 M NaCl, 50 mM Tris–HCl pH 7.2, 10 μM $MnCl_2$, 1 μM DTT) and subsequent sonication the cell lysate was cleared by centrifugation at 18,000× *g*. The protein concentration of each sample was determined using Roti®-Quant (Carl Roth GmbH&Co.KG) according to manufacturer's protocol. 200 μl cleared cell lysate were mixed with 700 μl bgaH buffer and the reaction was started by addition of 150 μl ONPG solution (6 mg/ml in bgaHa-buffer). Absorbance was measured at 420 nm at RT. In the control reaction the cell lysate was replaced by 200 μl double distilled water. The amount of substrate turned over was calculated

as described [55,56]: $\Delta A/\Delta t = \varepsilon$ x d x $\Delta c/\Delta t$ (d = 0.4 cm and ε for ONP at λ_{420} is 3300 M^{-1} cm^{-1}). One unit of BgaH activity is the amount of enzyme that catalyses the hydrolysis of 1 μmol ONPG per minute.

2.6. RNASeq data for the lsm gene

RNA was isolated from *Haloferax* strain H119 (grown to an OD_{650} of 0.8) and separated into two fractions of each 40 μg (fraction A and B). Ribosomal RNA was removed from RNA fraction A using the RiboZero kit (Epicentre). Both RNA fractions were sent to vertis Biotechnologie AG (Martinsried) for sequencing. RNA fraction B was treated with terminal exonuclease (TEX) to remove all 5′-monophosphorylated RNAs (including rRNAs). cDNA was prepared and sequenced using an Illumina HiSeq 2000. The resulting cDNA libraries were termed -TEX for fraction A (without treatment with terminal exonuclease but rRNA removed with RiboZero) and +TEX for RNA fraction B (treated with terminal exonuclease). Mapping of the reads has been done with bwa using standard parameters after quality trimming [57]. Visualization for the figure and further analysis was performed with ReadXplorer [58]. The coverage of the +TEX library in Fig. 2 has been normalized by a factor of 0.55 to account for the much higher number of mapped reads in that data set (~45,000,000 in +TEX versus ~25,000,000 in the −TEX library).

2.7. Swarm plate assay

Precultures were grown in synthetic medium with casamino acids as carbon source to early exponential growth phase (OD_{600} = 0.3 ± 0.1). To equalize cell densities, cells were collected by centrifugation, washed once with basal salts, and resuspended in basal salts to yield an OD_{600} of 0.375. Swarm plates contained 25 ml medium with 0.3% (w/v) agar and glucose or glycerol as carbon source. An inoculum of 2 μl cell suspension of deletion mutant ΔSm1 or the parent strain H119, respectively, was added to the middle of swarm plates. The plates were sealed in a plastic bag and incubated at 42 °C. The swarming radii were measured daily.

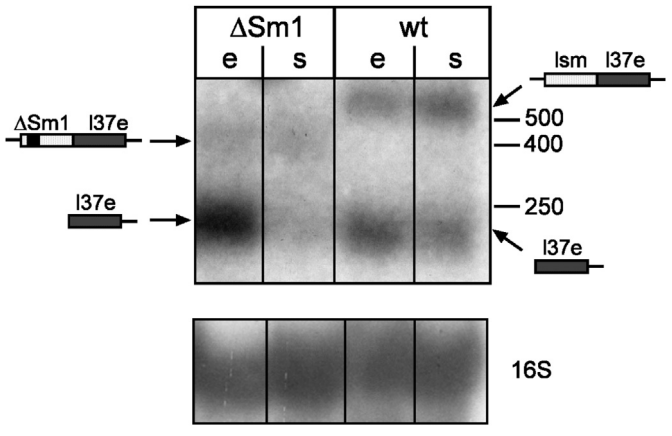

Fig. 2. Expression of the *lsm* and *rpl37R* genes. A Northern blot hybridised with a probe against the *rpl37R* mRNA shows that the *lsm* and *rpl37R* genes are expressed together as a dicistronic mRNA in exponential as well as stationary phase (lanes wt, RNA of about 500 nts) (see also Supplementary Fig. 1). The deletion strain ΔSm1 generates a slightly shorter mRNA due to the deletion of the Sm1 domain. An additional shorter RNA is detected with only about 200 nucleotides that corresponds to a monocistronic *rpl37R* transcript. Lanes wt: RNA isolated from wild type *Haloferax* cells, lanes ΔSm1: RNA isolated from the deletion strain; lanes e and s: RNA isolated from cells grown to exponential and stationary phase, respectively. The mRNAs are shown schematically at the sides, a size marker is shown at the right in nucleotides. Lower panel 16S: the membrane was hybridised with a probe against the 16S rRNA.

Three biological replicates were performed, and average values and standard deviations were calculated.

2.8. Phenotypic characterization

For phenotypic comparison, parent strain *H. volcanii* H119 and the deletion mutant ΔSm1 were grown under different conditions in 96 well microtiter plates as described previously [59]. Growth was monitored at 600 nm using a microtiter plate reader (Spectramax 340, Molecular Devices, Sunnyvale, CA). Three biological replicates were performed and the average values of growth yield, growth rate, the length of lag-phase and their standard deviations were calculated.

3. Results and discussion

3.1. Expression of the rpl37R gene is driven by an additional promoter located in the lsm reading frame

The genes for the Lsm protein and the 50S ribosomal L37.eR protein overlap by four nucleotides in *Haloferax volcanii* (Fig. 1). This overlap is conserved in all haloarchaeal genomes analysed and several other archaeal genomes outside of the haloarchaeal class (Supplementary Table 1). Overlapping gene pairs occur frequently in haloarchaea, e.g. *H. volcanii* contains 352 gene pairs that are annotated to overlap by four nucleotides. Northern analysis using a probe against the *rpl37R* mRNA revealed two RNAs, one longer molecule of about 500 nucleotides and a shorter one of about 200 nucleotides, confirming earlier observations, that two *rpl37R* transcripts of different sizes exist (Fig. 2 and Supplementary Fig. 1) [46]. The reading frames for the Lsm and L37.eR proteins together are 404 nucleotides long, thus the long RNA corresponds to the dicistronic mRNA, whereas the shorter RNA corresponds to a monocistronic *rpl37R* mRNA, since the reading frame for the L37.eR protein is 177 nucleotides (Fig. 2, lanes wt). Upon using a probe against the *lsm* RNA, only a single transcript is detected, corresponding in length with approximately 500 nucleotides to the dicistronic mRNA of *lsm* and *rpl37R* (Supplementary Fig. 1). These data suggest that either an additional promoter for *rpl37R* expression is located in the *lsm* gene or that the dicistronic transcript is processed yielding a stable *rpl37R* transcript and a labile *lsm* transcript. According to a system biology study analysing transcription factor binding in *Halobacterium salinarum,* many transcription factor binding events were observed in coding regions, suggesting that transcription initiation is not exclusively happening in intergenic regions but very often also in coding regions [60].

RNASeq performed with *Haloferax* RNA revealed a transcription start site in the *lsm* reading frame located 6 nucleotides upstream of the *rpl37R* ATG start codon, generating an almost leaderless mRNA (Figs. 1B and 3) (Hilker et al. in preparation). In addition, a high amount of 5′ monophosphorylated RNA is also detected. A processing at this site is highly unlikely since this would remove 10 nucleotides from the *lsm* ORF at the respective 3′ end mRNA. Therefore this observation might suggest that dephosphorylation of the primary *rpl37R* transcript is very fast. Closer inspection showed, that upstream of the *rpl37R* gene a potential promoter could be located in the *lsm* gene (Fig. 1). The potential promoter motif consists of [AA … TTAT …… AA] and matches the conserved elements determined for *Haloferax*: transcription factor B recognition element (BRE) sequence (position -35), the TATA box (position -27/-28) and the WW element (position -10/-11) (W = A or T) [61]. Therefore, the existence of an additional promoter seemed to be likely, and its putative presence was analysed experimentally.

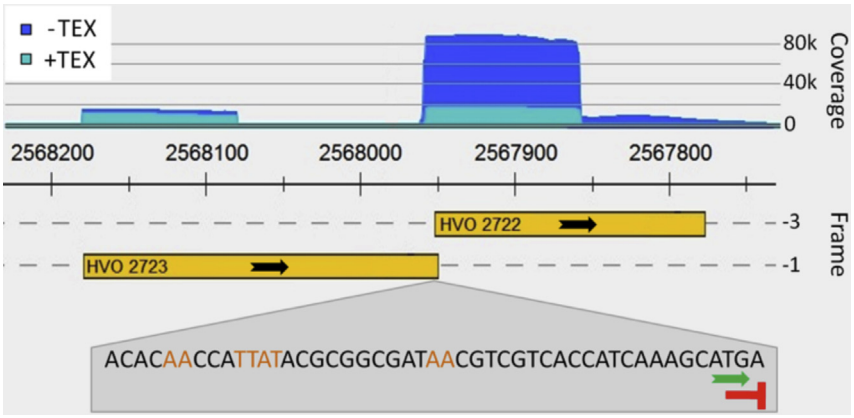

Fig. 3. RNASeq analysis for the *rpl37R* gene. Although the amount of 5′ monophosphorylated RNAs detected upstream of the *rpl37R* start codon is quite high primary ends are also detected suggesting that a promoter is located here. The genomic region of both genes, *lsm* and *rpl37R*, is shown, open reading frames are indicated in orange, the orientation of the genes is depicted by black arrows. Read coverage data from two RNASeq libraries ("+TEX" treated with TEX to accumulate primary transcripts, "−TEX" without TEX treatment to obtain all RNA 5′ ends, see Material and Methods) is shown for the genomic area containing both genes and reveals activity of the *lsm* promoter and strong activity of the *rpl37R* promoter. Cyan indicates the coverage of the TEX treated library (~11.000 in the *lsm* and ~17.000 in the *rpl37R* transcript start region). Blue represents expression data (coverage) from the library without TEX treatment (~15.000 in the *lsm* and ~88.000 in the *rpl37R* transcript start region). Coverage difference between +TEX and −TEX is ~3.800 reads in the *lsm*, and ~71.000 in the *rpl37R* transcript start region. The zoom shows the internal promoter sequence (orange), the start codon of *rpl37R* (green arrow) and the stop codon of *lsm* (red mark).

3.2. Analysis of promoter activities with a reporter gene

To confirm that the *lsm* ORF contains an active promoter for the downstream *rpl37R* gene, we cloned the promoter region located upstream of the *rpl37R* frame (Fig. 1B) in front of the reporter gene β-galactosidase using the vector pTA919. In addition, we wanted to confirm the observation made in the northern blot analysis that the signal strength of the *rpl37R* mRNA is higher in the exponential phase than in the stationary phase (Fig. 2). The *Haloferax* strain H555 was transformed with the construct and soluble protein extracts as well as RNAs were extracted from cultures grown to exponential and stationary phase. Transcripts were analysed using northern blots, which were hybridised with a probe against the β-galactosidase mRNA (Fig. 4A). The northern blot shows that the *rpl37R* upstream fragment can initiate transcription. Expression regulated by the *rpl37R* promoter varies during growth: in exponential phase expression is much stronger than in stationary phase. Since the northern blot only shows the steady state level of the RNAs, this might be due either to different promoter strength in the different growth phases or to the varying stability of the β-galactosidase mRNA in exponential and stationary phase. To investigate whether the effect is likewise visible on the protein level we measured and compared the activity of the protein extracts from the cultures with the promoter constructs grown to exponential and stationary phase (Fig. 4B). The highest β-galactosidase activity was found in cultures transformed with the *rpl37R* promoter construct. This observation differs from the result obtained on the RNA level, where the promoter psyn is more active than the *rpl37R*

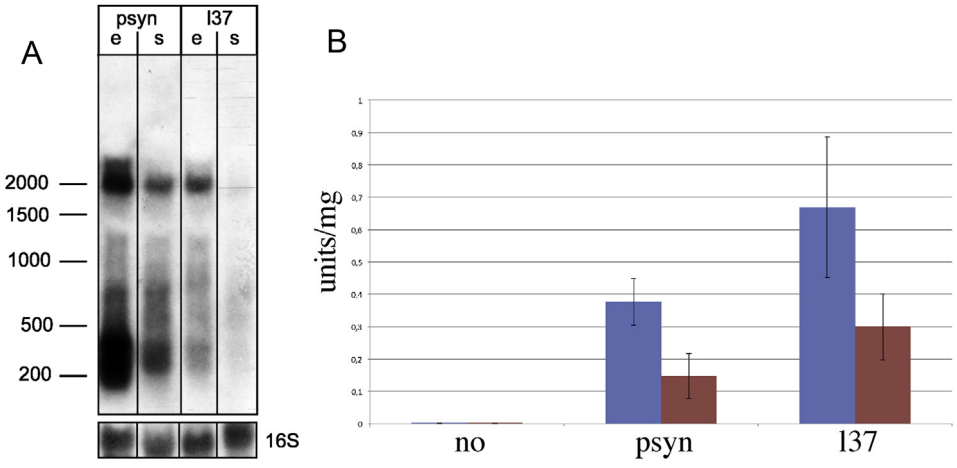

Fig. 4. Analysis of the promoter-gene-fusion. A. Northern analysis. RNA isolated from *Haloferax* cells transformed with the *rpl37R* promoter fusion plasmid was separated on 8% PAGE and transferred to a membrane. As control RNA was isolated from a strain that expressed the β-galactosidase gene from a synthetic promoter (psyn). Hybridisation with a probe against the β-galactosidase mRNA revealed the expression levels, the β-galactosidase mRNA has a size of 1992 nucleotides. In all four RNAs a signal for the β-galactosidase mRNA can be detected. For both strains the signal is stronger in RNA isolated from cells grown to exponential phase than in RNA from cells grown to stationary phase. Lower panel 16S: the membrane was hybridised with a probe against the 16S rRNA. Lanes psyn: expression was driven by the synthetic promoter; lanes l37 expression from the *rpl37R* promoter; lanes e and s: RNA isolated from cells grown to exponential phase and stationary phase, respectively. An RNA size marker is shown at the left in nucleotides. **B. Analysis of β-galactosidase activity.** Soluble protein extracts were isolated from strains transformed with a plasmid carrying the β-galactosidase gene without a promoter (pTA919, columns "no"), with a synthetic promoter (psyn, columns "psyn") and with the *rpl37R* promoter (columns "l37"). Activity tests with the β-galactosidase substrate ONPG were performed and the specific activities of β-galactosidase for each protein fraction are shown (values are given in units/mg, y-axis). Blue columns and red columns are data from extracts isolated from cells grown to exponential and stationary phase, respectively. Activity tests with the β-galactosidase substrate ONPG showed that the protein extract from cells transformed with the *rpl37R* promoter construct had highest β-galactosidase activity.

promoter. The *rpl37R* promoter fusion contains not only the promoter but also the codons for the first five amino acids from the L37.eR protein resulting potentially in a more stable protein.

Taken together these data show that the L37.eR expression is driven by two independent promoters: on the one hand it is expressed from the *lsm* promoter resulting in the dicistronic mRNA and on the other hand it is expressed from the promoter located in the *lsm* frame, which is specific for the *rpl37R* gene only, and ensures strong expression in exponential phase.

3.3. Deletion of the Sm1 motif

In a previous study we deleted the complete reading frame of the *lsm* gene (*Haloferax* strain Δ*lsm*) leaving only the nucleotides that overlap with the downstream *rpl37R* gene (ATGA) in the chromosome [46]. In this deletion strain the *rpl37R* specific promoter was also deleted since it is part of the *lsm* reading frame [46]. The resulting deletion mutant showed a reduction in growth upon incubation at lower temperatures [46]. Complementation of the Δ*lsm* deletion strain with *lsm* and *rpl37R* can rescue the phenotype, while complementation with the *lsm* gene alone does not rescue the phenotype (data not shown). Taking together these data suggest that expression of the *rpl37R* gene is influenced in the Δ*lsm* deletion strain because the *rpl37R* gene promoter is also removed. Therefore we aimed to generate a new *lsm* deletion mutant, which leaves the *rpl37R* specific promoter intact and the *rpl37R* expression unchanged and thus allows the investigation of the Lsm protein function. Members of the Sm/Lsm protein family are characterized by two conserved regions: the Sm1 and Sm2 motifs, which are separated by a variable region [17]. Since the promoter for the downstream *rpl37R* gene is located in the coding region for the Sm2 motif (Fig. 1A), we selected the Sm1 motif for deletion (Fig. 1A) and generated the deletion strain ΔSm1 (Fig. 5). Northern analysis of the resulting ΔSm1 strain showed that expression of *rpl37R* is not changed and that the *rpl37R* mRNA is still transcribed separately, maybe even slightly stronger than in the wild type (Fig. 2, lanes ΔSm1). Comparison of growth showed that the ΔSm1 deletion strain did not grow differently from the wild type strain under several conditions, e.g. several different salt concentrations and carbon sources (data not shown).

3.4. Influence of the Sm1 deletion on sRNA expression

In an earlier work we could show that several sRNA co-purify with a FLAG-tagged Lsm protein [46]. We wanted to analyse whether the deletion of the Sm1 domain resulted in a change in expression of those sRNA, that co-purified with the Lsm protein. We isolated RNA from wild type and ΔSm1 cells and analysed these RNAs using northern blots. Probes against sRNAs 30, 34, 45 and 132 were used to investigate changes in expression in the deletion strain (data not shown). No changes were observed between wild type and ΔSm1, thus the deletion of the Sm1 domain does not influence the expression of the interacting sRNAs under the conditions tested. A similar observation has been made with a *hfq* deletion strain in *B. subtilis*, where the deletion of the *hfq* gene also did not influence expression of sRNAs [27]. However, since we only investigated the presence of the sRNAs under standard conditions we can not rule out that expression of these sRNA in the ΔSm1 strain is changed under specific conditions as for instance stress conditions.

3.5. General changes in the transcriptome of the ΔSm1 strain

To investigate the overall effect of the deletion of the Sm1 domain we compared the transcriptome of the ΔSm1 strain with

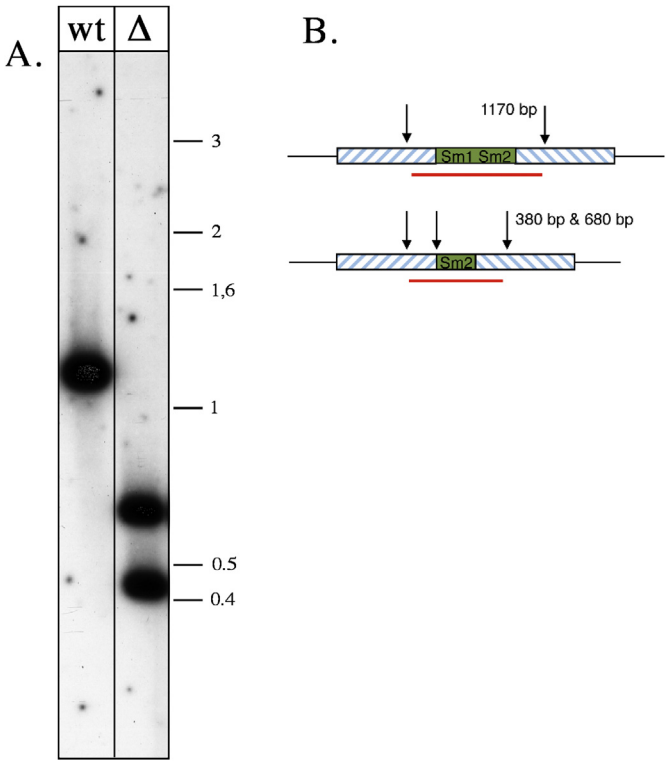

Fig. 5. Generation of the ΔSm1 strain. To keep the promoter for the downstream *rpl37R* gene intact, only the coding region for the Sm1 domain was removed. **A**. Southern blot showing wild type DNA (lane wt) and DNA form the ΔSm1 strain (lane Δ) after digestion with *Nco*I and *Bam*HI and hybridisation with a probe against part of the *lsm* gene (see B.). **B**. A probe against the *lsm* gene was used to detect wild type an deletion mutant DNA fragments. The red bar shows the location of the probe. The probe hybridises with a 1170 bp long fragment in the wild type and two fragments (380 and 680 bp) in the deletion mutant. The location of the recognition sites for the restriction endonucleases are indicated by arrows.

that of the wild type strain. cDNA was generated from RNA isolated from both strains and analysed by high throughput sequencing. Comparison of the obtained sequences from both strains showed that several genes are down-regulated in the deletion strain and many genes are up-regulated (Table 1, Supplementary Tables 1 and 2). Table 1A and B list all affected genes which show transcription level differences with a log2 value greater than 3 and lower than −3, respectively. Twelve genes are down-regulated, including the *lsm* gene (Table 1A). Since only 90 nucleotides of the 230 nucleotide long *lsm* mRNA were deleted, the transcript is still detected. Furthermore, a whole gene cluster shows lower expression: HVO_A0145–HVO_A0148, all four genes of this cluster are reduced by almost the same rate. The neighbouring HVO_A0149 gene, which is encoded on the opposite strand is likewise down-regulated. In addition, the large and small subunits of a carbamoyl-phosphate synthase, the bacterio-opsin activator-like protein as well as three hypothetical proteins are affected. Sixteen genes are up-regulated with a log2 value lower than −3 in the ΔSm1 strain (Table 1B): a putative isomerase, $tRNA^{Glu}$, a xylose dehydrogenase, two subunits of a glycerol-3-phosphate dehydrogenase, a putative glucose fructose oxidoreductase, the purine-binding chemotaxis protein CheW, aminotransferase class III, two proteins involved in ABC-type transport systems, a ygbK domain protein and five hypothetical proteins.

To confirm the changes in the expression observed with the transcriptome analysis, we investigated the expression of two down-regulated genes, HVO_A0147: ABC-type transport system permease protein and HVO_A0149: glucan 14-alpha-glucosidase,

Table 1
Genes that are up- and down-regulated in the ΔSm1 strain, respectively, are listed. The gene name, the encoded protein and the LOG_2 values are shown.

Gene name	Protein encoded	LOG_2
A. Genes down-regulated in ΔSm1		
HVO_A0147	ABC-type transport system permease protein (probable substrate sugar), tsgC4	5.0
HVO_A0148	ABC-type transport system periplasmic substrate-binding protein (probable substrate sugar), tsgD4	4.9
HVO_A0146	ABC-type transport system permease protein (probable substrate sugar), tsgB4	4.5
HVO_2723	Lsm protein (snRNP homolog)	3.8
HVO_2361	Carbamoyl-phosphate synthase, large subunit, carB (glutamine-hydrolysing)	3.7
HVO_A0145	ABC-type transport system ATP-binding protein (probable substrate sugar), tsgA4	3.5
HVO_1116	HTH domain protein	3.5
HVO_A0149	Glucan 14-alpha-glucosidase	3.3
HVO_1747	Hypothetical protein	3.2
HVO_0489	Hypothetical protein	3.2
HVO_2508	Carbamoyl-phosphate synthase, small subunit, carA (glutamine-hydrolysing)	3.1
HVO_1117	HTH-10 family transcription regulator, boa2	3.1
B. Genes up-regulated in ΔSm1		
HVO_B0210	Putative isomerase, YfiH family protein	−4.1
HVO_3058	$tRNA^{Glu}$	−3.9
HVO_B0028	Xylose dehydrogenase (NAD/NADP), xdh1	−3.7
HVO_1844	Hypothetical protein	−3.5
HVO_A0271	Glycerol-3-phosphate dehydrogenase subunit C, glpC2	−3.5
HVO_B0103	Glucose fructose oxidoreductase, putative	−3.4
HVO_A0607	Purine-binding chemotaxis protein CheW, cheW2	−3.4
HVO_B0070	Pyridoxal phosphate-dependent aminotransferase	−3.4
HVO_1843	Hypothetical protein	−3.4
HVO_A0089	Small CPxCG-related zinc finger protein	−3.3
HVO_2136	Hypothetical protein	−3.2
HVO_A0270	Glycerol-3-phosphate dehydrogenase subunit B, glpB2	−3.2
HVO_A0173	Hypothetical protein	−3.1
HVO_2444	ABC-type transport system ATP-binding protein (probable substrate dipeptides/oligopeptides), dppF5	−3.1
HVO_2112	DUF1537 family protein	−3.1
HVO_B0150	ABC-type transport system periplasmic substrate-binding protein (probable substrate iron-III)	−3.0
C. Motility related transcript changes in ΔSm1		
flaD, HVO_1203	fla cluster protein FlaD1	−1.91
cheB, HVO_1224	Protein-glutamate methylesterase CheB	−1.93
flaF, HVO_1214	fla cluster protein FlaF	−2.16
pilB1, HVO_620	Type IV pilus biogenesis complex ATPase subunit	−2.19
flgA1, HVO_1210	flagellin A1	−2.24
cheR, HVO_1222	Protein-glutamate O-methyltransferase CheR	−2.33
flgA2, HVO_1211	Flagellin A2	−2.40
cheD, HVO_1205	Taxis cluster protein CheD	−2.97
flaG, HVO_1215	fla cluster protein FlaG	−3.01
cheW2, HVO_A0607	Purine-binding chemotaxis protein CheW	−3.42

Genes belonging to the gene cluster for ABC-type transport system proteins are shaded in grey.

using a northern blot (Fig. 6). The highest down regulation according to the transcriptome data was observed for the mRNA of the ABC-type transport system permease protein (gene HVO_A0147) ($LOG_2 = 5.0$). The gene HVO_A0147 overlaps with the downstream HVO_A0146 gene that in turn abuts the HVO_A0145 gene, in addition HVO_A0147 is located close to the upstream HVO_A0148 gene. According to this genomic location all four genes might be transcribed into a single transcript. Transcriptome data also show down regulation of mRNAs for the neighbouring genes (HVO_A0145: LOG_2 3.5, HVO_A0146: LOG_2 4.5, HVO_A0148: LOG_2 4.9). The northern blot confirms that the genes are transcribed together and that the polycistronic RNA is down regulated in the ΔSm1 strain (Fig. 6). In addition the northern blot shows that the mRNA for the HVO_A0149 gene is monocistronic and that it is down-regulated in the ΔSm1 strain. Together the northern analyses confirm the transcriptome data.

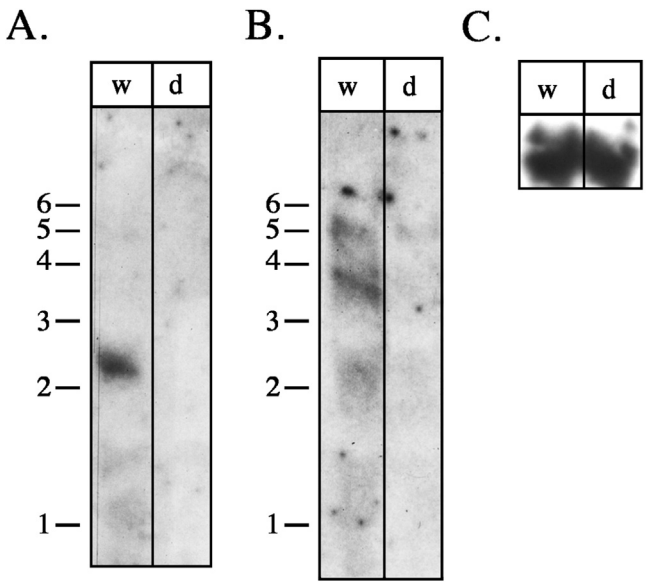

Fig. 6. Northern blot analyses of down-regulated transcripts. The two mRNAs for the glucan 14-alpha-glucosidase (HVO_A0149) (panel A) and ABC-type transport system permease protein (HVO_A0147) (panel B) are according to the transcriptome analyses down-regulated in the ΔSm1 strain. **A**. A northern blot carrying RNA from wild type (lane w) and ΔSm1 deletion strain (lane d) was hybridised with a probe against the mRNAs for glucan 14-alpha-glucosidase. An RNA of about 2.3 kb is detected in the RNA from the wild type strain but not in the RNA from the ΔSm1 deletion strain confirming the transcriptome data. The ORF of the glucan 14-alpha-glucosidase is 2009 nucleotides long, thus the mRNA is monocistronic. **B**. After stripping the blot was hybridised with a probe against the mRNA for ABC-type transport system permease protein. The RNAs of about 4–6 kb are detected in the RNA from a wild type strain but not in the RNA from the ΔSm1 deletion strain. If the four ABC-type transport genes (A0145–A0148) are transcribed together an mRNA of about 6 kb would be expected. The different RNAs hybridising with the probe might be the polycistronic transcript encoding all four proteins and processing product thereof. An RNA size marker is given at the left in kilo nucleotides. **C**. The northern blot was hybridised with a probe against the 16S rRNA, showing that in both lanes the same amount of RNA was loaded.

3.6. *Swarming behaviour of Haloferax wild type and ΔSm1 deletion mutant strains*

Since the chemotaxis protein CheW (Hvo_A0607) is up-regulated in ΔSm1 (Table 1B) we investigated whether other genes involved in motility are likewise up-regulated. Altogether we found 10 proteins connected with motility to be up-regulated (Table 1C, Supplementary Table 1). The different transcript levels of motility genes prompted us to compare the swarming properties of wild type and ΔSm1 strains (Fig. 7). The ΔSm1 was more active in swarming than the wild type, again confirming the transcriptome data. Deletion in the *lsm* gene might be a stress for the cell inducing a search for better environmental conditions by movement. Future experiments will have to show whether the Lsm protein is more directly involved in the regulation of motility or whether the deletion of the Sm1 domain affects a downstream regulator like for instance a small regulatory sRNA. Regulation of mobility genes by sRNAs might be possible since it has recently been shown that five sRNA gene deletion mutants in *Haloferax* showed reduced swarming, while one deletion mutant exhibited enhanced swarming compared to the parent strain [62].

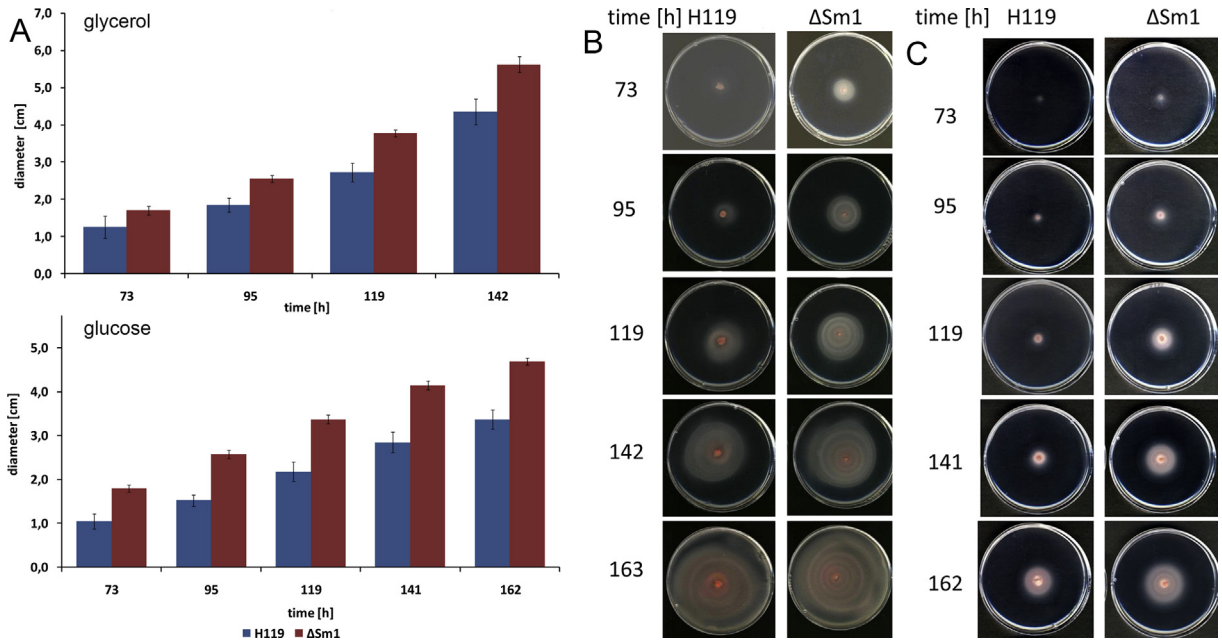

Fig. 7. Deletion strain ΔSm1 is more active in swarming. *Haloferax* wild type and deletion strain were compared with respect to their swarming activity. **A**. The swarming radii were measured for growth in medium with glycerol or glucose as carbon source. **B. & C**. Photos of swarming plates were taking at different time points after inoculation, swarming was analysed on plates with glycerol (B.) or glucose (C.) as carbon source.

3.7. Conclusion

Taken together we could show that the *lsm* gene and the overlapping *rpl37R* gene are co-transcribed but that expression of the *rpl37R* gene is in addition regulated independently by a second promoter located in the *lsm* ORF, which ensures strong expression in exponential growth phase. Deletion of the Sm1 motif results in major transcriptome changes, but sRNAs that were shown to interact with the Lsm protein are not affected in their expression.

Conflict of interest

No potential conflicts of interest were disclosed.

Acknowledgements

We are grateful to Elli Bruckbauer for her expert technical help. Work was funded by the German Research Council (Deutsche Forschungsgemeinschaft) through the priority program SPP 1258 “Sensory and regulatory RNAs in Prokaryotes” (grants DFG Ma1538/11-2, Be2121/5-2 and So264/14-2) and through DFG grants MA1538/18-1 and So264/21-1. We thank members of the priority programm for helpful discussions.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.biochi.2015.02.023.

References

[1] D.G. Scofield, M. Lynch, Evolutionary diversification of the Sm family of RNA-associated proteins, Mol. Biol. Evol. 25 (2008) 2255–2267, http://dx.doi.org/10.2210.1093/molbev/msn2175. Epub 2008 Aug 2257.
[2] B. Seraphin, Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs, EMBO J. 14 (1995) 2089–2098.
[3] J. Vogel, B.F. Luisi, Hfq and its constellation of RNA, Nat. Rev. Microbiol. 9 (2011) 578–589, http://dx.doi.org/10.510.1038/nrmicro2615.
[4] D. Wilson, R. Pethica, Y. Zhou, C. Talbot, C. Vogel, M. Madera, C. Chothia, J. Gough, SUPERFAMILY–sophisticated comparative genomics, data mining, visualization and phylogeny, Nucleic Acids Res. 37 (2009) D380–D386, http://dx.doi.org/10.310.1093/nar/gkn1762. Epub 2008 Nov 1026.
[5] P. Valentin-Hansen, M. Eriksen, C. Udesen, The bacterial Sm-like protein Hfq: a key player in RNA transactions, Mol. Microbiol. 51 (2004) 1525–1533.
[6] C.J. Wilusz, J. Wilusz, Eukaryotic Lsm proteins: lessons from bacteria, Nat. Struct. Mol. Biol. 12 (2005) 1031–1036.
[7] G.G. Carmichael, K. Weber, A. Niveleau, A.J. Wahba, The host factor required for RNA phage Qbeta RNA replication in vitro. Intracellular location, quantitation, and purification by polyadenylate-cellulose chromatography, J. Biol. Chem. 250 (1975) 3607–3612.
[8] M.T. Franze de Fernandez, L. Eoyang, J.T. August, Factor fraction required for the synthesis of bacteriophage Qbeta-RNA, Nature 219 (1968) 588–590.
[9] D. Schuppli, G. Miranda, H.C. Tsui, M.E. Winkler, J.M. Sogo, H. Weber, Altered 3'-terminal RNA structure in phage Qbeta adapted to host factor-less *Escherichia coli*, Proc. Natl. Acad. Sci. U S A 94 (1997) 10239–10242.
[10] J.S. Nielsen, A. Boggild, C.B. Andersen, G. Nielsen, A. Boysen, D.E. Brodersen, P. Valentin-Hansen, An Hfq-like protein in archaea: crystal structure and functional characterization of the Sm protein from *Methanococcus jannaschii*, RNA 13 (2007) 2213–2223. Epub 2007 Oct 2224.
[11] S. Gottesman, G. Storz, Bacterial small RNA regulators: versatile roles and rapidly evolving variations, Cold Spring Harb. Perspect. Biol. 3 (12) (2011) a003798, http://dx.doi.org/10.003710.001101/cshperspect.a003798.
[12] K.M. Wassarman, F. Repoila, C. Rosenow, G. Storz, S. Gottesman, Identification of novel small RNAs using comparative genomics and microarrays, Genes. Dev. 15 (2001) 1637–1651.
[13] E. Hajnsdorf, P. Regnier, Host factor Hfq of *Escherichia coli* stimulates elongation of poly(A) tails by poly(A) polymerase I, Proc. Natl. Acad. Sci. U S A 97 (2000) 1501–1505.
[14] B.K. Mohanty, V.F. Maples, S.R. Kushner, The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in *Escherichia coli*, Mol. Microbiol. 54 (2004) 905–920.
[15] M. Bouvier, C.M. Sharma, F. Mika, K.H. Nierhaus, J. Vogel, Small RNA binding to 5' mRNA coding region inhibits translational initiation, Mol. Cell 32 (2008) 827–837.
[16] B. Vecerek, M. Beich-Frandsen, A. Resch, U. Blasi, Translational activation of rpoS mRNA by the non-coding RNA DsrA and Hfq does not require ribosome binding, Nucleic Acids Res. 38 (2010) 1284–1293, http://dx.doi.org/10.1210.1093/nar/gkp1125. Epub 2009 Dec 1286.
[17] M.A. Schumacher, R.F. Pearson, T. Møller, P. Valentin-Hansen, R.G. Brennan, Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein, EMBO J. 21 (2002) 3546–3556.
[18] T. Møller, T. Franch, P. Hojrup, D.R. Keene, H.P. Bachinger, R.G. Brennan, P. Valentin-Hansen, Hfq: a bacterial Sm-like protein that mediates RNA–RNA interaction, Mol. Cell 9 (2002) 23–30.
[19] X. Sun, I. Zhulin, R.M. Wartell, Predicted structure and phyletic distribution of the RNA-binding protein Hfq, Nucleic Acids Res. 30 (2002) 3662–3671.
[20] A. Zhang, K.M. Wassarman, J. Ortega, A.C. Steven, G. Storz, The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs, Mol. Cell 9 (2002) 11–22.

[21] C. Sauter, J. Basquin, D. Suck, Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from *Escherichia coli*, Nucleic Acids Res. 31 (2003) 4091–4098.
[22] H.C. Tsui, H.C. Leung, M.E. Winkler, Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in *Escherichia coli* K-12, Mol. Microbiol. 13 (1994) 35–49.
[23] E. Masse, S. Gottesman, A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*, Proc. Natl. Acad. Sci. U S A 99 (2002) 4620–4625. Epub 2002 Mar 4626.
[24] T. Møller, T. Franch, C. Udesen, K. Gerdes, P. Valentin-Hansen, Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon, Genes. Dev. 16 (2002) 1696–1706.
[25] D.D. Sledjeski, C. Whitman, A. Zhang, Hfq is necessary for regulation by the untranslated RNA DsrA, J. Bacteriol. 183 (2001) 1997–2005.
[26] A. Zhang, S. Altuvia, A. Tiwari, L. Argaman, R. Hengge-Aronis, G. Storz, The OxyS regulatory RNA represses rpoS translation and binds the Hfq (HF-I) protein, EMBO J. 17 (1998) 6061–6068.
[27] H. Hammerle, F. Amman, B. Vecerek, J. Stulke, I. Hofacker, U. Blasi, Impact of Hfq on the Bacillus subtilis transcriptome, PLoS One 9 (2014) e98661, http://dx.doi.org/10.98610.91371/journal.pone.0098661 eCollection 0092014.
[28] J. Vogel, A rough guide to the non-coding RNA world of Salmonella, Mol. Microbiol. 71 (2009) 1–11, http://dx.doi.org/10.1111/j.1365-2958.2008.06505.x. Epub 02008 Nov 06504.
[29] Y. Chao, J. Vogel, The role of Hfq in bacterial pathogens, Curr. Opin. Microbiol. 13 (2010) 24–33, http://dx.doi.org/10.1016/j.mib.2010.1001.1001. Epub 2010 Jan 1014.
[30] N. De Lay, D.J. Schu, S. Gottesman, Bacterial small RNA-based negative regulation: hfq and its accomplices, J. Biol. Chem. 288 (2013) 7996–8003, http://dx.doi.org/10.7910.1074/jbc.R7112.441386. Epub 442013 Jan 441329.
[31] C. Mura, P.S. Randolph, J. Patterson, A.E. Cozen, Archaeal and eukaryotic homologs of Hfq: a structural and evolutionary perspective on Sm function, RNA Biol. 10 (2013) 636–651, http://dx.doi.org/10.610.4161/rna.24538. Epub 22013 Apr 24531.
[32] B.M. Collins, S.J. Harrop, G.D. Kornfeld, I.W. Dawes, P.M. Curmi, B.C. Mabbutt, Crystal structure of a heptameric Sm-like protein complex from archaea: implications for the structure and evolution of snRNPs, J. Mol. Biol. 309 (2001) 915–923.
[33] C. Mura, D. Cascio, M.R. Sawaya, D.S. Eisenberg, The crystal structure of a heptameric archaeal Sm protein: implications for the eukaryotic snRNP core, Proc. Natl. Acad. Sci. U S A 98 (2001) 5532–5537. Epub 2001 May 5531.
[34] I. Törö, S. Thore, C. Mayer, J. Basquin, B. Séraphin, D. Suck, RNA binding in an Sm core domain: X-ray structure and functional analysis of an archaeal Sm protein complex, EMBO J. 20 (2001) 2293–2303.
[35] A.D. Omer, T.M. Lowe, A.G. Russell, H. Ebhardt, S.R. Eddy, P.P. Dennis, Homologs of small nucleolar RNAs in Archaea, Science 288 (2000) 517–522.
[36] G. Puhler, H. Leffers, F. Gropp, P. Palm, H.P. Klenk, F. Lottspeich, R.A. Garrett, W. Zillig, Archaebacterial DNA-dependent RNA polymerases testify to the evolution of the eukaryotic nuclear genome, Proc. Natl. Acad. Sci. U S A 86 (1989) 4569–4573.
[37] P. Khusial, R. Plaag, G.W. Zieve, LSm proteins form heptameric rings that bind to RNA via repeating motifs, Trends Biochem. Sci. 30 (2005) 522–528.
[38] C. Mura, M. Phillips, A. Kozhukhovsky, D. Eisenberg, Structure and assembly of an augmented Sm-like archaeal protein 14-mer, Proc. Natl. Acad. Sci. U S A 100 (2003) 4539–4544. Epub 2003 Mar 4531.
[39] J. Salgado-Garrido, E. Bragado-Nilsson, S. Kandels-Lewis, B. Seraphin, Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin, EMBO J. 18 (1999) 3451–3462.
[40] T. Achsel, H. Stark, R. Lührmann, The Sm domain is an ancient RNA-binding motif with oligo(U) specificity, Proc. Natl. Acad. Sci. U S A 98 (2001) 3685–3689. Epub 2001 Mar 3620.
[41] S. Thore, C. Mayer, C. Sauter, S. Weeks, D. Suck, Crystal structures of the *Pyrococcus abyssi* Sm core and its complex with RNA. Common features of RNA binding in archaea and eukarya, J. Biol. Chem. 278 (2003) 1239–1247. Epub 2002 Oct 1229.
[42] I. Törö, J. Basquin, H. Teo-Dreher, D. Suck, Archaeal Sm proteins form heptameric and hexameric complexes: crystal structures of the Sm1 and Sm2 proteins from the hyperthermophile *Archaeoglobus fulgidus*, J. Mol. Biol. 320 (2002) 129–142.
[43] C. Mura, A. Kozhukhovsky, M. Gingery, M. Phillips, D. Eisenberg, The oligomerization and ligand-binding properties of Sm-like archaeal proteins (SmAPs), Protein Sci. 12 (2003) 832–847.
[44] J. Babski, L.K. Maier, R. Heyer, K. Jaschinski, D. Prasse, D. Jager, L. Randau, R.A. Schmitz, A. Marchfelder, J. Soppa, Small regulatory RNAs in Archaea, RNA Biol. 11 (2014) 484–493. Epub 2014 Mar 2031.
[45] J. Straub, M. Brenneis, A. Jellen-Ritter, R. Heyer, J. Soppa, A. Marchfelder, Small RNAs in haloarchaea: identification, differential expression and biological function, RNA Biol. 6 (2009) 281–292.
[46] S. Fischer, J. Benz, B. Späth, L.-K. Maier, J. Straub, M. Granzow, M. Raabe, H. Urlaub, J. Hoffmann, B. Brutschy, T. Allers, J. Soppa, A. Marchfelder, The archaeal Lsm protein binds to small RNAs, J. Biol. Chem. 285 (2010) 34429–34438.
[47] T. Allers, H.P. Ngo, M. Mevarech, R.G. Lloyd, Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the leuB and trpA genes, Appl. Environ. Microbiol. 70 (2004) 943–953.
[48] U. Johnsen, J.M. Sutter, A.C. Schulz, J.B. Tastensen, P. Schonheit, XacR - a novel transcriptional regulator of D-xylose and L-arabinose catabolism in the haloarchaeon *Haloferax volcanii*, Environ. Microbiol. 20 (2014) 1462–2920.
[49] G. Bitan-Banin, R. Ortenberg, M. Mevarech, Development of a gene knockout system for the halophilic archaeon *Haloferax volcanii* by use of the pyrE gene, J. Bacteriol. 185 (2003) 772–778.
[50] J. Sambrook, D. Russell, Molecular Cloning: a Laboratory Manual, Cold Spring Harbour Press, New York, 2001.
[51] D.T. Nieuwlandt, J.R. Palmer, D.T. Armbruster, Y.-P. Kuo, W. Oda, C.J. Daniels, A rapid procedure for the isolation of RNA from *Haloferax volcanii*, in: F.T. Robb, A.R. Place, K.R. Sowers, H.J. Schreier, S. DasSarma, E.M. Fleischmann (Eds.), Archaea: a Laboratory Manual, Cold Spring Harbour Press, New York, 1995, p. 161.
[52] S. Delmas, L. Shunburne, H.P. Ngo, T. Allers, Mre11-Rad50 promotes rapid repair of DNA damage in the polyploid archaeon *Haloferax volcanii* by restraining homologous recombination, PLoS Genet. 5 (2009) e1000552. Epub 1002009 Jul 1000510.
[53] M.L. Holmes, M.L. Dyall-Smith, Sequence and expression of a halobacterial beta-galactosidase gene, Mol. Microbiol. 36 (2000) 114–122.
[54] A. Large, C. Stamme, C. Lange, Z. Duan, T. Allers, J. Soppa, P.A. Lund, Characterization of a tightly controlled promoter of the halophilic archaeon *Haloferax volcanii* and its use in the analysis of the essential cct1 gene, Mol. Microbiol. 66 (2007) 1092–1106. Epub 2007 Oct 1031.
[55] M.L. Holmes, R.K. Scopes, R.L. Moritz, R.J. Simpson, C. Englert, F. Pfeifer, M.L. Dyall-Smith, Purification and analysis of an extremely halophilic beta-galactosidase from *Haloferax alicantei*, Biochim. Biophys. Acta 1337 (1997) 276–286.
[56] D. Gregor, F. Pfeifer, In vivo analyses of constitutive and regulated promoters in halophilic archaea, Microbiology 151 (2005) 25–33.
[57] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, Bioinformatics 25 (2009) 1754–1760, http://dx.doi.org/10.1710.1093/bioinformatics/btp1324. Epub 2009 May 1718.
[58] R. Hilker, K.B. Stadermann, D. Doppmeier, J. Kalinowski, J. Stoye, J. Straube, J. Winnebald, A. Goesmann, ReadXplorer–visualization and analysis of mapped sequences, Bioinformatics 30 (2014) 2247–2254, http://dx.doi.org/10.2210.1093/bioinformatics/btu2205. Epub 2014 Apr 2230.
[59] K. Jantzer, K. Zerulla, J. Soppa, Phenotyping in the archaea: optimization of growth parameters and analysis of mutants of *Haloferax volcanii*, FEMS Microbiol. Lett. 322 (2011) 123–130.
[60] T. Koide, D.J. Reiss, J.C. Bare, W.L. Pang, M.T. Facciotti, A.K. Schmid, M. Pan, B. Marzolf, P.T. Van, F.Y. Lo, A. Pratap, E.W. Deutsch, A. Peterson, D. Martin, N.S. Baliga, Prevalence of transcription promoters within archaeal operons and coding sequences, Mol. Syst. Biol. 5 (2009) 285, http://dx.doi.org/10.1038/msb.2009.1042. Epub 2009 Jun 1016.
[61] M. Brenneis, O. Hering, C. Lange, J. Soppa, Experimental characterization of Cis-acting elements important for translation and transcription in halophilic archaea, PLoS Genet. 3 (2007) e229.
[62] K. Jaschinski, J. Babski, M. Lehr, A. Burmester, J. Benz, R. Heyer, M. Dorr, A. Marchfelder, J. Soppa, Generation and phenotyping of a collection of sRNA gene deletion mutants of the haloarchaeon *Haloferax volcanii*, PLoS One 9 (2014) e90763, http://dx.doi.org/10.90710.91371/journal.pone.0090763 eCollection 0092014.