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 Maiera, Juliane Benza, Susan Fischer a, Martina Alstetter a, Katharina Jaschinskib, Rolf Hilker c, Anke Becker d, Thorsten Allers e, Jörg Soppab, Anita Marchfeldera,*

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.

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
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 Methanocaldococcus jannaschii [10,21]. Pliотropic 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 hfg 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, an important role being required for mRNA degradation, mRNA decapping, RNA stimulation, 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 be 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].

Archaebacteria can encode up to three different types of Lsm proteins (Lsm1–3) [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 Pyrococcus abyssi [38], a similar 14mer complex has also been observed for Lsm1 proteins [43]. In the halophilic archaean 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 L37eR [31]. Here, we show that the gene for the Haloferax Lsm protein contains a promoter for the downstream encoded L37eR 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 (ΔpyrE2, ΔtrpA, ΔleuB) [47], H555 [48] and ΔSm1 (ΔpyrE2, ΔtrpA, ΔleuB, 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 ΔSm1 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.d2 and Lsm37.u1 and Lsm37.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 KpnI and BamHI and subsequently cloned into the integrative pTA131 (digested with KpnI
and BamHI), yielding plasmid pTA131-Lsm37up. pTA131-Lsm37up was digested with NotI and BamHI to ligate fragment Lsm37down (digested with NotI and BamHI) 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 NotI and BamHI 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 Readprime™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 transcriptase and random primite kit Readprime™II (GE Healthcare). 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 RPMK 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-32P-dCTP and random prime kit Readprime™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 OD550 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 (bgaha). This chromosomal deletion prevents homologous recombination between the two similar β-galactosidase genes. H555 is a derivative of H25 [47], it was generated using the bgaha deletion plasmid pTA617. To construct pTA617, a HindIII-BsrgI fragment of pTA128 [52] containing the bgaha gene was cloned in pTA131 [47] at Asp718 and Noti sites (blunt-ended) to generate pTA506, then an EcoRV-BamHI fragment containing the bgaha 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 Ndel 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 Clal and Xbal 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 rpl37k gene was amplified from Haloferax genomic DNA with PCR using the following primers: ProL37#2 and ProL37#3. After digestion with Apal and Ndel the resulting DNA fragments were cloned into pTA919 (digested with Apal and Ndel) yielding the rpl37k 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 Apal and Ndel the resulting DNA fragments were cloned into pTA919 (digested with Apal and Ndel) yielding the rpl37k 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. To verify the results the expression profile of the β-galactosidase (HVO_A0149) mRNA using qRT-PCR was analysed. For qPCR, cDNA was reverse transcribed using the QuantiTect kit (Qiagen, Hilden, Germany). The expression of bgaha was normalised to the reference gene rpl37k (HVO_A0127), which was used as an internal control. The primer pair for bgaha is: bgahaF (5′-ACCTTCCTGGGCTTCAGTCA-3′, nucleotide 106–129) and bgahaR (5′-GCTCTGCTCTGGTTTGGTG-3′, nucleotide 679–698). The primer pair for rpl37k is: rpl37kF (5′-GTGGGGCTACCGGTGCTG-3′, nucleotide 184–206) and rpl37kR (5′-GTGTCAGTGAGCCCTGTCG-3′, nucleotide 1411–1433). The qPCR conditions were as follows: 15 s at 95°C for initial denaturation, followed by 45 cycles of 15 s at 95°C and 1 min at 60°C. The Melting curves were recorded from 60°C to 95°C at an increment of 0.5°C. The relative expression levels were calculated using the ΔΔCt method [55]. The data were analysed using the Prism software (Graphpad).
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 (Epigenetica). 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 RNA 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.

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.
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 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 “137”). 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 L37eR protein resulting potentially in a more stable protein.

Taken together these data show that the L37eR 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 Δlsm) 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 Δlsm 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 Δlsm 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 hfg deletion strain in B. subtilis, where the deletion of the hfg gene also did not influence expression of sRNAs [27]. However, since we only investigated the presence of the sRNAs under standard conditions we cannot 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 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, rRNA13L, 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,
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) (LOG2 = 5.0). The gene HVO_A0147 overlaps with the downstream gene HVO_A0146 gene that in turn abuts the HVO_A0145 gene in the 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: LOG2 3.5, HVO_A0146: LOG2 4.5, HVO_A0148: LOG2 4.9). The 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. After stripping the blot the probe might be 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 RNA encoding 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. The northern blot was hybridised with a probe against the 16S rRNA, showing that in both lanes the same amount of RNA was loaded.

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.

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].
interact with the Lsm protein are not affected in their expression. major transcriptome changes, but sRNAs that were shown to in exponential growth phase. Deletion of the Sm1 motif results in promoter located in the rpl37R gene is in addition regulated independently by a second overlapping rpl37R gene is co-transcribed but that expression of the gene are co-transcribed but that expression of 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 program for helpful discussions.

Appendix A. Supplementary data

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

References


[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 US snRNPs, EMBO J. 14 (1995) 2089–2098.


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 taken 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.


