Identification of a Small Regulatory RNA UspS Associated with the Universal Stress Protein in *Lactobacillus* Species

Zarah M. Fowler,[1,2] Sasha S. Bronovitskiy,[1,3] Finn K. Rose,[1,4] and Brian M. Lee[1*]

[1]Department of Chemistry, Coastal Carolina University, Conway, SC; [2]Department of Biology, Coastal Carolina University, Conway, SC; [3]School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA; [4]Horry County Schools Scholars Academy, Conway, SC

The gut microbiome is a complex habitat with many bacterial species, each playing crucial roles in regulating various physiological processes in the body. As the use of probiotics to combat human disease continues to increase, it is important to understand the mechanisms by which probiotic bacteria regulate their interactions with other bacteria and their host. Our exploration of the physiological functions of probiotic bacteria hopes to elucidate the role of small regulatory RNA (sRNA) in regulating gene expression within the microbiome. The goal of this project was to characterize the structure and function of the sRNA, UspS, which is found in probiotic, lactic acid bacteria. In *Lactobacillus*, UspS is closely associated with a downstream universal stress protein and contains an orphaned Lacto-*usp* RNA motif of unknown function. Computational methods have been used to study the UspS sRNA sequences from two *Lactobacillus* species in order to predict the secondary structures, generate 3D models, and search for potential mRNA interactions. Comparative sequence alignments and covariance analysis within the secondary structures predict a pseudoknot structure. The UspS sequence was isolated from two *Lactobacillus* species and sRNAs were synthesized by *in vitro* transcription with a T7 RNA polymerase. In preliminary studies, differential scanning fluorimetry of the UspS sRNA was able to confirm the presence of stable secondary structures. Future work will be focused on the structure of the pseudoknot region of UspS and its role in regulating the expression of the downstream universal stress protein.

Introduction

The completion of the human genome project in 2003 revealed that the human genome consists of about 20,000 genes.[1–3] This project marked a historical scientific achievement in studying the human body and advanced the field of medicine by accelerating our understanding of human diseases. Some of the major advancements resulting from the sequencing and identification of gene locations have been the improved comprehension of diseases such as breast cancer, cystic fibrosis, age-related macular degeneration, Alzheimer's disease, and Huntington's disease.[2,4,5] Two decades later, scientists continue to dedicate extensive efforts to studying these genes and their impact on human health, but in doing so we have overshadowed nearly eight million other genes present in the human body (Fig. 1).[6] These overlooked genes originate from the gut microbiome, which refers to the presence of small microbial species residing on and within the body.

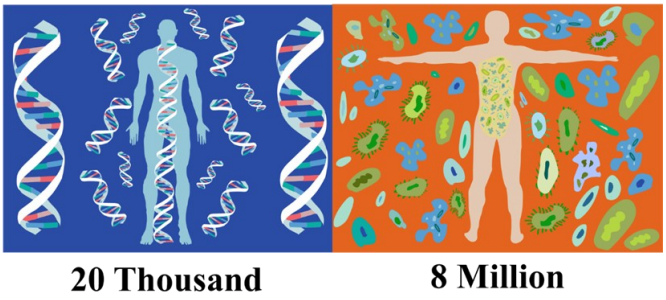

Fig. 1. The number of human genes in the body (~22,000), and the number of microbial genes in the human body (~8,000,000). Graphical representation prepared in Adobe Illustrator by Zarah Fowler.[55]

The composition of the microbiome is influenced by an array of factors including diet, environment, antibiotic usage, familial genetics, and age.[7] Among the many types of bacteria inhabiting the microbiome are probiotic bacteria. These beneficial microorganisms are frequently encountered in everyday food staples such as yogurt, pickles, bread, milk, and cheese (Fig. 2). Their ubiquity in our diet adds an intriguing dimension to the intricate workings of our bodies.

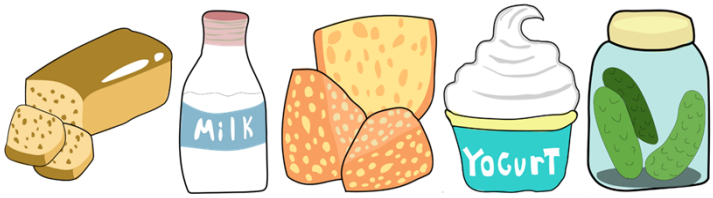

Fig. 2. Examples of probiotic foods: bread, milk, cheese, yogurt, and pickles. Graphical representation prepared in Adobe Illustrator by Zarah Fowler.[55]

The production of these foods involves a fermentation process that utilizes yeast or lactic acid bacteria. These tiny microorganisms have been recognized for their ability to benefit the host through various mechanisms, such as reducing intestinal pH, safeguarding against pathogenic organisms by means of competition, and regulating the immune response.[8,9] Furthermore, there is the proposition that probiotics can serve as preventative and therapeutic agents in combating human diseases, such as acute diarrhea, *Clostridium difficile* infection, and Crohn's disease.[8] One possible mechanism by which probiotic bacteria influence human molecular physiology is through small non-coding RNAs or sRNAs.

The central dogma of molecular biology describes the process by which DNA is transcribed into RNA, and, in turn, RNA is translated into proteins.[10] These proteins play vital roles in the molecular mechanisms of both human bodies and bacteria. However, recent research has unveiled a fascinating discovery: not all RNA molecules undergo translation into proteins; instead, they persist as non-coding RNA, typically ranging from 50 to 400 nucleotides in length.[11,12] These sRNAs assume the role of regulators within bacteria, exerting post-transcriptional control over gene expression and post-translational control over protein function. By interacting with mRNA through base-pairing, sRNA can modulate gene expression by influencing the termination of transcription, the accessibility for translation, or the long-term stability of the mRNA (Fig. 3).[13]

Extensive research spanning several decades had primarily focused on studying sRNAs in Gram-negative bacteria, such as *Escherichia coli* and *Salmonella enterica*.[14] In Gram-negative pathogenic bacteria, sRNAs exert significant influence on cell physiology, virulence, and pathogenesis systems.[13] Interestingly, recent investigations suggest that bacteria-derived sRNAs can impact the host by being exported within extracellular vesicles.[12] In the case of pathogenic bacteria, extracellular vesicles, transmembrane proteins, and gap junctions serve as conduits for delivering sRNAs to hosts, thereby suppressing host immunity. Consequently, hosts may respond by expressing their own sRNAs to counteract the virulence.[15] Considering that Gram-negative bacteria and

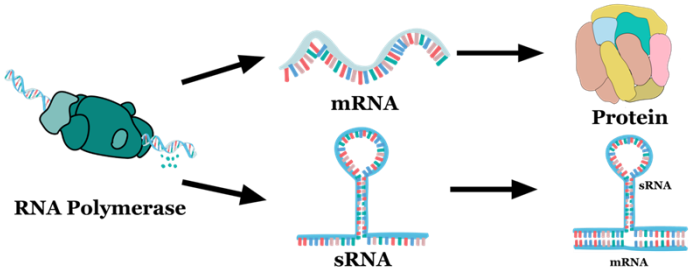

Fig. 3. The central dogma of molecular biology and transcription of sRNA. The central dogma depicts transcription of DNA into mRNA by RNA polymerase and translation of mRNA into proteins. Alternatively, DNA can be transcribed into sRNA. This sRNA then interacts with mRNA to control gene expression or proteins to control function. Graphical representation prepared in Adobe Illustrator by Zarah Fowler.[55]

host employ mechanisms for cross-kingdom exchange of sRNAs, it is reasonable to assume that Gram-positive bacteria also possess the potential to interact with the host through similar mechanisms.

Given the growing interest in probiotics and their widespread usage, it becomes crucial to delve into the realm of sRNAs within lactic acid bacteria. There exists limited research focused on sRNAs in Gram-positive bacteria, leaving it unclear whether sRNAs in Gram-positive bacteria are homologous to those in Gram-negative bacteria. However, recent studies in Gram-positive bacteria have shown a similar reliance on mRNA-sRNA interactions, but limited use of RNA-binding proteins to chaperone these interactions.[14] These findings highlight the significance of investigating sRNAs across a diverse range of bacterial species.

Considering the profound impact of the human gut microbiome on human health, it becomes imperative to unravel the intricate mechanisms employed by various bacteria in controlling gene expression and protein function. For the scope of this study, we have included two prominent probiotic Gram-positive bacteria frequently found in the gut and in consumer products. In this research project, we selected an sRNA found in *Lactobacillus bulgaricus* and *Lactobacillus acidophilus*. This sRNA was named UspS due to the presence of a downstream *usp* gene that codes for the universal stress protein, UspA. The purpose of this project was to identify and explore the conservation of function and structure of the UspS sRNA within *L. bulgaricus* and *L. acidophilus*.

The identification of sRNA roles often necessitates expensive and time-consuming experiments. As a result, computational methods have been developed to identify and characterize sRNAs before experimental techniques are undertaken.[16] In this study, a variety of bioinformatic techniques were employed to explore the function and structure of UspS. First, NCBI blast analysis was successful in identifying homologues of UspS in sixteen *Lactobacillus* species.[17] These homologous sequences were aligned within Jalview to examine nucleotide conservation and identify the promoter elements.[18] Subsequently, the secondary structure of UspS in *L. bulgaricus* and *L. acidophilus* was predicted via mfold and RNApdbee, while the tertiary structures were generated using FARFAR2 algorithm in Rosetta.[19–22]

Analysis of the secondary structure of UspS revealed structural similarity in the P4 region, which is predicted to form a pseudoknot structure (Fig. 4). A sequence alignment of UspS exhibited 56% conservation in the P4 region, which suggests that the P4 region may be evolutionarily conserved, serving as an essential site for mRNA interactions. Previous studies have noted the similarity between the P4 region of UspS and 6S RNA, which binds to the sigma subunit of RNA polymerase and inhibits transcript under nutrient stress conditions.[23–25] Alternate regulatory functions of UspS have been explored using IntaRNA and CopraRNA to predict potential sRNA-mRNA interactions.[26–30] The results indicated that UspS may initiate changes in bacterial membrane structure, nucleotide metabolism, and amino acid metabolism that is associated with a response to environmental stress. In preliminary structural studies, the UspS sequence was isolated from both *L. bulgaricus* and *L. acidophilus* using PCR to create DNA templates for transcription. The UspS sRNA has been synthesized by *in vitro* transcription using T7 RNA polymerase. Initial characterization by differential scanning fluorimetry shows evidence of stable secondary structures within the UspS sRNA transcripts, which can now be used for structural studies and interaction assays.

Methods

Selection of the UspS sRNA Sequence

The UspS sequence was selected as part of a broad search for regulatory RNA elements in lactic acid bacteria that have not been fully characterized in terms of structure and function. The UspS sequence includes the Lacto-*usp* motif, which was first reported by Weinberg and others, based on a computational screen that used covariation analysis to identify bacterial RNA motifs.[31] The motif was first identified in six sequences from the order Lactobacillales and appears within the 5'-untranslated region (UTR) of the *usp* gene, which codes for the universal stress protein.[31] Based on some structural similarity, the Lacto-*usp* motif was proposed to play the role of 6S RNA, which inhibits the activity of RNA polymerase during stationary growth to ensure cell survival.[31] However, the Lacto-*usp* motif is smaller than other described 6S RNA motifs and may instead be a *cis*-regulatory element for the expression of the universal stress protein.[31]

The UspS sequence was characterized using computational methods to delineate the boundaries of the regulatory RNA elements. The nucleotide BLAST algorithm from NCBI was used to find homologous UspS sequences that contain the Lacto-*usp* motif within intergenic regions of lactic acid bacteria.[17] Representative UspS sequences were identified in sixteen different species of *Lactobacillus*: *L. acidophilus, L. delbrueckii* subsp. *bulgaricus* (a.k.a. *L. bulgaricus*), *L. crispatus, L. amylovorus, L. ultunensis, L. helveticus, L. acetotolerans, L. jensenii, L. delbrueckii* subsp. *indicus, L. delbrueckii* subsp. *lactis, L. delbrueckii* subsp. *jakobensii, L. johnsonii, L. gasseri, L. paragasseri, L. taiwanensis,* and *L. kefiranofaciens.* The representative UspS sequences from *Lactobacillus* were aligned within Jalview using the Clustal algorithm.[32,33] The approximate transcription start site for each UspS sequence was identified by mapping the promoter sequences. The consensus promoter sequences of *Lactobacillus* include two hexameric sequences at -35 (TTGACA) and -10 (TATAAT) nucleotides upstream of the transcription start site and allowed for mapping of the start site within 5-6 nucleotides without experimental data.[34] For each of the UspS sequences from *Lactobacillus*, a rho-independent intrinsic terminator was identified from a preliminary prediction of the RNA secondary structure using mfold.[35]

UspS sRNA Secondary Structure Prediction and 3D Structure Modeling

UspS sequences from *L. bulgaricus* and *L. acidophilus* were selected for secondary structure predictions and modeling of the 3D structure. The web-based server for mfold was used to predict secondary structures of the UspS sequences with 5' end set at the transcription start site and the 3' end set immediately after the intrinsic terminator.[35] The core secondary structures were also predicted for each UspS sequence with the intrinsic terminator hairpin loop excluded. The lowest energy structures from mfold were used to generate a dot-bracket format file to describe the core secondary structure prediction results. The R-scape optimized structure of the Lacto-*usp* motif in Rfam was used to identify two pseudoknot contacts within the UspS sequences.[36,37] For each UspS sequence, both pseudoknot (PK) contacts were included in the dot-bracket format file with square brackets to denote the PK1 base pairing and curly brackets to denote the PK2 base pairing. In some cases, the predicted base pairing from mfold was adjusted to match the predictions from R-scape in order to include the pseudoknot contacts. The secondary structure figures were created with RNApdbee to include lines indicating the predicted pseudoknot contacts.[38,39] The 3D structures of the core UspS sequences from both *L. bulgaricus* and *L. acidophilus* were modeled using the FARFAR2 algorithm within the Rosetta molecular modeling suite.[40] The input for FARFAR2 included the FASTA format sequence files and modified dot-bracket format files to define the base-pairing and pseudoknot contacts. The secstruct_general flag was used to

Fig. 4. A sequence alignment of the P4 region in UspS orthologous sequences from sixteen different bacterial species made using Jalview multiple sequence alignment.[32] The alignment of *L. bulgaricus* and *L. acidophilus* showed 56% sequence conservation.

obligate base pairing as specified within the dot-bracket format file. For each 3D model, 10,000 structures were calculated and the lowest energy structure was extracted as the best model for each. The 3D models were visualized in PyMOL (Schrödinger) using the RiboVis representation of the RNA backbone and nucleotide structure (Das lab, Stanford).[41]

Identification of Potential UspS sRNA:mRNA Interactions

The UspS sequence appears to be contained within the 5'-UTR of the *usp* mRNA and may play a *cis*-regulatory role. However, in most of the UspS sequences, the intrinsic terminator resides downstream of the Lacto-*usp* motif, which would allow this UspS transcript to act as a regulator sRNA. In order to identify potential regulatory targets, CopraRNA and IntaRNA were used to search the genomes of both *L. bulgaricus* and *L. acidophilus* for genes with mRNA sequences that are complementary to the respective UspS sRNA transcripts from each.[26–30] The predicted interactions were screened using the CopraRNA web server to determine the potential regulatory regions within the UspS sRNA and to identify genes that may be targeted for regulation.

Growth of Lactic Acid Bacteria

Lactobacillus delbrueckii subsp. *bulgaricus* strain Lb14 (ATCC 11842) and *Lactobacillus acidophilus* strain Scav (ATCC 4356) were purchased as freeze-dried stocks from American Type Culture Collection (Manassas, VA). Both bacterial strains were initially propagated in Lactobacilli MRS media (Difco). An aliquot of the freeze-dried bacterial cells was removed with a pipet tip and transferred to 0.5 mL of MRS broth for resuspension. For each bacterial strain, a streak plate was prepared from the resuspended cells with an inoculation loop using MRS agar. The remaining resuspended cells were then added to 5 mL of MRS broth in 14 mL polypropylene culture tube with a dual position snap-cap (Falcon). All plates and broth cultures were incubated overnight at 37 °C under anaerobic conditions using an 0.4 L anaerobic jar, a gas generating sachet, and an oxygen indicator (Mitsubishi). The growth of bacterial cultures was monitored by measuring culture turbidity at 600 nm using a UV-Vis spectrophotometer (Thermo Scientific). The growth cultures were harvested during the exponential growth phase. Overnight growth cultures were used for genomic DNA extraction, detailed below. For glycerol stocks, 0.5 mL of growth culture was combined with 0.5 mL of 50% glycerol in polypropylene screw-top tube and stored at -74 °C.

Extraction of Genomic DNA from Lactic Acid Bacteria

Genomic DNA was extracted from both *L. bulgaricus* and *L. acidophilus* using the Wizard Genomic DNA Purification Kit (Promega). For each strain, a 1 mL aliquot of exponentially growing culture was harvested by centrifuging for 2 minutes at 16,000 × *g*. The protocol for isolating genomic DNA from Gram Positive bacteria was followed. After removing the supernatant, each cell pellet was resuspended in 480 µL of 50 mM EDTA pH 8.0 and 120 µL of 10 mg/mL of lysozyme (Sigma) was added to each microfuge tube. The lysozyme reactions were incubated at 37 °C for 45-50 minutes, then centrifuged for 2 minutes at 16,000 × *g*. The supernatant was removed and the cells were resuspended in 600 µL of Nuclei Lysis solution (Promega). The solution was incubated at 80 °C for 5 minutes and cooled to room temperature. After cooling, 3 µL of RNase A solution (Promega) was added and the sample was incubated at 37 °C for 15 minutes, then cooled to room temperature. After cooling, 200 µL of the Protein Precipitation solution (Promega) was added, and the solution was vortexed followed by incubation in ice for 5 minutes. The sample was centrifuged at 16,000 × *g* for 3 minutes. The supernatant was mixed with room temperature isopropanol and centrifuged at 16,000 × *g* for 2 minutes. After removing the supernatant, the DNA pellet was washed by with 600 µL of room temperature 70% ethanol and then centrifuged. Pellets were left to air dry for 12-20 minutes and then resuspended in 100-200 µL of DNA Rehydration solution (Promega). After incubation at room temperature overnight, the genomic DNA samples were stored at 4 °C.

Isolation of the UspS sRNA Genes by PCR

The UspS sRNA genes were isolated and amplified from the genomic DNA of both *L. bulgaricus* (Ldb) and *L. acidophilus* (Lac) using polymerase chain reaction (PCR) to prepare DNA templates for *in vitro* transcription. Two template constructs were designed for the UspS sRNA genes from each *Lactobacillus* species. The P16 constructs include all predicted paired regions of UspS from P1 through P6. The P15 constructs exclude the P6 stem-loop, which is predicted to form an intrinsic rho-independent terminator. The DNA primers (IDT) were designed using Primer3 to optimize specificity and minimize PCR byproducts.[42] The forward and reverse primers include the restriction sites, BamHI and HindIII, respectively, to allow for cloning into a pUC18 vector. The forward primer includes a 17-nucleotide promoter element for T7 RNA polymerase. The reverse primer includes an additional restriction site, either DraI or SwaI, to allow for blunt end cutting of the DNA template prior to *in vitro* transcription. The PCR reactions were prepared in 0.2 mL thin-walled tubes (Neptune) with 20 mL total volume with 1 mL of genomic DNA template, 0.5 mM of each primer, 0.2 mM of each dNTP (NEB), 0.4 units of Phusion High-Fidelity DNA polymerase (NEB), and the supplied HF buffer (NEB). The reaction protocol started with 30 seconds at 95 °C prior to 30 cycles of melting for 10 seconds at 95 °C, annealing for 30 seconds at the optimal temperature for each primer pair (from 59.9 °C to 63.9 °C), and polymerization for 45 seconds at 72 °C. The amplification cycles were followed by 5 minutes of extension time at 72 °C. The optimal annealing temperatures were maintained using the gradient heating block of the T100 Thermocycler (Bio-Rad). After completion, all reactions were stored at 4 °C.

Purification of UspS PCR Products

Purification of the four PCR products (Ldb P16, Ldb P15, Lac P16, and Lac P15) was done using the Wizard SV Gel and PCR Clean-Up System (Promega). For each 20 µL PCR reaction, an equal volume of Membrane Binding solution was added. Each sample was transferred to the silica membrane of an SV mini-column inserted into a collection tube. The protocol for the Wizard SV Gel and PCR Clean-Up System (Promega) was followed. After loading the sample onto the column, the silica membrane was washed twice with 700 µL and 500 µL of Membrane Wash solution and centrifuged at 16,000 × *g* for 1 minute and 4 minutes, respectively. After the last wash, the column assembly was centrifuged at 16,000 × *g* for 2 minutes with the lids open to allow any remaining ethanol to evaporate. Each DNA sample was incubated with 35 µL of ultrapure water (Millipore) at room temperature for 10 minutes. Samples were eluted from the column by centrifugation at 16,000 × *g* for 1 minute. The purified PCR products were stored at -20 °C. The four PCR products were assayed by gel electrophoresis using a 1.5% agarose gel prepared in TAE buffer with 0.25 mg/mL of ethidium bromide (EtBr) for staining. For each sample, 1 µL of PCR product was dilute to

6 µL with water and Purple gel loading dye (NEB). The gel electrophoresis was run at 80 V for 45 minutes after loading all four samples and a 100 bp DNA ladder (NEB) for a size reference. The gel was placed on a UV transilluminator to visualize the EtBr stained DNA bands and capture images using a Molecular Imager ChemiDoc XRS+ system (Bio-Rad).

Preparation of UspS DNA Templates for Transcription

DNA templates for transcription were prepared from the purified PCR products by using either DraI or SwaI restriction enzymes (NEB) to create blunt ends defining the 3' end of the transcript. For the Ldb P1-6 and Lac P1-6 constructs, 15 µL of each PCR product (about 1.5 µg) was digested with 20 units of DraI with the supplied rCutSmart buffer and the total volume adjusted to 30 µL with ultrapure water (Millipore). The DraI reactions were incubated overnight at 37 °C. For the Ldb P1-5 and Lac P1-5 constructs, 15 µL of each PCR product (about 1.5 µg) was digested with 10 units of SwaI with the supplied NEBuffer r3.1 and the total volume adjusted to 30 µL with ultrapure water (Millipore). The SwaI reactions were incubated overnight at room temperature, 23 °C. The digest reactions were purified using the Wizard SV Gel and PCR Clean-Up System (Promega) following the protocol described above for the four PCR products. The volume of the membrane binding solution was increased to match the 30 µL volume of each digest reaction. All other steps in the protocol were unchanged. The DNA was eluted with 35 µL of ultrapure water (Millipore).

Synthesis of UspS sRNA Transcripts

The four UspS sRNA constructs (Ldb P16, Ldb P15, Lac P16, and Lac P15) were synthesized by *in vitro* transcription using T7 RNA polymerase (NEB) and the supplied T7 reaction buffer (40 mM TrisHCl pH 7.9, 6 mM $MgCl_2$, 2 mM spermidine, and 1 mM DTT). For each construct, two 10 µL reactions were setup with variations in the final concentration of $MgCl_2$ (24 mM and 36 mM). The total concentration of NTPs for all reactions was 25 mM. The proportion of UTP, GTP, CTP, and ATP added to each reaction was adjusted to match base composition of the synthesized RNA. For each reaction, 0.25 µg of the blunt end digest DNA template and 1 unit of RNasin, ribonuclease inhibitor (Promega). The transcription reactions were incubated at 37°C for one hour. All transcription reactions were treated with 4 units of RNase-free TURBO DNase I with the supplied buffer (Invitrogen) adjusted with autoclaved ultrapure water (Millipore) to a total volume of 100 µL. The DNase digestion was incubated for 1 hour at 37°C.

Gel Electrophoresis of UspS sRNA Transcripts

Both UspS sRNA transcripts and DNA templates were assayed together by native polyacrylamide gel electrophoresis (PAGE) using a 4-20% Tris-glycine gel with the supplied buffer (Invitrogen). Each of the RNA samples included 5 µL of the DNase treated transcription reaction with RNA gel loading dye (Thermo Scientific). The DNA samples included 1 µL of the purified, digested DNA template with Purple gel loading dye (NEB). Both the RiboRuler low range RNA ladder (Thermo Scientific) and 100 bp DNA ladder (NEB) were included in the gel for size referencing. The gel electrophoresis was run at 200 V for 45 minutes. The gel was stained with SYBR Green II (Molecular Probes), which was diluted 1:10,000 with the native Tris-glycine gel running buffer (Invitrogen). After staining for 1 hour with agitation on a rocking platform, the gel was destained for another hour with ultrapure water (Millipore). The gel was visualized using a UV transilluminator with images captured by a Molecular Imager ChemiDoc XRS+ system (Bio-Rad).

Purification of the UspS sRNA Transcripts

The UspS sRNA transcripts of the four constructs (Ldb P16, Ldb P15, Lac P16, and Lac P15) were purified using the RNeasy MinElute Cleanup Kit (QIAGEN). For each construct, the DNase I treated transcription reactions at two different concentrations of $MgCl_2$ (24 mM and 36 mM) were combined and the volume of each combined construct sample was adjusted to 200 µL with ultrapure water (Millipore). For each construct sample, 700 µL of guanidine-thiocyanate lysis buffer RLT (QIAGEN) and 500 µL of ethanol were added with mixing after each addition. The protocol for the RNeasy MinElute Cleanup Kit was followed. Each RNA sample was eluted from the silica membrane with 14 µL of RNase-free water (QIAGEN). The purified RNA for each construct was stored at –20°C.

Structural Characterization of the UspS Transcripts

The tertiary structures of the UspS sRNA transcripts for all four constructs (Ldb P16, Ldb P15, Lac P16, and Lac P15) were measured by differential scanning fluorimetry (DSF), also known as thermal melt assays.[43] For each construct, two different $MgCl_2$ concentrations were used: 0 mM and 0.5 mM. Each sample included 1 mg of RNA in 50 mL of total volume. The sample buffer was 20 mM TrisHCl at pH 7.5 with 10 mM NaCl. A fluorescent reporter dye, SYBR Green II RNA (Molecular Probes), was added to each sample at 1X concentration with max excitation at 497 nm and max fluorescence at 520 nm. The samples were prepared on ice in a clear 96-well clear polypropylene PCR plate (Bio-Rad) and then sealed with optically clear film (Bio-Rad). The plate was centrifuged briefly to remove bubbles and collect sample volume at the bottom of the wells. The plate was transferred to a CFX96 Touch RT-PCR (Bio-Rad) to measure the fluorescence intensities with the plate initially incubated at 5°C. The temperature was increased in steps of 0.5° C from 5°C to 95°C. At each 0.5°C increment, the temperature was held for 30 seconds before measuring the fluorescence intensity of every sample in the plate using the SYBR Green settings for excitation (450-490 nm) and detection (515-530 nm). The raw data was processed with the CFX Maestro software (Bio-Rad) using the melt curve analysis features to plot the first derivative of the fluorescence versus temperature. The temperature with the maximum change in fluorescence was used to determine the melting temperature for each sample.

Results and Discussion

Conservation of Sequence and Secondary Structure

The investigation into the sRNA UspS in *Lactobacillus* species has yielded significant findings regarding its conservation and potential molecular functions. Through multiple sequence alignment, a high level of conservation is observed among the sixteen *Lactobacillus* species examined. Sequence alignment of the P4 region of *L. bulgaricus* and *L. acidophilus* show 56% sequence conservation (Fig. 4). Structural conservation of UspS in the P4 region between *L. bulgaricus* and *L. acidophilus* is confirmed through secondary structure predictions using mfold, which highlights two matching internal loops, three stem areas, and one hairpin loop (Fig. 5).[35] The structural conservation of UspS in the P4 region is visually shown highlighted in pink. This conservation strongly indicates that UspS sRNA is essential for some molecular function of *Lactobacillus* and may be a significant site for mRNA interactions.

To further explore the structural characteristics of the P4 region, a secondary structure schematic was created using the conserved regions identified through multiple sequence alignment and secondary structure predictions (Fig. 6). In this figure, nucleotides that are >93% conserved between all sixteen *Lactobacillus* species are highlighted with a blue bubble. The nucleotides that are identical between *L. bulgaricus* and *L. acidophilus* are highlighted with a yellow circle. Additionally, pseudoknot sequences are described using a blue bracket notation. The high degree of conservation within the predicted PK1 contacts suggests that this interaction is important for tertiary structure within the P4 region. The variability within the PK2 contacts may play a functional role allowing for other factors to bind to the internal loop or terminal hairpin of P4.

Modeling the Tertiary Structure of UspS

The tertiary structures of the P4 region of UspS were modeled using the FARFAR2 algorithm within the Rosetta suite of molecular modeling programs (Fig. 7).[40] The predicted secondary structure and pseudoknot contacts were used as constraints to build the model structures. The models are colored to highlight pseudoknot contacts with purple (PK1) and pink (PK2) nucleotides. These 3D models, along with mapping regions of sequence conservation, enabled the identification of potential sites for regulation of protein expression. The presence of pseudoknots gives P4 a unique 3D structure as pseudoknots are caused by nucleotides

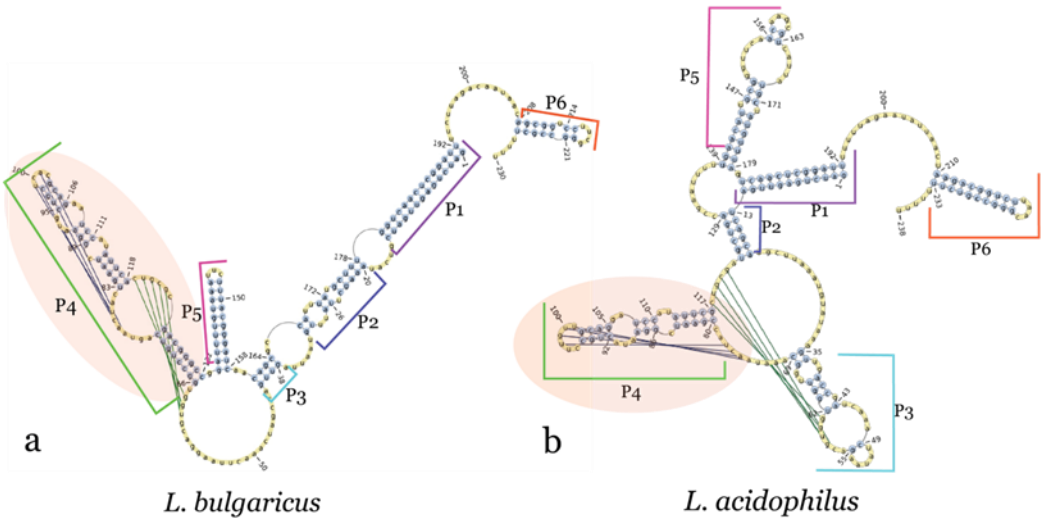

Fig. 5. Secondary structure comparison of UspS sRNA sequences. Structural regions are labeled P1 through P6. The P4 regions highlighted in pink are highly conserved between species. Pseudoknot sequences are identified by attaching lines on the P4 region. Secondary structures showed conservation of the P4 region in a) *L. bulgaricus* and b) *L. acidophilus*. Figures made using mfold and RNApdbee.[35,38,39,56,57]

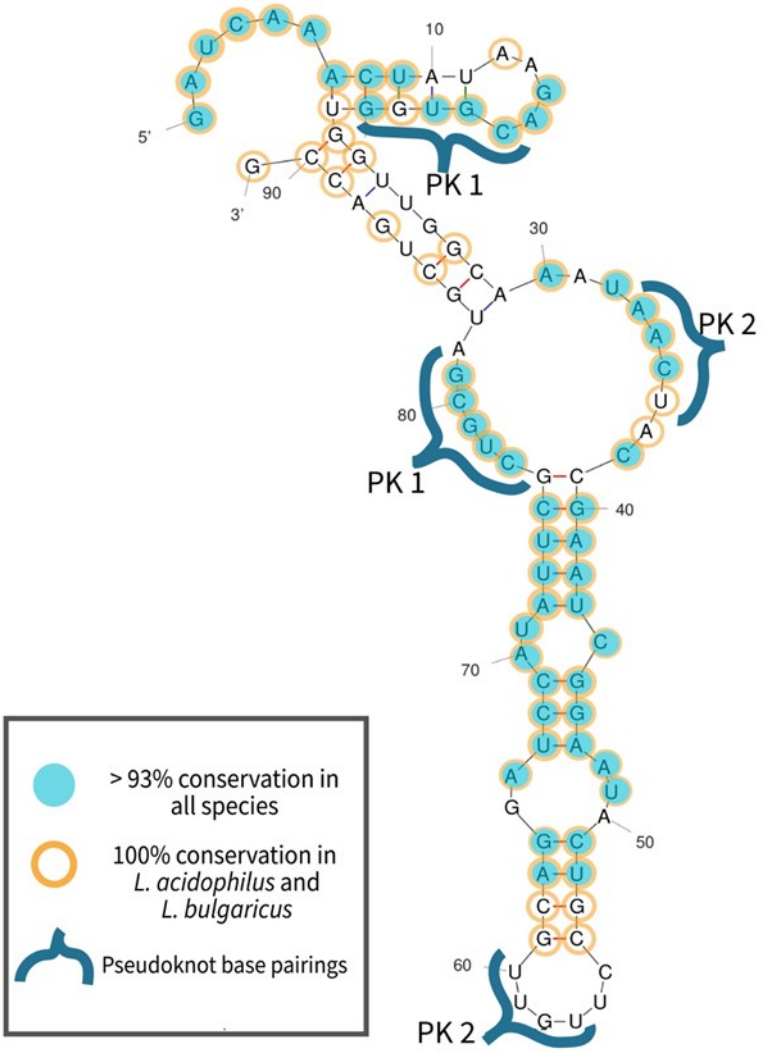

Fig. 6. Conservation of the UspS P4 region in sixteen species from the pairwise alignment using Jalview. The high degree of conservation in the predicted PK1 region suggests this interaction is important for tertiary structure in the P4 region, and variability within PK2 suggests it may play a functional role in allowing other factors to bind. Graphical representation prepared in Adobe Illustrator by Zarah Fowler based on secondary structure predictions with mfold.[32,35,55]

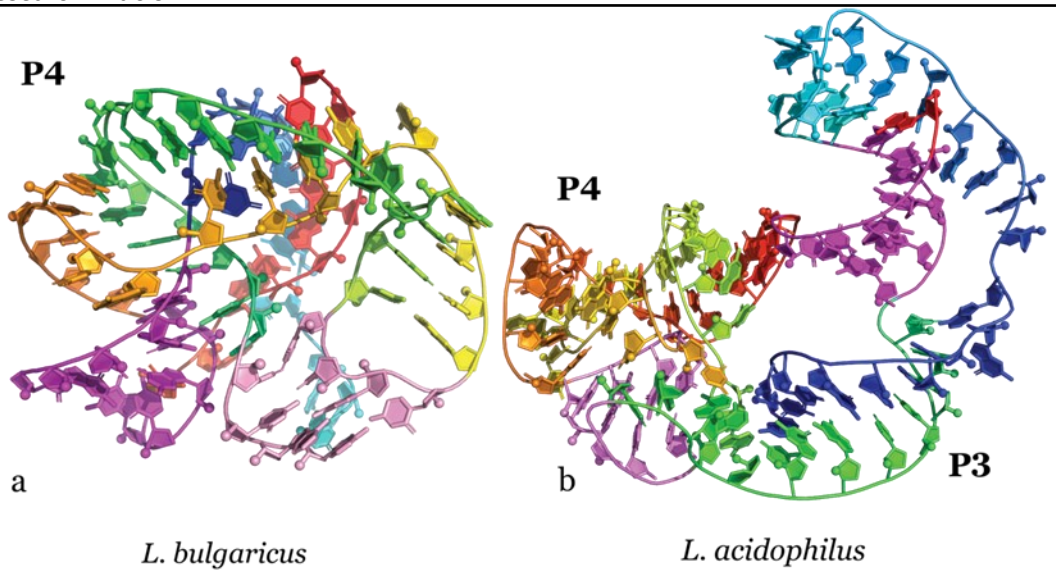

Fig. 7. Tertiary structures of the P4 pseudoknot regions of a) *L. bulgaricus* and b) *L. acidophilus* from FARFAR2 modeling in Rosetta with pseudoknot contacts highlighted in pink and purple.[40] Tertiary structures identified nucleotides exposed for interactions with regulatory proteins or targeted mRNA sequences. The P4 RNA models were visualized in PyMOL (Schrödinger) using RiboVis (Das lab, Stanford) for representation of the nucleic acid structures.[41]

from both sides of the large internal loop that interact with the region preceding the P4 domain (PK1) and with the terminal hairpin loop that defines the middle of the P4 domain (PK2).[44,45]

The 3D models of UspS from *L. bulgaricus* and *L. acidophilus* were evaluated using PyMOL (Schrödinger) to identify nucleotides that are exposed for interactions with regulatory proteins or targeted mRNA sequences.[41] In the model of UspS P4 region from *L. bulgaricus*, both pseudoknot regions are close to each other within a compact structure that includes most of the large internal loop and the terminal hairpin loop. However, there are two loop regions of highly conserved nucleotides that are available for interactions in the tertiary structure. Using the numbering from Fig. 6, both Ade30 and Ura32 bases within the large internal loop are exposed. Within the 2:1 internal loop, both Ade70 and Ura71 bases are also exposed. All four positions are completely conserved in the sixteen sequences from *Lactobacillus*. The model of the UspS P4 region of *L. acidophilus* appears quite different within a very open structure involving both the P3 and P4 regions of the predicted secondary structure. The pseudoknot contacts of PK2 are quite similar in both models. However, the large internal loops that involve PK2 have very different predicted secondary structures, which dictates the more open structure in the model of *L. acidophilus*. Additional experimental data is needed to confirm either or both of these models of the tertiary structure of UspS, which will help to identify potential regulatory interactions.

Predicting sRNA:mRNA Interactions

CopraRNA and IntaRNA were used to predict potential sRNA-mRNA interactions within the P4 region of the UspS sequence.[26–30] The top ten sRNA and mRNA interactions (Table 1) were identified for *L. bulgaricus* and *L. acidophilus*. For both searches, the sRNA query was limited to the P4 region that contains the Lacto-*usp* motif. The genomic DNA sequence of either *L. bulgaricus* or *L. acidophilus* was used as the search basis to map potential mRNA targets with the query limited to ±200 nucleotides around the translation start site. A common theme emerged with nearly all of the top ten interactions of each sRNA mapping to the pseudoknot contacts of PK1. Within the UspS sRNA of *L. bulgaricus*, most of the mRNA interactions mapped to the region 1-21 just preceding the P4 structure. For *L. acidophilus*, the majority of predicted mRNA interactions were mapped to the PK1 side of the internal loop, encompassing residues 73-89.

The potential mRNA targets for regulation by the UspS sRNA are generally related to a stress response involving activation of proteolytic enzymes and amino acid transporters. Notably, there are no regulated genes associated with carbohydrate metabolism or the stringent response. The activation of proteolytic agents may be related to either resource scavenging or remodeling the organism to adapt to a changing environment. Other potential mRNA targets include membrane associated factors such as YfhO family proteins[46] and S-layer proteins[47] that may mediate remodeling of the cell surface in response to environmental changes. The regulation of remodeling by UspS sRNA would then correlate with the predicted interactions with the mRNA genes for primosomal protein N, dihydroorotate dehydrogenase, and carbamoyl phosphate synthase, which may inhibit DNA replication while activating nucleotide synthesis in response to an environmental stress.[48]

The universal stress protein is a nucleotide binding domain that plays a multifaceted role in bacterial stress response (Fig. 8).[49,50] The coding region of a *usp* gene in Lactobacillus is located just downstream of the Lacto-*usp* motif. The UspS sequence is predicted to be within the 5'-UTR of the *usp* gene and may act as a *cis*-regulator from the expression of *usp*. However, the presence of an intrinsic terminator prior to the start of the *usp* gene would also allow the UspS sequence to act as an independent sRNA. Either mode of RNA-directed regulation of *usp* expression would require interactions between the UspS sequence and the *usp* mRNA. The IntaRNA algorithm was used to predict potential interactions between the P4 region of UspS and the coding region of the *usp* mRNA.[28,29] For *L. acidophilus,* IntaRNA predicted three possible interaction sites within the P4 region at positions 53-61 (5'-GCCUUGUUG-3'), 69-75 (5'-CAUAUUC-3'), and 35-41 (5'-CUACCGA-3').[26–30] In the case of *L. bulgaricus*, two possible interaction sites were predicted at positions 13-23 (5'-GGACGUGGUGG-3') and 45-57 (5'-GGAAUGCUGCUCA-3'). For both species, the predicted interaction with the most favorable energy involves the terminal hairpin loop of P4 that is also predicted to form the PK2 pseudoknot contacts. As such, the pseudoknot structures of the P4 region of UspS could mediate a switching mechanism to control sRNA to mRNA interactions.

Structural Homology between UspS and 6S RNA

The P4 region of UspS exhibits a structural resemblance to 6S RNA found in *E. coli*.[31] In other species, 6S RNA is typically associated with increased expression of universal stress proteins.[51] Interestingly, all of the *Lactobacillus* species that possess UspS, lack a predicted 6S RNA.[31] The predicted 6S RNA is only observed in some species of *Lactobacillus*, namely *L. sp. 1.1424, L. plantarum, L. backi, L. salivarius,* and *L. fermentum.* Although 6S RNA was among the first non-coding RNA molecules to be sequenced, its function remained a

Table 1 Predicted UspS sRNA:mRNA interactions within *Lactobacillus*

Rank[a]	*L. bulgaricus* ATCC 11842		*L. acidophilus* ATCC 4356	
	Position sRNA[b]	mRNA associated gene	Position sRNA[b]	mRNA associated gene
1	1-20	glutamine ABC transporter	1-10	glutamine ABC transporter
2	12-21	type I methionyl aminopeptidase	56-66	type I methionyl aminopeptidase
3	13-19	YfhO membrane protein	78-88	YfhO membrane protein
4	1-17	dihydroorotate dehydrogenase	1-17	dihydroorotate dehydrogenase
5	12-20	C40 family peptidase	76-89	C40 family peptidase
6	9-21	primosomal protein N	74-87	SLAP domain surface protein
7	13-91	putative ABC transporter	73-89	primosomal protein N
8	1-21	YqgQ serine endopeptidase	37-58	purine permease
9	66-80	HD domain serine endopeptidase	52-88	carbamoyl phosphate synthase
10	13-19	ATP subunit of Clp protease	1-21	putative ABC transporter

[a] Predicted interactions were ranked by the CopraRNA p-value.
[b] The UspS numbering of the sRNA query sequence matches the numbering in Fig. 6 at the 5' end. Note, that the *L. bulgaricus* sequence has two less residues in the terminal hairpin loop of P4, so the 3' end numbering after residue 54 will differ by two residues.

Table 2 Primers for isolation of UspS and construction of transcription templates

Construct	PCR Primer Sequence (5'-3')[a]	Restriction Enzymes		Size	T_M
		Cloning	Blunt-end	(bp)	(°C)
Ldb P16 forward	GACGAATTCTAATACGACTCACTATAGGATCTGAGTTGATCGTCATAAAGCTTATTTGTA	EcoRI		272	62.5
reverse	CGGCTGCAGGTTTAAAGCGGCCCGAAGGACC	PstI	DraI		62.7
Ldb P15 forward	GACGAATTCTAATACGACTCACTATAGGATCTGAGTTGATCGTCATAAAGCTTATTTGTA	EcoRI		238	62.5
reverse	CGGCTGCAGGATTTAAATCCGAGTTGACCAAAGCATAAATGG	PstI	SwaI		62.8
Lac P16 forward	GACGAATTCTAATACGACTCACTATAGGATCTGAGTTGATCAGCAAGCTTAAATGT	EcoRI		277	62.3
reverse	CGGCTGCAGGTTTAAAGACGGCCCGTTAAGACCG	PstI	DraI		62.9
Lac P15 forward	GACGAATTCTAATACGACTCACTATAGGATCTGAGTTGATCAGCAAGCTTAAATGT	EcoRI		239	62.3
reverse	CGGCTGCAGGATTTAAATCCGAGTTGACTAAACATAGCATAATGA	PstI	SwaI		61.6

[a] The PCR primer sequences include restriction sites for cloning (red), the promoter for T7 RNA polymerase (green), and blunt-end restriction sites (cyan) to prepare the DNA template for *in vitro* transcription. The PCR product sizes and annealing temperatures are shown in the table.

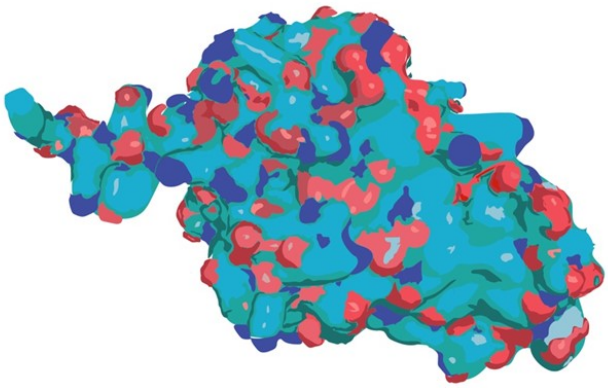

Fig. 8. Universal stress protein is associated with an mRNA that is predicted to interact with UspS in *L. bulgaricus* and *L. acidophilus*. CopraRNA and IntaRNA were used to find the predicted the sRNA:mRNA interactions. NCBI Blast and Uniprot were used to identify the protein functions. The graphical representation was prepared in Adobe Illustrator by Zarah Fowler based on the Alpha Fold model of the universal stress protein from *L. bulgaricus* strain ATCC 11842 (UniProt Q1GAV8).[17,26,41,49,50,55]

mystery for more than three decades.[52,53] In response to starvation in *E. coli*, the increasing accumulation of 6S RNA allows it to form a complex with the sigma specificity factor (σ^{70}) bound to the holoenzyme of RNA polymerase.[53] As the cells enter stationary phase, 6S RNA inhibits the transcription of σ^{70}-dependent genes by RNA polymerase.[53] The large internal loop of the 6S RNA structure mimics the transcription bubble allowing RNA polymerase to bind and even synthesize short pRNA transcripts by RNA-directed transcription.[25] While there is some resemblance between 6S RNA and the Lacto-*usp* motif, the predicted structure of entire the UspS sequence is both smaller and more complex. The central region of 6S RNA forms a large 13:17 nucleotide loop that mimics that transcription bubble. The P4 region of UspS in *L. bulgaricus* has a conserved internal loop at the junction of two stems, but the size of this 9:6 nucleotide loop may be too small to mimic a transcription bubble. The unpaired residues on both strands of this internal loop are predicted to form pseudoknot contacts with other parts of the P4 pseudoknot structure (Fig. 5 and Fig. 6). The other predicted large internal loops in both UspS sequences contain three way junctions, which are unlikely to be transcription bubble mimics (Fig. 5).

Synthesis and Characterization of UspS RNA

For initial structural characterization of UspS, the genomic DNA of *L. bulgaricus* (Ldb) and *L. acidophilus* (Lac) was extracted from cultures grown under anaerobic conditions. The UspS constructs were generated by PCR from genomic DNA using primers that incorporated the promoter for T7 RNA polymerase and a blunt-end restriction site (Table 2). Two constructs were made for each species, one full length that included P1 through P6 regions (P16) and one truncated that removed the intrinsic terminator at P6 (P15). The PCR products were assayed on an agarose gel and the bands observed were consistent with the expected sizes with Ldb P16 at 272 bp, Ldb P15 at 238 bp, Lac P16 at 277 bp, and Lac P15 at 238 bp (Fig. 9).

Prior to *in vitro* transcription, the downstream end of each construct was cleaved with a blunt-end restriction enzyme to define the end point of transcription. RNA was synthesized by *in vitro* transcription using T7 RNA polymerase (NEB). Both the RNA transcripts and the DNA templates were assayed simultaneously on a polyacrylamide gel with appropriate size markers for each (Fig. 10). The observed bands for the cleaved DNA templates were as expected. The RNA transcripts for both constructs from *L. acidophilus* (Lac) were consistent with the expected sizes for Lac P16 at 235 nt and Lac P15 at 195 nt. However, the observed sizes of the RNA transcripts from *L. bulgaricus* (Ldb) were significantly longer than the expected lengths for Ldb P16 at 229 nt and Ldb P15 at 194 nt. Both bands for the *L. bulgaricus* transcript were 50-60 nt larger than the expected size, which may be due to self-priming at the 3' end of the transcript followed by RNA template directed extension.[54] The Ldb P16 and P15 constructs will be redesigned with

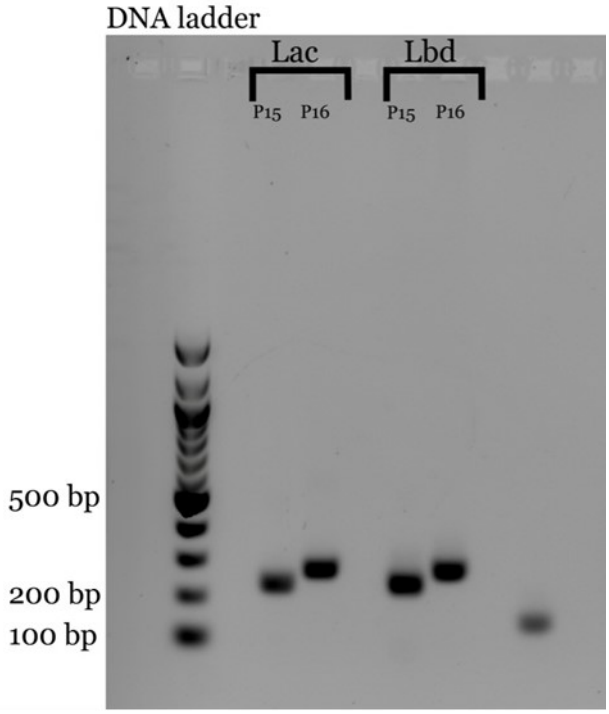

Fig. 9. The UspS sRNA genes were isolated by PCR from *L. acidophilus* (P15 and P16) and *L. bulgaricus* (P15 and P16). The PCR products were assayed by electrophoresis using a 1.5% agarose gel in TAE with ethidium bromide staining (Bio-Rad), run for 45 minutes at 80 V. Gel loading dye Purple (NEB) was included in the samples. The 100 bp DNA Ladder (NEB) was used for reference.

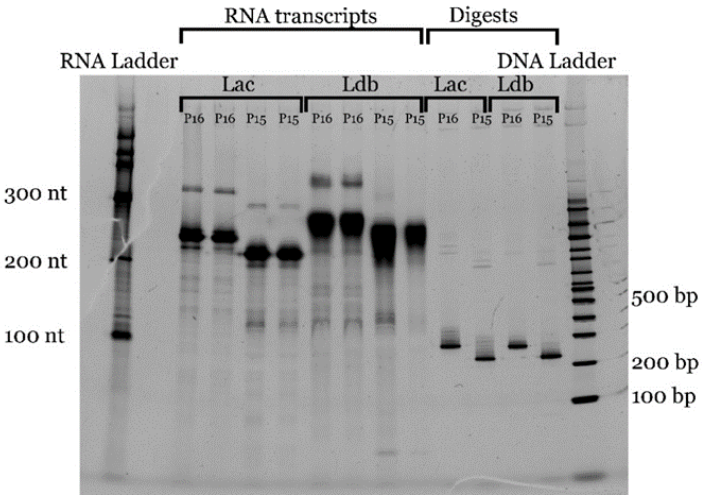

Fig. 10. Transcription reactions and DNA templates were assayed by native PAGE for four UspS constructs: Lac P16 and P15, and Ldb P16, and P15. RNA transcripts: Reactions for each construct at 24 mM and 36 mM $MgCl_2$. Digests: DNA templates digested by a blunt-end restriction enzyme prior to transcription. Both RNA and DNA products were assayed by electrophoresis on a 4-20% Tris-glycine native PAGE gel with Tris-glycine buffer (Invitrogen) for 45 minutes at 200 V. The RiboRuler Low Range RNA Ladder (Thermo Scientific) and the 100 bp DNA Ladder (NEB) were used as reference. The gel was stained with SYBR Green II (Molecular Probes).

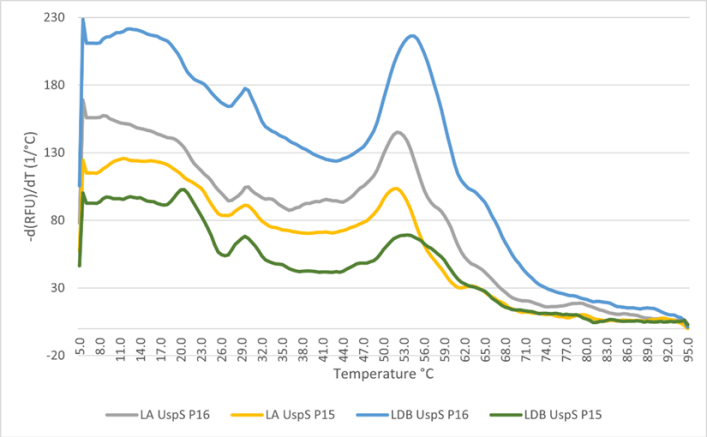

Fig. 11. Differential scanning fluorimetry of UspS sRNA transcripts. RNA transcripts: Lac P16 and P15, and Ldb P16 and P15. All assays shown were performed with 0 mM $MgCl_2$.[43] The non-linear change in fluorescence intensity with temperature is correlated with the thermal denaturation of the structure of the sRNA. The peaks indicate the melting temperature of secondary or tertiary structure within the sRNA. Well-defined unfolding transitions were found near 54°C for both Ldb constructs and at 52° C for both Lac constructs.

sequence variations at the 3' end to find a better template for these RNA transcripts.

Structural Stability of UspS RNA

The initial characterization of the RNA transcripts used differential scanning fluorimetry to assay for secondary or tertiary structure.[43] All four constructs were assayed for thermal stability in the presence of a fluorescent reporter dye at two different concentrations of $MgCl_2$. The fluorescence intensity was observed for each RNA sample as the temperature increased. A non-linear change in the decrease fluorescence intensity with temperature is correlated with the thermal denaturation of a structured RNA molecule either due to the loss of base-pairing interactions or tertiary contacts between nucleotides. The analysis of the thermal melt curves uses the negative of the first derivative of the change in fluorescence intensity with temperature to create and easy to interpret graph where peaks may indicate the melting temperature of various secondary or tertiary structure domains within an RNA molecule. For all four UspS constructs, the first derivative graphs show a well-defined unfolding transition near 54°C for both *Lab* constructs and at 52°C for both (Lac) constructs. For both species, the addition of the intrinsic terminator (P6) does not affect the melting temperature but does improve the intensity of the derivative peak, perhaps by minimizing single-stranded regions of the construct that may influence the dynamics of interactions with the fluorescence dye. The thermal stability of the UspS constructs indicates the presence of the predicted secondary structural elements, which can be further assayed by other methods, such as RNase T_1 digests to map the loop regions. Smaller UspS constructs that focus on the P4 pseudoknot domain can be assayed for stability by differential scanning fluorimetry in preparation for structural studies by either X-ray crystallography or NMR spectroscopy.

Conclusion

The investigation into the sRNA UspS in *Lactobacillus* species has provided valuable insight into its conservation and potential molecular functions. Multiple sequence alignment revealed significant conservation of sequences between *Lactobacillus* species, suggesting that this sRNA is essential for the molecular functions of lactobacilli.[32] Furthermore, secondary structure predictions show structural conservation of UspS P4 region between *L. acidophilus* and *L. bulgaricus* suggesting this is a significant site for mRNA interactions.[26–30] The conserved structure of the P4 region suggests a potential correspondence between UspS and 6S RNA in *Lactobacillus* species, as most lactobacilli lack a predicted 6S RNA. Consequently, UspS may function similarly to 6S RNA in *E. coli* by forming a complex with the sigma subunit of RNA polymerase and inhibiting transcription.[25,36,51] Through three-dimensional modeling and genetic mapping, potential sites for translational control of protein synthesis were identified, providing further insight into the functional implications of UspS.[40] Notably, the presence of a universal stress protein downstream of UspS in both *L. acidophilus* and *L. bulgaricus* suggests that UspS may interact with the mRNA of a universal stress protein. Other predicted mRNA interactions indicate that UspS may activate a change in the bacterial membrane structure in response to stress.[26–30] In future studies, the structure of the P4 pseudoknot will be further characterized for a better understanding of the possible homology with 6S RNA.

Acknowledgements

The authors gratefully acknowledge Dr. Gabriela C. Pérez Alvarado who provided invaluable assistance with laboratory experiments and insightful discussions about research. This project was supported by grant P20GM103499 (SC INBRE) from the National Institute of General Medical Sciences of the National Institutes of Health, and by startup funds to B. M. Lee from the Gupta College of Science at Coastal Carolina University.

Notes and References

**Corresponding author email:* brianlee@coastal.edu

1. Nurk S, Koren S, Rhie A, Rautiainen M, Bzikadze AV, Mikheenko A, Vollger MR, Altemose N, Uralsky L, Gershman A, et al. The complete sequence of a human genome. Science. 2022;376(6588):44–53. doi:10.1126/science.abj6987
2. 2003: Human Genome Project Completed. Genome.gov. 2022 Sep 14 [accessed 2023 Jun 14]. https://www.genome.gov/25520492/online-education-kit-2003-human-genome-project-completed
3. Gilbert JA, Blaser MJ, Caporaso JG, Jansson JK, Lynch SV, Knight R. Current understanding of the human microbiome. Nature Medicine. 2018;24(4):392–400. doi:10.1038/nm.4517
4. Gibbs RA. The Human Genome Project changed everything. Nature Reviews. Genetics. 2020;21(10):575–576. doi:10.1038/s41576-020-0275-3
5. Cooke Bailey JN, Pericak-Vance MA, Haines JL. The Impact of the Human Genome Project on Complex Disease. Genes. 2014;5(3):518–535. doi:10.3390/genes5030518
6. NIH Human Microbiome Project defines normal bacterial makeup of the body. National Institutes of Health (NIH). 2015 Aug 31 [accessed 2023 Jun 9]. https://www.nih.gov/news-events/news-releases/nih-human-microbiome-project-defines-normal-bacterial-makeup-body
7. Hasan N, Yang H. Factors affecting the composition of the gut microbiota, and its modulation. PeerJ. 2019;7:e7502. doi:10.7717/peerj.7502
8. Williams NT. Probiotics. American Journal of Health-System Pharmacy. 2010;67(6):449–458. doi:10.2146/ajhp090168
9. Domínguez Rubio AP, D'Antoni CL, Piuri M, Pérez OE. Probiotics, Their Extracellular Vesicles and Infectious Diseases. Frontiers in Microbiology. 2022;13:864720. doi:10.3389/fmicb.2022.864720
10. Crick FH. On protein synthesis. Symposia of the Society for Experimental Biology. 1958;12:138–163.
11. Liu G, Chang H, Qiao Y, Huang K, Zhang A, Zhao Y, Feng Z. Profiles of Small Regulatory RNAs at Different Growth Phases of Streptococcus thermophilus During pH-Controlled Batch Fermentation. Frontiers in Microbiology. 2021;12:765144. doi:10.3389/fmicb.2021.765144
12. Felden B, Augagneur Y. Diversity and Versatility in Small RNA-Mediated Regulation in Bacterial Pathogens. Frontiers in Microbiology. 2021;12:719977. doi:10.3389/fmicb.2021.719977
13. Djapgne L, Oglesby AG. Impacts of Small RNAs and Their Chaperones on Bacterial Pathogenicity. Frontiers in Cellular and Infection Microbiology. 2021;11:604511. doi:10.3389/fcimb.2021.604511
14. Jørgensen MG, Pettersen JS, Kallipolitis BH. sRNA-mediated control in bacteria: An increasing diversity of regulatory mechanisms. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. 2020;1863

(5):194504. doi:10.1016/j.bbagrm.2020.194504
15. Cai Q, Qiao L, Wang M, He B, Lin F-M, Palmquist J, Huang S-D, Jin H. Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes. Science. 2018;360(6393):1126–1129. doi:10.1126/science.aar4142
16. Sansen J, Thebault P, Dutour I, Bourqui R. Visualization of sRNA-mRNA Interaction Predictions. In: 2016 20th International Conference Information Visualisation (IV). Lisbon: IEEE; 2016. p. 342–347. https://ieeexplore.ieee.org/document/7557950/. doi:10.1109/IV.2016.14
17. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990;215(3):403–410. doi:10.1016/S0022-2836(05)80360-2
18. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009;25(9):1189–1191. doi:10.1093/bioinformatics/btp033
19. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research. 2003;31(13):3406–3415. doi:10.1093/nar/gkg595
20. Zok T, Antczak M, Zurkowski M, Popenda M, Blazewicz J, Adamiak RW, Szachniuk M. RNApdbee 2.0: multifunctional tool for RNA structure annotation. Nucleic Acids Research. 2018;46(W1):W30–W35. doi:10.1093/nar/gky314
21. Antczak M, Zok T, Popenda M, Lukasiak P, Adamiak RW, Blazewicz J, Szachniuk M. RNApdbee—a webserver to derive secondary structures from pdb files of knotted and unknotted RNAs. Nucleic Acids Research. 2014;42(W1):W368–W372. doi:10.1093/nar/gku330
22. Watkins AM, Rangan R, Das R. FARFAR2: Improved De Novo Rosetta Prediction of Complex Global RNA Folds. Structure. 2020;28(8):963-976.e6. doi:10.1016/j.str.2020.05.011
23. Weinberg Z, Wang JX, Bogue J, Yang J, Corbino K, Moy RH, Breaker RR. Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. Genome Biology. 2010;11(3):R31. doi:10.1186/gb-2010-11-3-r31
24. Barrick JE, Sudarsan N, Weinberg Z, Ruzzo WL, Breaker RR. 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. RNA. 2005;11(5):774–784. doi:10.1261/rna.7286705
25. Wassarman KM. 6S RNA, a Global Regulator of Transcription Storz G, Papenfort K, editors. Microbiology Spectrum. 2018;6(3):6.3.06. doi:10.1128/microbiolspec.RWR-0019-2018
26. Wright PR, Georg J, Mann M, Sorescu DA, Richter AS, Lott S, Kleinkauf R, Hess WR, Backofen R. CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains. Nucleic Acids Research. 2014;42(W1):W119–W123. doi:10.1093/nar/gku359
27. Wright PR, Richter AS, Papenfort K, Mann M, Vogel J, Hess WR, Backofen R, Georg J. Comparative genomics boosts target prediction for bacterial small RNAs. Proceedings of the National Academy of Sciences. 2013 [accessed 2023 Jun 13];110(37). https://pnas.org/doi/full/10.1073/pnas.1303248110. doi:10.1073/pnas.1303248110
28. Mann M, Wright PR, Backofen R. IntaRNA 2.0: enhanced and customizable prediction of RNA–RNA interactions. Nucleic Acids Research. 2017;45(W1):W435–W439. doi:10.1093/nar/gkx279
29. Busch A, Richter AS, Backofen R. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. Bioinformatics. 2008;24(24):2849–2856. doi:10.1093/bioinformatics/btn544
30. Raden M, Ali SM, Alkhnbashi OS, Busch A, Costa F, Davis JA, Eggenhofer F, Gelhausen R, Georg J, Heyne S, et al. Freiburg RNA tools: a central online resource for RNA-focused research and teaching. Nucleic Acids Research. 2018;46(W1):W25–W29. doi:10.1093/nar/gky329
31. Weinberg Z, Wang JX, Bogue J, Yang J, Corbino K, Moy RH, Breaker RR. Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. Genome Biology. 2010;11(3):R31. doi:10.1186/gb-2010-11-3-r31
32. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009;25(9):1189–1191. doi:10.1093/bioinformatics/btp033
33. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology. 2011;7(1):539. doi:10.1038/msb.2011.75
34. McCracken A, Turner MS, Giffard P, Hafner LM, Timms P. Analysis of promoter sequences from Lactobacillus and Lactococcus and their activity in several Lactobacillus species. Archives of Microbiology. 2000;173(5–6):383–389. doi:10.1007/s002030000159
35. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research. 2003;31(13):3406–3415. doi:10.1093/nar/gkg595
36. Kalvari I, Nawrocki EP, Ontiveros-Palacios N, Argasinska J, Lamkiewicz K, Marz M, Griffiths-Jones S, Toffano-Nioche C, Gautheret D, Weinberg Z, et al. Rfam 14: expanded coverage of metagenomic, viral and microRNA families. Nucleic Acids Research. 2021;49(D1):D192–D200. doi:10.1093/nar/gkaa1047
37. Rivas E, Clements J, Eddy SR. A statistical test for conserved RNA structure shows lack of evidence for structure in lncRNAs. Nature Methods. 2017;14(1):45–48. doi:10.1038/nmeth.4066
38. Antczak M, Zok T, Popenda M, Lukasiak P, Adamiak RW, Blazewicz J, Szachniuk M. RNApdbee—a webserver to derive secondary structures from pdb files of knotted and unknotted RNAs. Nucleic Acids Research. 2014;42(W1):W368–W372. doi:10.1093/nar/gku330
39. Zok T, Antczak M, Zurkowski M, Popenda M, Blazewicz J, Adamiak RW, Szachniuk M. RNApdbee 2.0: multifunctional tool for RNA structure annotation. Nucleic Acids Research. 2018;46(W1):W30–W35. doi:10.1093/nar/gky314
40. Watkins AM, Rangan R, Das R. FARFAR2: Improved De Novo Rosetta Prediction of Complex Global RNA Folds. Structure. 2020;28(8):963-976.e6. doi:10.1016/j.str.2020.05.011
41. Warren L. DeLano. The PyMOL Molecular Graphics System. 2021. https://pymol.org/2/
42. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. Primer3—new capabilities and interfaces. Nucleic Acids Research. 2012;40(15):e115–e115. doi:10.1093/nar/gks596
43. Silvers R, Keller H, Schwalbe H, Hengesbach M. Differential Scanning Fluorimetry for Monitoring RNA Stability. ChemBioChem. 2015;16(7):1109–1114. doi:10.1002/cbic.201500046
44. Peselis A, Serganov A. Structure and function of pseudoknots involved in gene expression control: Structure and function of pseudoknots. Wiley Interdisciplinary Reviews: RNA. 2014;5(6):803–822. doi:10.1002/wrna.1247
45. Antczak M, Popenda M, Zok T, Zurkowski M, Adamiak RW, Szachniuk M. New algorithms to represent complex pseudoknotted RNA structures in dot-bracket notation Valencia A, editor. Bioinformatics. 2018;34(8):1304–1312. doi:10.1093/bioinformatics/btx783
46. Rismondo J, Percy MG, Gründling A. Discovery of genes required for lipoteichoic acid glycosylation predicts two distinct mechanisms for wall teichoic acid glycosylation. Journal of Biological Chemistry. 2018;293(9):3293–3306. doi:10.1074/jbc.RA117.001614
47. Johnson BR, Hymes J, Sanozky-Dawes R, Henriksen ED, Barrangou R, Klaenhammer TR. Conserved S-Layer-Associated Proteins Revealed by Exoproteomic Survey of S-Layer-Forming Lactobacilli Nojiri H, editor. Applied and Environmental Microbiology. 2016;82(1):134–145. doi:10.1128/AEM.01968-15
48. Bonilla CY, Grossman AD. The Primosomal Protein DnaD Inhibits Cooperative DNA Binding by the Replication Initiator DnaA in Bacillus subtilis. Journal of Bacteriology. 2012;194(18):5110–5117. doi:10.1128/JB.00958-12
49. The UniProt Consortium, Bateman A, Martin M-J, Orchard S, Magrane M, Ahmad S, Alpi E, Bowler-Barnett EH, Britto R, Bye-A-Jee H, et al. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Research. 2023;51(D1):D523–D531. doi:10.1093/nar/gkac1052
50. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583–589. doi:10.1038/s41586-021-03819-2
51. Neusser T, Polen T, Geissen R, Wagner R. Depletion of the non-coding regulatory 6S RNA in E. coli causes a surprising reduction in the expression of the translation machinery. BMC Genomics. 2010;11(1):165. doi:10.1186/1471-2164-11-165
52. Brownlee GG. Sequence of 6S RNA of E. coli. Nature New Biology. 1971;229(5):147–149. doi:10.1038/newbio229147a0

53. Wassarman KM, Storz G. 6S RNA Regulates E. coli RNA Polymerase Activity. Cell. 2000;101(6):613–623. doi:10.1016/S0092-8674(00)80873-9
54. Gholamalipour Y, Karunanayake Mudiyanselage A, Martin CT. 3′ end additions by T7 RNA polymerase are RNA self-templated, distributive and diverse in character—RNA-Seq analyses. Nucleic Acids Research. 2018;46 (18):9253–9263. doi:10.1093/nar/gky796
55. Adobe Illustrator. 2023. https://www.adobe.com/products/illustrator.html
56. Peselis A, Serganov A. Structure and function of pseudoknots involved in gene expression control: Structure and function of pseudoknots. Wiley Interdisciplinary Reviews: RNA. 2014;5(6):803–822. doi:10.1002/ wrna.1247
57. Antczak M, Popenda M, Zok T, Zurkowski M, Adamiak RW, Szachniuk M. New algorithms to represent complex pseudoknotted RNA structures in dot-bracket notation Valencia A, editor. Bioinformatics. 2018;34(8):1304–1312. doi:10.1093/bioinformatics/btx783