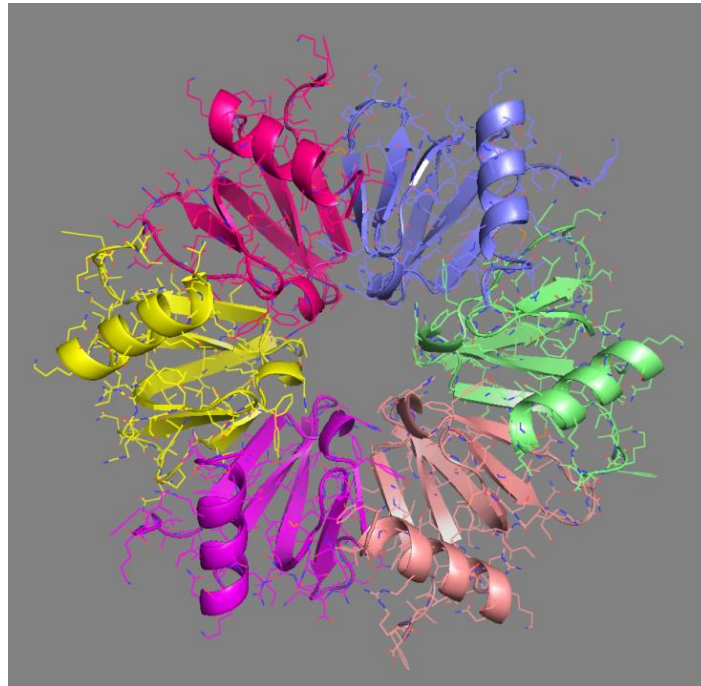


The Hfq RNA chaperone in the deep-branching Thermotogales lineage:

Attempts to reveal its biological role



**Master's Thesis in Biology
Specialization in Microbiology**

Oda Barth Vedøy



University of Bergen

2014

Abstract

Bacteria frequently use small RNAs (sRNAs) as part of their regulation of gene expression at the post-transcriptional level. Many of these sRNAs have been shown to depend on the RNA chaperone protein Hfq for their regulatory properties. Homologs of Hfq are found in Archaea and Eukarya, suggesting a common evolutionary predecessor in the last universal common ancestor. The function of Hfq has predominantly been studied in model species and pathogenic bacteria. Yet, no studies have been made on the role of Hfq in deep phylogenetic lineages of bacteria or in thermophiles. The aim of this study was to determine the role of Hfq in gene expression in Thermotogales, a deeply branching and thermophilic lineage. Hfq from *Thermosipho africanus* (*TaHfq*) was expressed in *Escherichia coli* as an insoluble protein and anti-Hfq antibodies were raised in rabbits. Western blot analysis using anti-*TaHfq* antiserum confirmed that *hfq* was expressed as a soluble protein in *T. africanus*. A high degree of conservation in amino acid sequence and protein motifs was confirmed by multiple sequence alignment, and a homology model of *TaHfq* indicated a similar 3D structure as determined by X-ray crystallography of Hfq from other bacteria, including the presence of a similar RNA-binding pocket. This suggests that Hfq has a role in gene expression in Thermotogales similar to that in model species. This is also supported by the fact that *hfq* is conserved in the genome-sequences of all the Thermotogales species.

Co-immunoprecipitation of Hfq from lysates of *T. africanus* cells with anti-*TaHfq* antiserum was performed in order to pull out sRNAs interacting with Hfq, followed by cDNA synthesis. Although the technical procedure appeared functional, it was not possible to isolate cDNA representing sRNAs from *T. africanus*. This may be the result of a very low expression level of Hfq in *T. africanus*, as subsequently demonstrated by Proteomics analysis. It is possible that *TaHfq* is only expressed under certain conditions, such as under stress conditions as found in several other bacteria, or in its natural habitat, where this organism is exposed to a variety of extreme conditions.

Acknowledgements

The work presented in this thesis was carried out during the period of August 2012 to November 2014 at the Department of Biology, University of Bergen.

I would like to thank my supervisor Professor Nils Kåre Birkeland for helping me find an exciting project where I have gained insight into several different microbiological and molecular methods. I started from zero, and would like to thank you for having the patience to see me through this thesis work.

I would also like to thank technicians Marit Madsen and especially Birte Töpper for all their guidance and help in the lab. Thanks also to Helge Erikstad and Antonio Garcia-Moyano, you always seem to have the time to help a frustrated master student. I am thankful to have been working in the nice environment of the microbiology group.

Thanks also to my family for doing their best in trying to seem interested when I've been sharing details of the challenges in my thesis work. For feed-back and encouragements; thanks to my aunt Tanja, Yngvild, Calum and Emiliy. Thanks to Anne Marie, Natalie, Leanna and Trude for company in the lab, long-lasting lunch breaks and study trips.

I especially want to thank Martin for taking so good care of Frida, making it no problem for me to return to the thesis early on. You have also lowered my frustration level significantly with your computer help. I admire your patience.

Table of contents

Abstract	3
Acknowledgements	4
1. Introduction.....	8
1.1 Gene expression and regulation in bacteria.....	8
1.1.1 The role of sRNAs in bacterial gene regulation	9
1.2 Hfq proteins.....	9
1.3 Experimental strategies for identifying RNA targets for Hfq	14
1.4 Thermotogae	17
1.4.1 <i>T. africanus</i>	19
1.4.2 <i>K. olearia</i>	20
1.4.3 <i>P. mobilis</i>	21
1.5 Aim of the study	21
2. Materials and methods	22
2.1 Bacterial strains and cultivation	23
2.2 Extraction and quantification of DNA.....	27
2.2.1 DNA extraction by the cetyltrimethylammonium bromide (CTAB) method.....	27
2.2.2 DNA quantification by NanoDrop.....	27
2.2.3 Agarose Gel Electrophoresis.....	28
2.3 Gene amplification	28
2.3.1 Primer design for polymerase chain reaction (PCR).....	28
2.3.2 PCR.....	30
2.3.3 Purification of the PCR product	30
2.4 Cloning.....	31
2.4.1 pET-21b(+) vector	31
2.4.2 Digestion of PCR-product and vector	31
2.4.3 Ligation	31
2.4.4 Transformation for sequencing.....	32
2.4.5 Plasmid clone purification	32
2.4.6 Confirmation of plasmid inserts	33
2.5 Overexpression of cloned genes	35
2.5.1 Transformation for gene expression	35
2.5.2 Induction of gene expression	35
2.5.3 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)	36

2.6	Testing the solubility of the Hfq protein	37
2.7	Antibody production	37
2.7.1	Western blotting.....	38
2.8	Identification of Hfq-bound RNAs	39
2.8.1	Estimation of exponential phase.....	39
2.8.2	Cultivation of <i>T. africanus</i> at threshold temperatures.....	40
2.8.3	Co-immunoprecipitation (co-IP).....	40
2.8.4	Isolation of RNA.....	41
2.8.5	cDNA synthesis	43
2.8.6	Sequencing of cDNA	46
2.8.7	Sequencing of single cDNAs	46
2.9	Ouchterlony double immunodiffusion	48
2.10	Proteomics analysis	48
2.11	Bioinformatics	49
2.11.1	Phylogenetic tree.....	49
2.11.2	Multiple sequence alignment.....	49
2.11.3	Prediction of protein structure.....	49
3.	Results	50
3.1	Cloning.....	50
3.1.1	Preparation of genomic DNA.....	50
3.1.2	Gene amplification	51
3.1.3	Digestion of PCR-products and vector.....	52
3.1.4	Ligation and transformation.....	53
3.2	Overexpression.....	55
3.3	<i>T. africanus</i> growth curve.....	57
3.4	Western blot analysis	57
3.5	RNA work.....	62
3.5.1	Pull down of Hfq/RNA complexes and cDNA synthesis.....	62
3.6	Sequencing of cDNA	64
3.7	Proteomics.....	67
3.8	Bioinformatics	68
3.8.1	Phylogenetic tree based on Hfq and 16S rRNA	68
3.8.2	Multiple sequence alignment of Hfq from Thermotogales and model species	69
3.8.3	Prediction of protein structure.....	71

4. Discussion	74
4.1 Discussion of materials and methods.....	74
4.1.1 Gene expression in <i>E. coli</i>	74
4.1.2 Protein solubility.....	75
4.1.3 RNA-extraction	76
4.2 Discussion of results	77
4.2.1 Exclusion of <i>K. olearia</i> and <i>P. mobilis</i>	77
4.2.2 Weak transfer of proteins	77
4.2.3 Cross reaction with secondary antibody.....	77
4.2.4 RNA extraction, cDNA synthesis and sequencing.....	78
4.2.5 Possible reasons for unsuccessful RNA precipitation.....	79
4.2.6 Bioinformatics.....	79
5. Conclusions.....	81
6. Further work.....	82
7. References.....	83
8. Appendices	88
8.1 Appendix A: Results.....	88
8.2 Appendix B: Solutions and buffers	93
8.3 Appendix C: Materials and methods.....	98

1. Introduction

1.1 Gene expression and regulation in bacteria

Gene expression is the process where DNA is transcribed into mRNA, which eventually is translated into proteins. This is an essential but energy demanding cellular process. Not all proteins are needed at all times, and to synthesize all proteins at a constant rate would be an extreme waste of energy, and cause metabolic chaos. Regulation of gene expression is therefore an important feature of all cells. Metabolic regulation can be performed at three levels; at the transcription level, post-transcription level and at the protein level. Regulation at the protein level is the most effective, as proteins can be activated or inactivated in seconds. This enables a quick response to a changing environment. Regulation at the transcriptional level is the most energy saving form of regulation because it prevents the cell from making transcripts and proteins which are not needed. Post-transcriptional regulation is regulation of mRNAs, and during the last 20 years post-transcriptional regulation has been recognized as a widespread regulatory mechanism in bacteria (Valverde 2012). Regulatory RNAs (regRNAs) are often shown to play a role in post-transcriptional regulation. The effects of these regRNAs can lead to up- or down regulation of gene expression as described below. There are two types of regRNAs; *cis*-encoded and *trans*-encoded. Examples of *cis*-encoded regRNAs are riboswitches and attenuators. *Cis*-encoded regRNAs are encoded on the same mRNA strand that is being regulated, and regulates in a 1:1 stoichiometric ratio. *Trans*-encoded regRNAs are transcribed from a separate DNA region. These regRNAs share short and imperfect base pairing with their target mRNAs, and may regulate several individual mRNAs transcribed from a separate region (Waters and Storz 2009). In bacteria, regulation by *trans*-encoded regRNAs often requires the RNA-binding protein Hfq, which is the focus of this thesis.

Trans-encoded regRNAs are also called small non-coding RNAs (sRNAs) because of their small size (approximately 50-300 base pairs) (Storz, Vogel et al. 2011), and the fact that they do not translate into proteins (as opposed to some regRNAs). In some publications, *cis*-encoded regRNAs are also referred to as sRNA (Storz and Haas 2007, Waters and Storz 2009, Liu and Camilli 2010, Faner and Feig 2013). The term used in this way can be misleading as a *cis*-encoded regRNA is a part of a larger mRNA. In this thesis only *trans*-encoded regRNAs are

referred to as sRNA, as was done in the review article by Sobrero and Valverde from 2012 (Sobrero and Valverde 2012).

1.1.1 The role of sRNAs in bacterial gene regulation

Trans-encoded sRNAs are known to regulate translation by base pairing with target mRNAs, thereby altering its secondary structure. This base pairing can lead to an up- or down regulation of gene expression through affecting translation, but most often a blocking of gene expression is observed. Blocking happens when sRNA binds to the ribosome binding site (rbs), thereby preventing binding of the ribosome to the transcript. In some cases binding of sRNA has shown to alter the secondary structure of an mRNA, making the rbs available for ribosome binding. This mechanism has been observed in *Vibrio cholerae* and *Escherichia coli* (Hammer and Bassler 2007, Urban and Vogel 2008). Blocking of the rbs by sRNA is a reversible form of regulation as translation can continue if the sRNA detaches from the rbs. An irreversible form of regulation occurs when binding of sRNA leads to degradation of the sRNA/mRNA complex by RNase. Hfq is a small protein that in many cases seems to be necessary for sRNA binding. This assumption is based on studies showing that sRNAs are unable or have an impaired ability to bind mRNA when Hfq is absent or limited, and also that the recruitment of RNase is often facilitated by the Hfq protein (Aiba 2007, Caron, Lafontaine et al. 2010).

1.2 Hfq proteins

The RNA-binding protein Hfq was first discovered in 1968 by Franze de Fernandez as a host factor for the RNA phage ϕ Q β in *E. coli* (Franze de Fernandez, Eoyang et al. 1968). Later studies have shown that Hfq is a highly conserved protein, with homologs found in Archaea and Eukarya (Sauter, Basquin et al. 2003). These homologous proteins are ascribed to the family of Sm or LSm-proteins. They were first discovered in eukaryotes, and called “smith antigen” or Sm proteins after Stephanie Smith, the patient in which they were first found (Notman, Kurata et al. 1975). The assumption of an evolutionary relatedness between these proteins is based on sequence comparisons, protein structure and RNA-binding sites. This is discussed in the 2011 review article by Murina and Nikulin (Murina and Nikulin 2011). A

sequence comparison revealed two conserved regions, one central region of about 29 amino acids (the Sm1 motif), and a second region with about 14 amino acids (the Sm2 motif) (Murina and Nikulin 2011). These highly conserved regions present in all domains of life suggests that a Hfq homolog may have been present in the LUCA (Last Universal Common Ancestor). Hfq acts as an RNA chaperone for sRNAs, and it is suggested that this protein may be encoded in up to half of all bacterial species (Valentin-Hansen, Eriksen et al. 2004). Hfq has been shown to play a key role in regulation of stress response (Peng, Soper et al. 2014), and also in the fitness and virulence of numerous bacterial pathogens (Chao and Vogel 2010, Liu, Wu et al. 2010, Schiano, Bellows et al. 2010). The protein is thought to have an impact on the expression of certain genes, and in some organisms regulation of up to 20% of the genes depend partly on Hfq proteins (Chao and Vogel 2010). When binding to small RNAs, Hfq proteins can alter and/or stabilize the secondary structure of the RNA leading to the formation of an alternative sRNA secondary structure that binds to mRNA (Liu, Wu et al. 2010). In this way, Hfq acts as a host factor for post-transcriptional regulation. It has also been recognized for having a stabilizing effect on small RNAs, and to stimulate RNA pairing between small RNAs and mRNAs (Soper, Mandin et al. 2010). It is widely accepted that Hfq can have five different modes of action, illustrated in Fig. 2. Hfq can act as a facilitator for small RNAs to bind to mRNA in *trans*, thereby stimulating or inhibiting translation (Fig. 2a and b). There are examples of both up- and down-regulation of gene expression by Hfq, but down-regulation is the most commonly observed effect. This happens when the Hfq-sRNA complex binds to the target mRNA thereby blocking or freeing the rbs for ribosome binding. Hfq binding may protect sRNA from degradation by ribonuclease cleavage, but has also been shown to lead to degradation of the sRNA-mRNA complex (Fig. 2c and d). Upon mRNA binding, Hfq may stimulate polyadenylation by poly(A) polymerase (PAP), leading to a 3'-5' degradation by an exoribonuclease (Fig. 2e).

Hfq is composed of six subunits, which come together in a doughnut-shaped formation (Soper, Mandin et al. 2010). Each subunit is composed of one α -helix and five β -sheets. The RNA-binding motifs of Hfq have been determined by crystal structure studies and binding and mutagenesis studies. These have shown that Hfq interaction preferentially occurs between uracil-rich RNAs on the proximal face, and with adenine-rich RNAs on the distal face (Link, Valentin-Hansen et al. 2009). The crystal structure of the RNA AU₅G bound to

Staphylococcus aureus Hfq revealed the binding properties of Hfq (Schumacher, Pearson et al. 2002). Hfq binds this RNA in a circular manner around its central pore on the proximal face. The adenine and uracils bind Hfq in nucleotide pockets made up by the Sm1 and Sm2 motifs from two neighbouring subunits of the Hfq hexamer. The guanine penetrates the central pore and binds there. The amino acids involved in binding are tyrosine, lysine and glutamine, whereas tyrosine from two neighbouring subunits forms the sides of the “pocket” where RNA binds. Lysine anchors the base to its pocket. Glutamine is the only amino acid that is 100% conserved within the α -helix of the Hfq protein and is involved in binding of the uracil base. The binding of adenine and uracil occurs in a similar way, but adenine has a stronger binding due to its larger size.

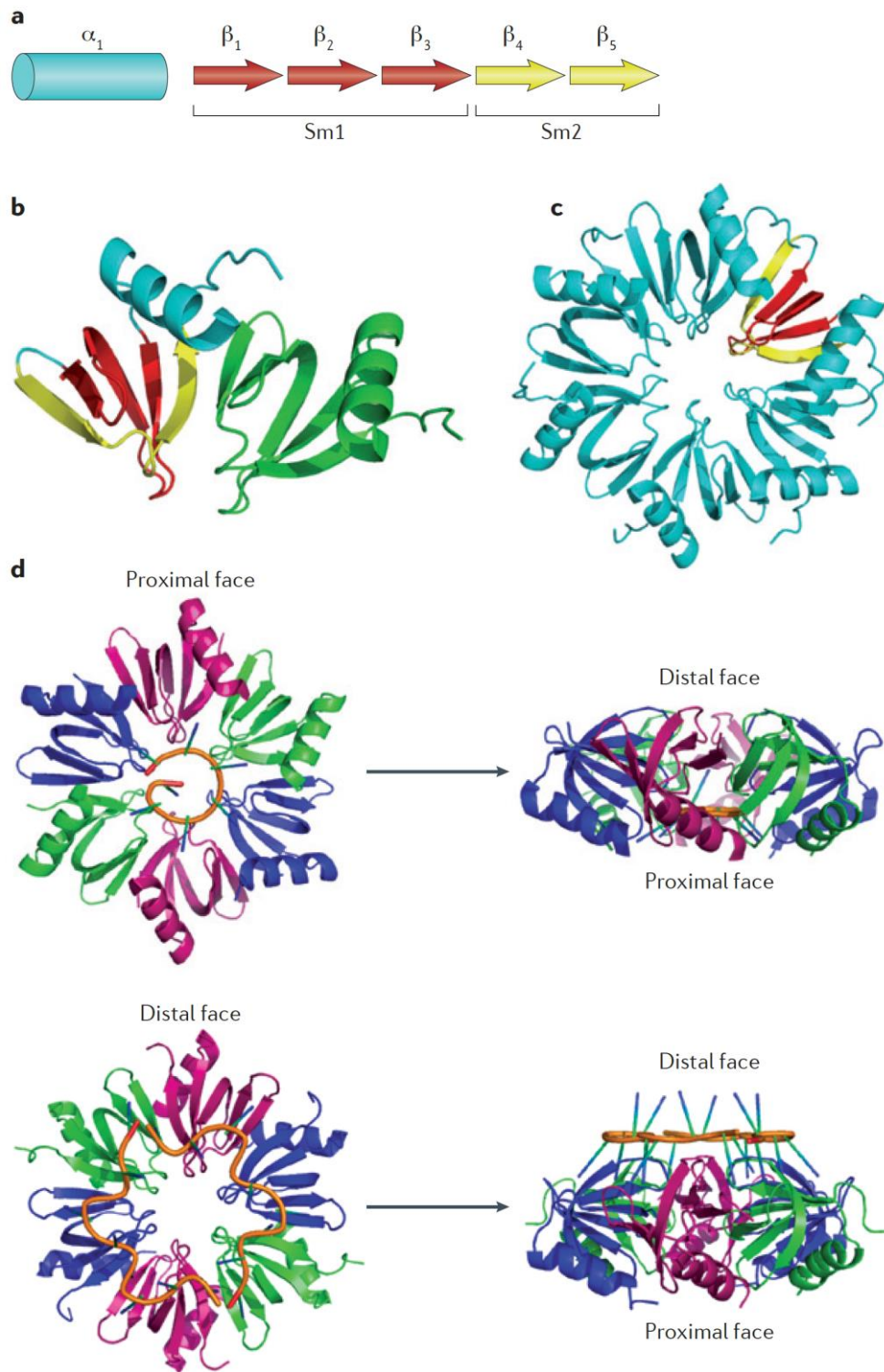


Figure 1: The structure of Hfq and its interactions with RNA (from Vogel and Luisi 2011). **a** | Secondary structure revealing the two conserved regions of the Hfq protein, the Sm1 motif and the Sm2 motif. **b** | Tertiary structure showing the interaction between the β_4 -sheet on one subunit and the β_5 -sheet on the neighbouring subunit. **c** | Quaternary structure revealing the ring structure of the Hfq protein, highlighting one subunit. Each subunit is composed of one α -helix and five β -sheets. The β_4 and β_5 sheets in the periphery are involved in binding of the subunits through hydrogen-bonding edges. **d** | Two sites for interactions with RNAs (orange) are located on opposite sides of the Hfq protein, on the proximal and the distal faces. The α -helices are exposed on the proximal face.

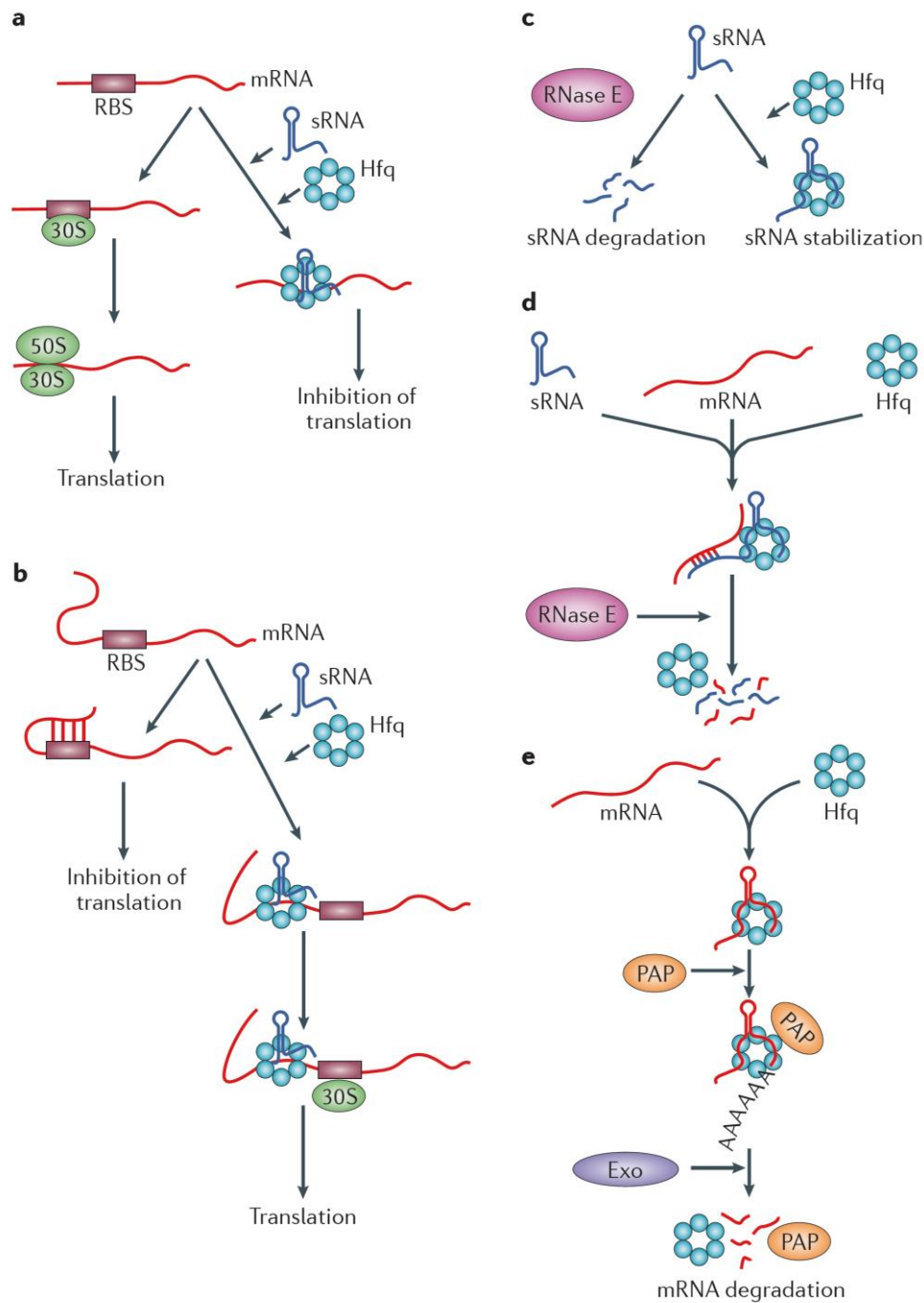


Figure 2: Widely accepted modes of Hfq activity (from Vogel and Luisi 2011). **a** | Hfq may bind sRNA in the ribosome binding site (rbs), hereby preventing ribosomes from binding the transcript. This action will lead to a down-regulation of gene expression. **b** | Hfq may bind sRNA in a region prior to the rbs, in the 5' end of a transcript, hereby exposing the rbs and allowing binding of a ribosome. This action will lead to an up-regulation of gene expression. **c** | By Hfq binding, sRNA may be stabilized and protected from ribonuclease cleavage, often carried out by ribonuclease E (RNase E). **d** | Hfq binding may lead to degradation of both sRNA and the target mRNA. **e** | Upon mRNA binding Hfq may stimulate polyadenylation by poly(A) polymerase (PAP), leading to a 3'-5' degradation by an exoribonuclease (Exo).

1.3 Experimental strategies for identifying RNA targets for Hfq

A method for identifying RNA targets for Hfq is to pull the bound complex out of a cell lysate, extract RNAs and identify the transcripts. There are several strategies for doing this and some of the most frequently used methods are illustrated in Fig. 3.

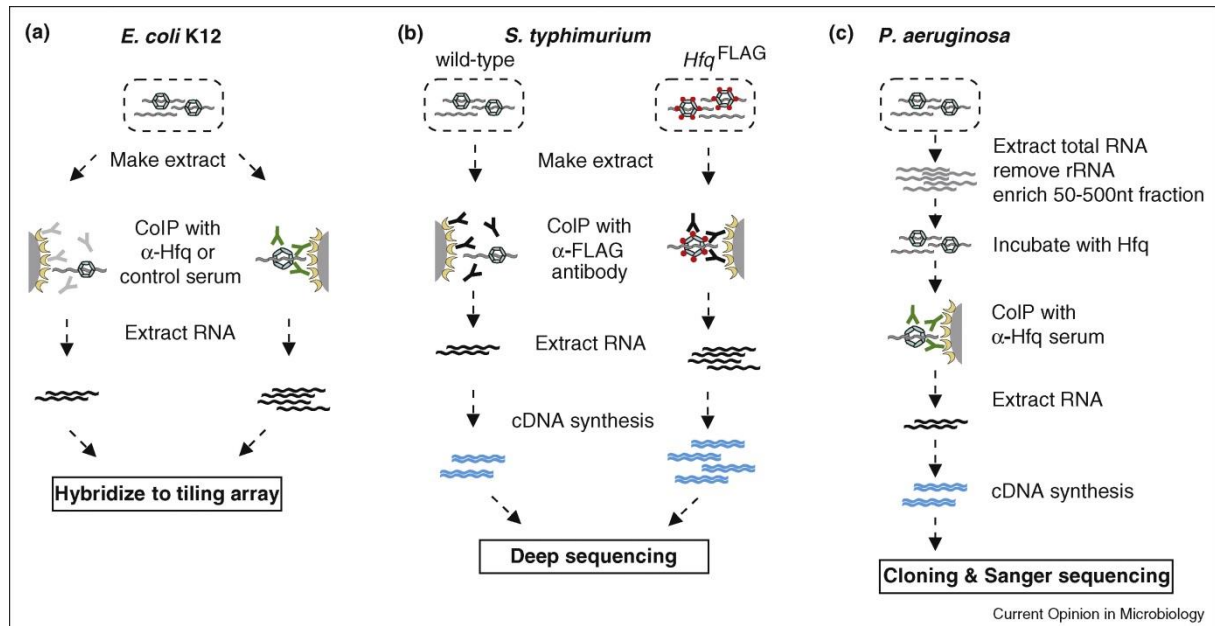


Figure 3: Experimental strategies for the identification of Hfq-associated sRNAs and mRNAs (from Chao and Vogel 2010). **a** | Hfq from *E. coli* K12 cell lysate is co-immunoprecipitated with polyclonal anti-Hfq serum; pre-immune serum is used in the control sample. Hfq-bound RNA is extracted and analysed by hybridization onto high-density oligonucleotide tiling arrays. **b** | The Hfq gene of *S. typhimurium* is genetically modified by a FLAG-epitope tag. A commercially available monoclonal anti-FLAG antibody is used to co-immunoprecipitate the tagged Hfq protein. An untagged wild type strain is used in a control reaction. Hfq-bound RNA is extracted and converted into cDNA and analysed by 454 pyrosequencing. **c** | Total RNA is extracted from *P. aeruginosa*. The extracted RNA can be size-fractionated by running it on a polyacrylamide gel. Hfq from *P. aeruginosa* is cloned into competent cells and expressed. The size-fractionated RNA is then incubated with purified Hfq protein (from the competent cells) *in vitro*, and a co-immunoprecipitation with anti-Hfq antibodies follows. Hfq-bound RNA is extracted and converted into cDNA before being cloned and sequenced.

Co-immunoprecipitation (co-IP) is a common method for identifying targets for Hfq. The goal is to bind and precipitate Hfq-bound complexes. This can be performed by using Hfq specific antibodies. To accomplish this cell lysate is mixed with antibodies against the protein of interest and the mix is left to precipitate. Specialized beads such as Sepharose A can help the precipitation of an antibody. These beads bind strongly to the Fc region of the immunoglobulin G (IgG) antibodies and can be pulled out by centrifugation (Fig. 4). This allows isolation of Hfq from cell extracts. Hfq-specific antibodies are only available for a few

bacterial species, and must in most cases be developed prior to co-IP. This time-consuming step is avoided when using an affinity or epitope bound Hfq. In this case Hfq is cloned into a plasmid where a specific tag or epitope is added, for example his-tag or FLAG-tag. The tagged protein is then expressed in *E. coli*, purified and incubated with a lysate of the original cells. From this mix, sRNA- and mRNA-bound Hfq can be pulled out with help of the specific tag, and the RNA can be isolated. A FLAG-tagged protein can be pulled out of a solution using a specific commercially available antibody. If a histidine tag has been attached to the protein, the complexes can be pulled out using a nickel column. This method requires that the recombinant protein is soluble as it is to be mixed with a cell lysate. The co-IP method may favour sRNAs that are stable. It is also important to be aware that some proteins are expressed under particular growth conditions and thus may go undetected in these kinds of experiments. One should include cells grown under various conditions to obtain the best possible result.

Identification of the Hfq-bound mRNAs and sRNAs is done by sequencing of cDNA. To create the template material for sequencing, the RNAs must first be reversely transcribed. There are several strategies for doing reverse transcription; one can use random hexaprimers, specific primers for known mRNAs, or the RNA can first be modified by addition a poly(A) tail before using oligo(dT) primers for first strand cDNA synthesis. In eukaryotes, mRNAs contain a poly(A) tail *in vivo*. This enhances their stability and translation efficiency. Poly(A) tails are not initially added to the mRNAs in prokaryotes, and will *in vivo* lead to degradation of the mRNA. Poly(A) tails can be added to prokaryotic mRNAs *in vitro* by poly(A) polymerase to enhance stability and translation efficiency. The advantage of this method compared to cDNA amplification using premade specific primers is that it allows cDNA amplification of genes that are not previously known and characterized. Another possibility for making cDNA is to use oligonucleotide hexaprimers made up of six random bases, making it possible to anneal to random sites on mRNA and be the basis for cDNA synthesis. In this way all transcripts can be reversely transcribed, including transcripts that have been partially degraded.

There are different methods for analysing cDNA; High throughput sequencing (HTS) and microarray are some of the most frequently used alternatives. Microarray is a common method for transcriptional profiling. Using this method for RNAs extracted from a wild type

and an *hfq* mutant allows examination of which genes are up and down regulated. This method requires a species specific microarray, meaning that the genome of the organism that is being studied must be sequenced and annotated. If the microarray has not already been established it is a time-consuming and expensive step.

HTS technologies generate a large amount of sequence data at a rapid pace and low cost compared to earlier methods. The efficiency of HTS is achieved by massive parallelisation of the sequencing process which is implemented differently on different technology platforms. HTS is likely the optimal method for identifying mRNAs and sRNAs today (Faner and Feig 2013). *Hfq* mutants have been made to assess the phenotypic response of bacteria to Hfq deficiency. Deletion strains can be made by transposon mutagenesis or by complete *hfq* deletion.

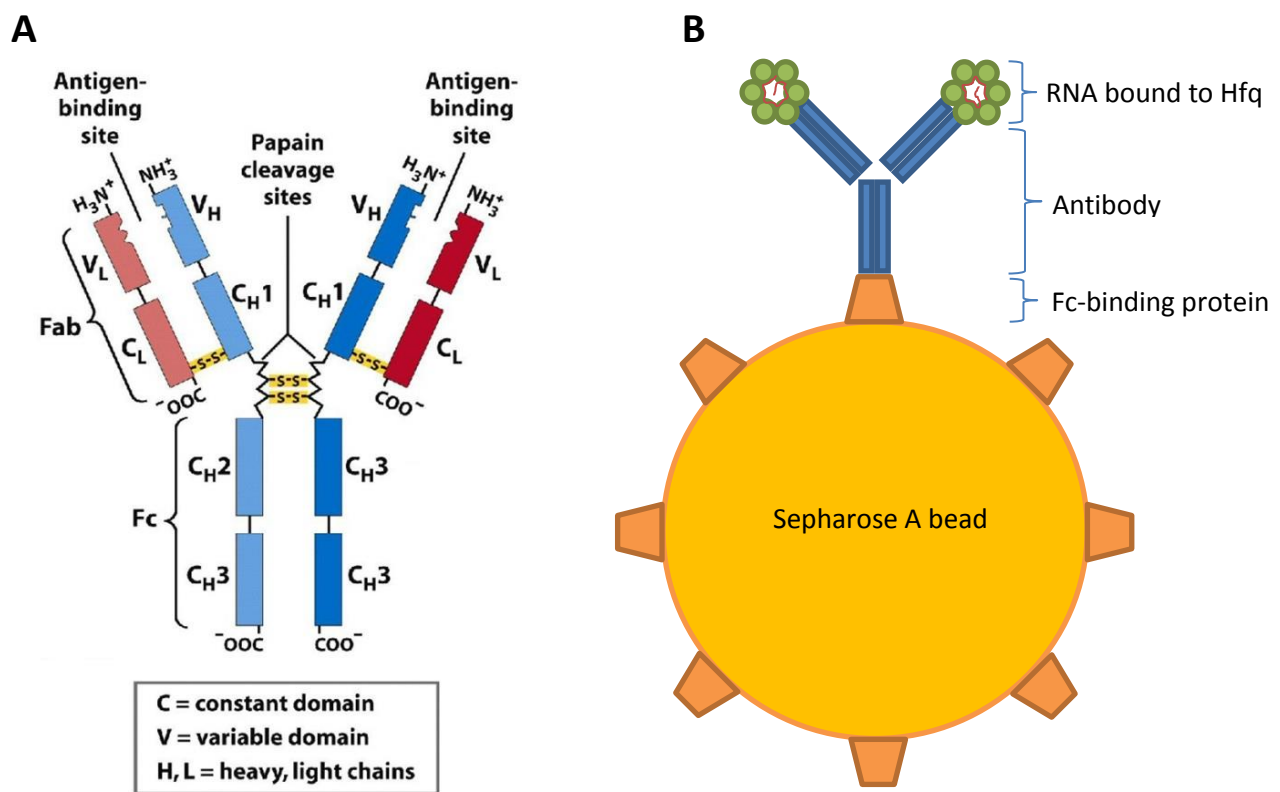


Figure 4A: IgG antibody (Nelson, Cox et al. 2008). Antigens bind to the Fab regions of the antibody and Sepharose A beads binds to the Fc region. **B:** Model showing the expected precipitate after co-IP with antiserum (incorrect aspect ratio). Sepharose A beads are coated with the Fc-binding Protein A from *Staphylococcus aureus* which allows binding of IgG antibodies. The anti-Hfq antibody will bind the Fc-binding protein of Sepharose A in the Fc-region, and Hfq in the Fab-region. Hfq-bound RNA can then be isolated from the precipitate.

1.4 Thermotogae

The Thermotogae phylum represents a deeply branching group of Bacteria consisting of only one class, one order and one family; Class Thermotogae, order Thermotogales, and family *Thermotogaceae*. The *Thermotogaceae* family currently consists of 11 genera; *Defluviitoga* (Hania, Godbane et al. 2012), *Fervidobacterium* (Patel, Morgan et al. 1985), *Geotoga* (Davey, Wood et al. 1993), *Kosmotoga* (Dipippo, Nesbø et al. 2009), *Marinitoga* (Wery, Lesongeur et al. 2001), *Mesoaciditoga* (Reysenbach, Liu et al. 2013), *Mesotoga* (Nesbø, Bradnan et al. 2012), *Oceanotoga* (Jayasinghearachchi and Lal 2011), *Petrotoga* (Davey, Wood et al. 1993), *Thermosipho* (Huber, Woese et al. 1989) and *Thermotoga* (Huber, Langworthy et al. 1986).

In a paper by Bhandari and Gupta (2013) a division into three orders was recently proposed; order *Thermotogales*, *Kosmotogales* ord. nov. and *Petrotogales* ord. nov., containing four families; *Thermotogaceae*, *Fervidobacteriaceae* fam. nov., *Kosmotogaceae* fam. nov. And *Petrotogaceae* fam. nov., and a new genus; *Pseudothermotoga* gen. nov. (Bhandari and Gupta 2013). The proposal is based on comparative genomic studies, particularly focused on conserved signature indels (CSIs). CSIs are insertions or deletions of amino acids within conserved regions. This is considered a rare evolutionary event as these insertions or deletions are likely to interfere with the function of the protein. If the protein is not translated correctly it might hinder or limit the growth or survival of the bacteria, hence these CSIs are unlikely to be transferred to a new generation. As these CSIs are rare evolutionary events, shared CSIs are a strong indication of common descent.

The first discovered Thermotogae bacteria was *Thermotoga maritima*, characterized in 1986 (Huber, Langworthy et al. 1986). *T. maritima* was first isolated from geothermally heated marine sediments in Italy and the Azores. Common for the members of phylum Thermotogae is that they are gram-negative, non-sporulating, rod-shaped bacteria enclosed by a characteristic sheath-like envelope called a toga. They are mostly obligate anaerobic chemoorganoheterotrophs with a fermentative metabolism. Members of Thermotogae have been found in marine hydrothermal vents and fresh water hot springs, usually at moderate pH and salinity. They have also been isolated from the deep subsurface, including oil wells. They use sulphur as an electron acceptor, generating hydrogen sulphide as a by-product of metabolism. As the name suggests, all Thermotogae were known to be extreme or

moderate thermophiles, until the recent discovery of mesophilic Thermotogae in mesophilic environments (Nesbø, Kumaraswamy et al. 2010).

16S rRNA gene sequencing has shown that the Thermotogae constitute an early branching lineage in the bacterial phylogenetic tree (Fig. 5). The bacteria have been subject to an extensive lateral gene transfer. This was first detected in *T. maritima*, where 24% of the genes were most similar to archaeal genes (Nelson, Clayton et al. 1999). Lateral gene transfer has also been detected in *Thermosipho africanus*, where 26% of the genes were most similar to Firmicutes, and 13% to Archaea (Nesbø, Bapteste et al. 2009).

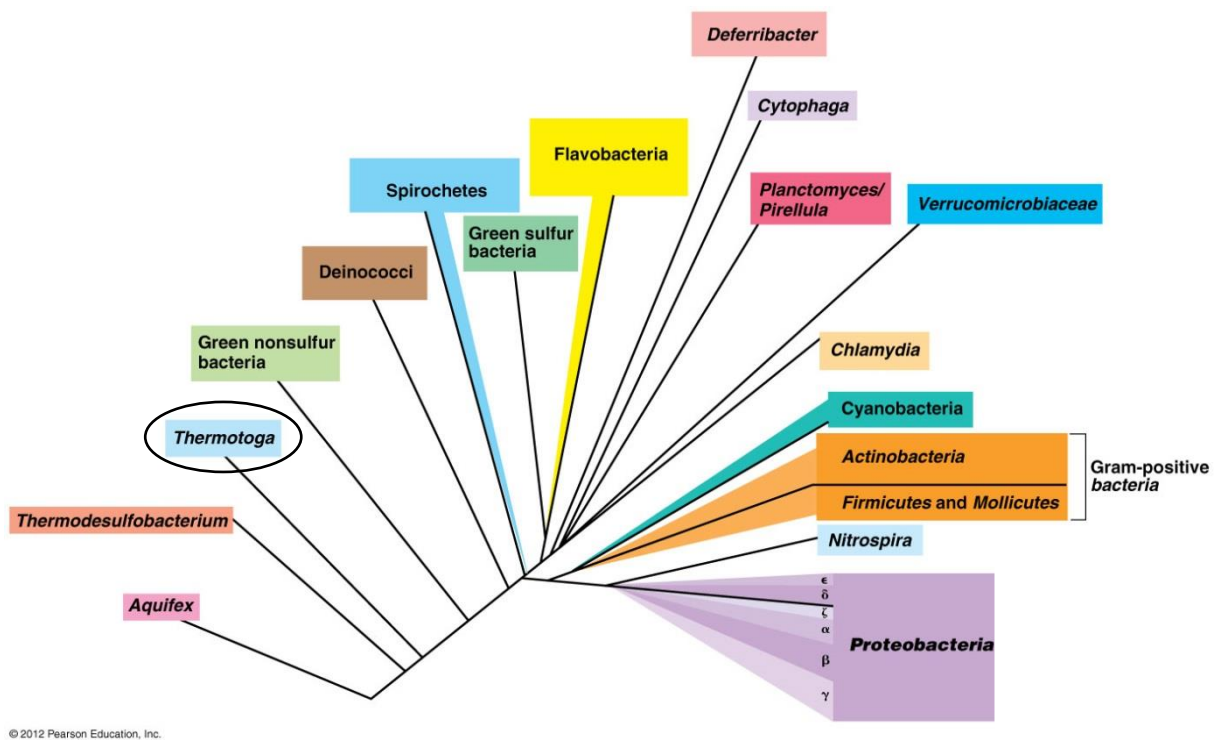


Figure 5: Schematic tree of the domain Bacteria, based on 16S rRNA (modified from Madigan, Martinko et al. 2012). Thermotoga (encircled) is deeply branched in the phylogenetic tree.

Three different bacterial strains, representing three different Thermotogae genera were assessed in this study, all of which have been isolated from North Sea oil well production waters. Several species were investigated to allow comparisons of potentially different functions of the Hfq protein. The evolutionary relationship of the Thermotogales species is shown in Fig. 6. The species that were studied in this thesis are encircled.

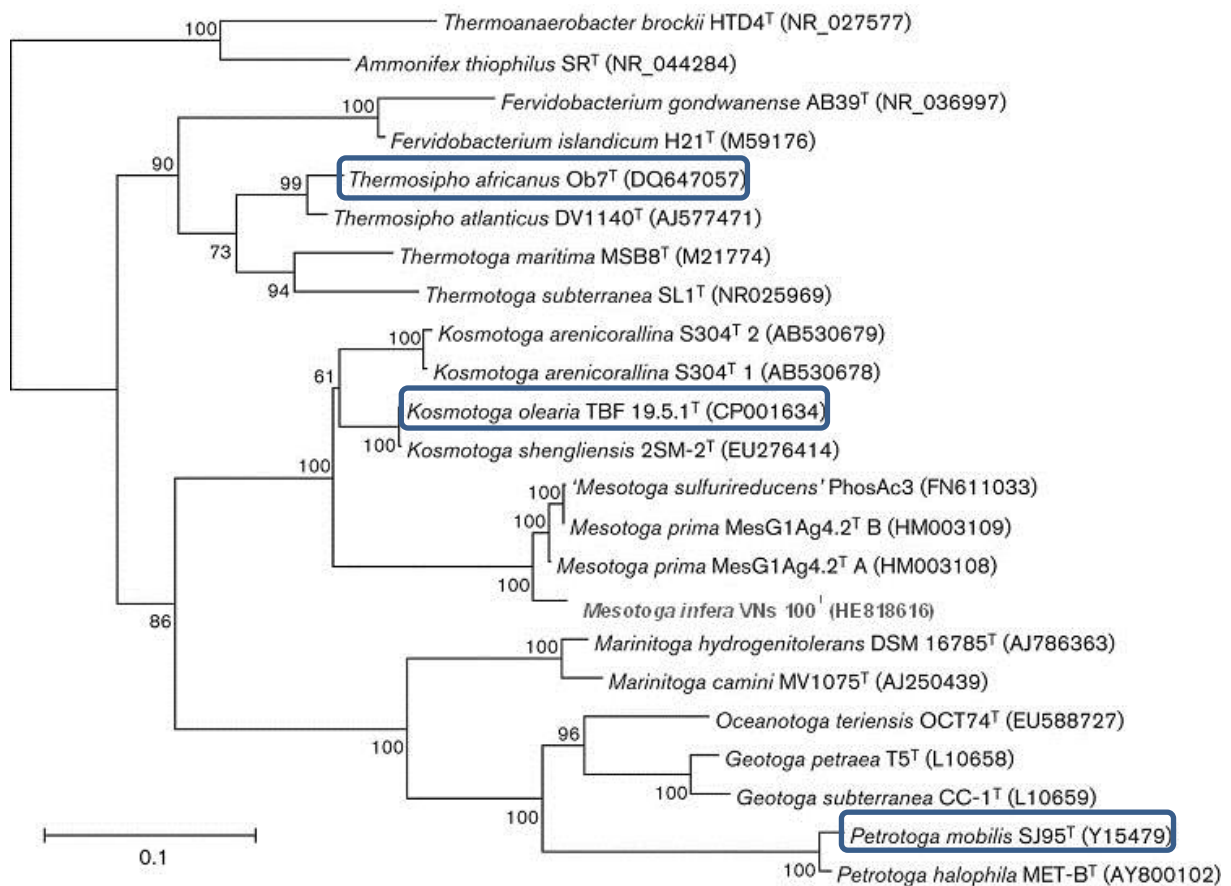


Figure 6: Maximum-likelihood phylogenetic tree based on 16S rRNA showing the relatedness between the *Thermotogales* (modified from Ben Hania, Postec et al. 2013). The bacteria used in this study are encircled. *Thermoanaerobacter Brockii* and *Ammonifex thiophilus* were used as outgroups. The numbers at each node show the bootstrap value, which gives the percentage of this particular branching in a phylogenetic tree when repeating the process 500 times. Bootstrap values below 50% have been removed. The bar indicates the number of substitutions per site.

1.4.1 *T. africanus*

T. africanus was first characterized by Hubert et al. in 1989 (Huber, Woese et al. 1989). The species name is derived from the place they were first isolated, in a marine hydrothermal area at Obock in Djibouti, Africa (isolate Ob7). *T. africanus* was the first species to be characterized belonging to the genus *Thermosipho*. The genus name reflects the physiological properties of the bacteria; it is a thermophile with a surrounding sheet-like structure often referred to as a toga (*thermê* meaning heat, *sipho* meaning little pipe to suck drinks through, a tube). *T. africanus* can form short chains of up to twelve individual cells surrounded by one toga (Ravot, Olliver et al. 1996). Fig. 7 is an electron micrograph of four *T. africanus* cells surrounded by a toga. *T. africanus* grows optimally at 75°C and pH 7.2. These

bacteria can grow in environments with temperatures ranging from 35-77°C, pH 6.0 to 8.0 and 0.11-3.6% NaCl (Huber, Woese et al. 1989).

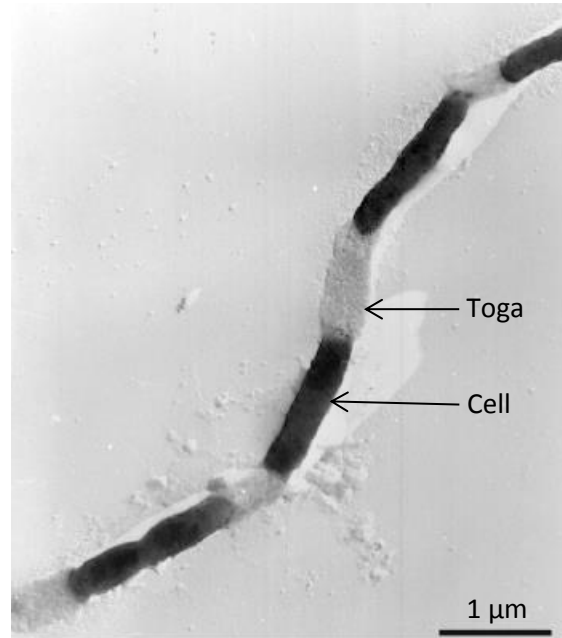


Figure 7: Electron micrograph of *T. africanus* (modified from Huber and Hannig 2006). The toga (in grey) surrounds each of the four rod-shaped cells (in black). Bar; 1 µm.

1.4.2 *K. olearia*

K. olearia was first characterized by Dipippo et al. (Dipippo, Nesbø et al. 2009). This strain (isolate TBF 19.5.1(T)) was isolated from oil production fluid originating from 1560 meters below the sea floor, at the Troll B platform in the North Sea. *K. olearia* was the first species to be characterized belonging to the genus *Kosmotoga*. The name is derived from *kosmos* meaning universe/world and *toga* referring to the sheet-like outer structure. Bacteria belonging to the *Kosmotoga* genus are non-motile and rod-shaped. *K. olearia* strain TBF 19.5.1(T) grows optimally at 65°C, pH 6.8 and 2.5-3% NaCl. These bacteria can grow in environments with temperatures ranging from 20-80°C, pH of 5.5 to 8.0 and 1-6% NaCl (Dipippo, Nesbø et al. 2009).

1.4.3 *P. mobilis*

The genus *Petrotoga* was first characterized by Davey et al. in 1993 (Davey, Wood et al. 1993). The name is derived from *petra* meaning rock/stone and *toga* again referring to the sheet-like outer structure. *P. mobilis* was first characterized by Lien et al. in 1998 (Lien, Madsen et al. 1998). This strain (isolate SJ95), was isolated from oil production water retrieved from water separator tanks on an off-shore oil platform in the North Sea. *Mobilis* means movable/motile and refer to the species ability to move by a subpolar flagellum. The cells can grow singularly or in chains of over 20 cells surrounded by a single toga. *P. mobilis* strain SJ95 grows optimally at 58-60°C, pH of 6.5-7.0 and 3-4% NaCl. These bacteria can grow in environments with temperatures ranging from 40-65°C, pH of 5.5-8.5 and 0.5-9.0% NaCl (Lien, Madsen et al. 1998).

1.5 Aim of the study

Earlier studies have demonstrated an important role of Hfq proteins in fitness and virulence of several bacterial strains belonging to Proteobacteria and gram-positive bacteria. The role of this highly conserved protein has not yet been investigated in phylogenetically deeply branched or thermophilic bacteria. The aim of this study was to investigate the role of Hfq proteins in gene expression in the deeply branching Thermotogales order, including the following sub-goals:

- 1- Cloning and overexpression of histidine-tagged Hfq from three different Thermotogae species.
- 2- Purification of recombinant tagged Hfq using affinity columns or production of anti-Hfq antibodies in rabbits.
- 3- Extraction and purification of Hfq proteins in complex with RNAs from cell lysates of the three strains.
- 4- Synthesis of cDNA from RNA co-purified with Hfq.
- 5- Identification of target genes that are regulated by Hfq and their interacting sRNAs by sequencing of the cDNAs.

The plan for the laboratory work is shown in Fig. 8.

2.1 Bacterial strains and cultivation

Species belonging to three different genera of the order Thermotogales were used in this study; *K. olearia* (DSM no.21960, type strain), *P. mobilis* (DSM no. 10674, type strain) and *T. africanus* (DSM no. 13782), which were purchased from DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH). Their phylogenetic relationship is shown in Fig. 6. All three strains have been isolated from North Sea oil well production waters (Table 1). The bacteria were cultured in mineral media as recommended by DSMZ or in Medium 2 (Table 2). The medium was prepared in two or three litre Erlenmeyer flasks that also served as dispenser. All the ingredients listed in Table 2 except bicarbonate and Na₂S were mixed and dissolved prior to autoclaving (Getinge VS 90) using a magnetic stirrer. The medium was dispensed into serum flasks using Hungate technique (Fig. 9). The dispenser flask was corked with a butyl rubber cork customized for anaerobic tapping. During cooling on ice, the medium was kept anoxic by flushing the flask with nitrogen gas while stirring with a magnetic stirrer. When the medium had cooled to approximately 40°C, bicarbonate, Na₂S and vitamin solution (Table 3) were added and the pH was adjusted to 7 with 6M HCl.

Table 1: An overview of the three strains of Thermotogales that were used in this study.

Name	DSM number	Optimal growth temperature (temperature range)	Reference
<i>Kosmotoga olearia</i>	21960 (type strain)	65°C (20-80)	(Dipippo, Nesbø et al. 2009)
<i>Petrotoga mobilis</i>	10674 (type strain)	58-60°C (40-65)	(Lien, Madsen et al. 1998)
<i>Thermosipho africanus</i>	13782	75°C (35-77)	(Nesbø, Bapteste et al. 2009)

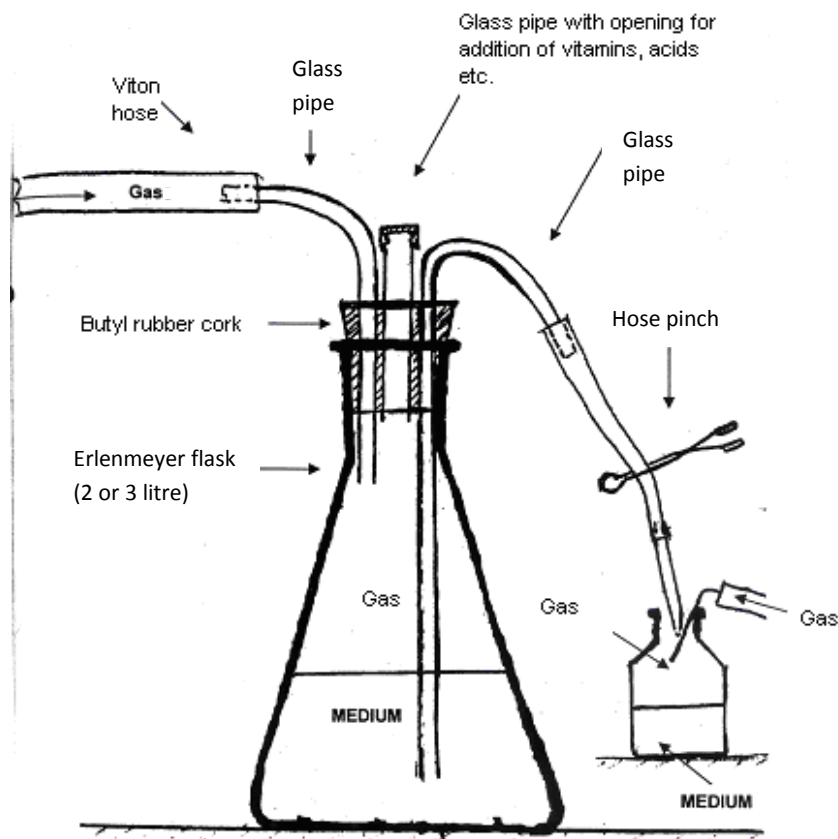


Figure 9: Setup for anaerobic dispensing of medium into serum flasks using Hungate technique (Modified from N. K. Birkeland, personal communication).

Approximately 50 ml aliquots were transferred to sterile 120 ml serum bottles. The bottles were corked with butyl rubber stoppers and capped with aluminium crimp seals using Hungate technique (Hungate 1950). Syringes were used to add substrates listed below (Table 5). The substrates were dissolved in anoxic water which was made by boiling Milli-Q water for approximately 20 minutes while flushing with nitrogen gas through a silicon tube that was placed into the solution. The bottles were corked as described above. A small amount of a dithionite solution was used for reducing the medium if the resazurin indicated oxic conditions. A 0.6M dithionite solution was prepared with anoxic water and sterile filtrated (0.2 μm) into serum bottles that had been flushed with nitrogen gas. Typically only one drop of dithionite was needed to make the medium anoxic.

The serum flasks were inoculated with one ml of an active culture from DSMZ or from stored cultures. *P. mobilis* was incubated at 55°C, *K. olearia* at 65°C and *T. africanus* at 70°C. The cultures were incubated from three days to one week before harvesting.

Table 2: Composition of the basal medium for cultivation of the different strains of Thermatogales.

Component	Medium for <i>Petrotoga</i> and <i>Thermosipho</i> (DSMZ no. 718)*	Medium for <i>Kosmotoga</i> (no. DSMZ 1163)*	Medium 2**
KCl	0.335 g	0.330 g	0.33 g
MgCl ₂ x 2 H ₂ O		0.900 g	
MgCl ₂ x 6 H ₂ O	4.000 g		0.90 g
MgSO ₄ x 7 H ₂ O	3.450 g	1.400 g	1.40 g
NH ₄ Cl	0.250 g	0.250 g	0.25 g
CaCl ₂ x 2 H ₂ O	0.140 g	0.140 g	0.14 g
K ₂ HPO ₄	0.140 g		
KH ₂ PO ₄		0.450 g	0.45 g
NaCl	18.000 g	30.000 g	20.00 g
Trace elements (Table 4)	10.000 ml	10.00 ml	1.0 ml
Fe(NH ₄) ₂ (SO ₄) ₂ x 6 H ₂ O	2.000 g		
Resazurin	1.000 g	0.5 mg	0.5 ml (0.02%)
NaHCO ₃	1.000 g		
Na ₂ S x 9 H ₂ O	0.500 g		
Na ₂ S (0.5M)			4 ml
1,4-Piperazinediethanesulfonic acid (PIPES (SIGMA))		6.700 g	
Vitamin solution (Table 3)	10.00 ml	10.00 ml	
Cystein-HCl x H ₂ O		0.500 g	
Milli-Q H ₂ O	1000 ml	1000 ml	1000 ml

* as recommended by DSMZ (www.dsmz.de).

**modified from Widdel, Kohring et al. (1983).

Table 3: Composition of the vitamin solution (modified from Widdel and Pfenning 1981).

Component	Amount
4-Aminobenzoic acid	8.0 mg
D(+)-biotin	2.0 mg
Nicotinic acid	20.0 mg
Ca-D(+) pantothenate	10.0 mg
Pyridoxamine x 2HCl	30.0 mg
Thiamine (Cl) ₂	20.0 mg
Milli-Q H ₂ O	1000 ml

Table 4: Composition of the trace element solution. Procedure and recipe were obtained from DSMZ (www.dsmz.de).

Component	Amount
Nitrilotriacetic acid	1.50 g
MgSO ₄ x 7H ₂ O	3.00 g
MnSO ₄ x H ₂ O	0.50 g
NaCl	1.00 g
FeSO ₄ x 7H ₂ O	0.10 g
CoCl ₂ x 6H ₂ O	0.18 g
CaCl ₂ x 2H ₂ O	0.10 g
ZnSO ₄ x 7H ₂ O	0.18 g
CuSO ₄ x 5H ₂ O	0.01 g
KAl(SO ₄) ₂ x 12H ₂ O	0.02 g
H ₃ BO ₃	0.01 g
Na ₂ MoO ₄ x 2H ₂ O	0.01 g
NiCl ₂ x 6H ₂ O	0.03 g
Milli-Q H ₂ O	1000 ml

Table 5: The substrates added to medium in Table 2 for the different strains.

Component	<i>T. africanus</i>	<i>P. mobilis</i>	<i>K. olearia</i>
10% yeast extract	0.5 ml	100 µl	0.5 ml
10% trypton	0.5 ml		
10 % maltose			0.5 ml
50% glucose	0.5 ml	0.5 ml	
1M Na+Thio	1 ml*		

**Sodium thionite was not used for Medium 2.

2.2 Extraction and quantification of DNA

Genomic DNA was extracted from the cultivated species to be used as a template for the amplification of the *hfq* gene.

2.2.1 DNA extraction by the cetyltrimethylammonium bromide (CTAB) method

Genomic DNA was extracted from the three bacterial species using a modified CTAB protocol (Ausubel, Brent et al. 1988). The modified method is started by spinning down three ml of culture at 13,000 rpm in Biofuge 13 (Heraeus Sepatech) for two minutes, and resuspension of the pellets in 567 μ l of 1X TE-buffer and one μ l of RNaseA (20 mg/ml). The samples were incubated in a water bath at 65°C for 60 minutes. Following incubation, 30 μ l of 10% sodium dodecyl sulphate (SDS) and three μ l of 20 mg/ml proteinase K were added. After incubation for 60 minutes at 37°C, 0.5 μ l of 3M sodium acetate (pH 4.6) and 100 μ l of 5M sodium chloride were added. The samples were then added 80 μ l of CTAB/NaCl, and incubated for 20 minutes at 65°C. Thereafter, 785 μ l of chloroform/isoamyl alcohol (24:1) was added to the samples, and the samples were centrifuged at 13,000 rpm for five minutes. Two layers were visible in the sample tubes from which the lower layer contained proteins and cell debris and the upper layer contained bacterial DNA. The upper layer was transferred to an Eppendorf tube and the DNA was precipitated by adding 0.6X volume (420 μ l) of isopropanol. After centrifugation at 13,000 rpm for 15 minutes a pellet was visible in each tube. The pellet was washed with 70% ethanol, and the tubes were left in a fume hood with their lids off to dry off potential remnants of ethanol. Finally, 50 μ l of 1X TE-buffer was added and the samples were left at 4°C overnight for the pellets to dissolve in the buffer.

2.2.2 DNA quantification by NanoDrop

DNA was quantified by using a NanoDrop ND-1000 Spectrophotometer. Nucleic acids have a maximum absorption at 260 nm, while proteins typically have a maximum absorption at 280 nm. The NanoDrop measures the DNA concentration based on the absorption at 260 nm, and also gives the relationship of the absorption between 260 nm and 280 nm. A 260/280 ratio close to 1.8 reflects pure DNA. Deviations from this number indicate the presence of

contaminants, such as proteins and other organic compounds. Calibrations were done with Milli-Q water or buffer, depending on the DNA-suspension.

2.2.3 Agarose Gel Electrophoresis

Agarose gel electrophoresis was performed to confirm the presence of nucleic acids, and to estimate the size of the fragments. A 1% agarose gel was made in 1X TAE-buffer with a total volume of 20 ml. The agarose was dissolved by heating, and when the temperature had dropped to approximately 60°C, 0.5 µl of Gel Red was added to the gel mix. The solution was gently poured into a small plastic form, a well comb inserted, and the gel was left for solidification for approximately 30 minutes. The gel was placed in an electrophoresis unit (GE Healthcare) and 1X TAE-buffer was poured into the unit until the gel was covered completely. Current is carried by the ions in the TAE-buffer. The buffer also assures a stable pH during electrophoresis. Three µl DNA ladder (Appendix B, table B.2) was loaded into well number one to allow estimation of the length of the fragments in the samples. DNA samples were mixed 3:1 with 6X loading buffer (Appendix B, table B.3). The 6X loading buffer contains sucrose which allows the sample to settle at the bottom of the well, and two dyes (bromophenol blue and xylene cyanol FF) which allow monitoring of the electrophoresis throughout the run. The electrophoresis was run at 80V for one hour, and the results visualized under UV light by a transilluminator (Molecular Imager ChemiDoc xRST with Image lab software, Bio-Rad).

2.3 Gene amplification

Specific primers were designed to allow amplification of the *hfq* genes from the Thermotogales species. After amplification, the amplicons were purified using the GenElute PCR Clean-UP Kit from Sigma-Aldrich.

2.3.1 Primer design for polymerase chain reaction (PCR)

Specific primers were designed to amplify the DNA-sequence coding for the *hfq* genes (ordered from Sigma-Aldrich). The sequences coding the *hfq* genes were obtained from the

National Center for Biotechnology Information (NCBI) online database (www.ncbi.nlm.nih.gov). Two primers were designed for each strain to match the 5' and 3' ends of the genes. To enable subsequent cutting of the product using specific restriction enzymes the forward primers were extended with an *NdeI* restriction site (5'-CATATG-3'), and the reverse primers with a *HindIII* restriction site (5'AAGCTT-3'). It was ensured that the *hfq* gene itself did not carry any of these restriction sites by searching for restriction sites using Emboss Explorer (<http://emboss.bioinformatics.nl>). Calculating the annealing temperature is an important step of primer design. In the annealing step of the PCR reaction the primers attach to the single stranded DNA. The annealing temperature is usually set to approximately 5°C below the melting temperature (T_m). The T_m is the temperature where about half of the double stranded DNA dissociates and becomes single stranded. This is dependent on the length of the primer, and the content of guanine and cytosine. The T_m values were calculated separately for all the primers, and it should be between 55°C and 65°C. It is important that the T_m value for the two primers of one PCR reaction is similar as they are to be run under the same conditions in the PCR-machine. The T_m value can be adjusted by adding or removing bases. The T_m values were calculated using the online program Oligo Calc (<http://www.basic.northwestern.edu/biotools/oligocalc.html>). The *hfq* genes were amplified from the three bacterial strains by performing PCR with the specific primers shown in Table 8.

Table 8: Primers used for amplifying of the *hfq* genes. The *NdeI* restriction site (CATATG) was added to the forward primers and the *HindIII* restriction site (AAGCTT) to the reverse primers. The sequence ATAT was added to the 5' end of the restriction sites. GATC was added prior to the *HindIII* site in the reverse primer for *K. oleria*. This is a site for the restriction enzyme *BamHI* which was included as an alternative to *HindIII*. As there were no difficulties in cutting with *HindIII*, the *BamHI* site was not included when designing the reverse primers for *P. mobilis* and *T. africanus*.

Name	Forward primer 5'--3'	T_m , forward primer	Reverse primer 5'--3'	T_m , reverse primer
<i>K. oleria</i>	ATAT <u>CATATG</u> GCTGAAAAATTCA ATCTTCAGG	62.0°C	ATATGATCAAGCTTCCCCTCCTTC TTATTCTGGACAG	61.0°C
<i>P. mobilis</i>	ATAT <u>CATATG</u> GCAGAAAAGTTCA ATTTACAAGAC	62.1°C	ATATAAGCTTCTCATCTTTAGACG GTGGTTCTTG	63.6°C
<i>T. africanus</i>	ATAT <u>CATATG</u> GCAGAAAATTTA ATTTACAAGACAG	62.1°C	ATATAAGCTTCTTTTCTTCATTTTC AGATTCTTCGTT	60.8°C

2.3.2 PCR

PCR was performed in 50 µl total reaction mixtures using 1 mM dNTPs, 0.5 mM of each primer, 1 Unit One Taq DNA polymerase and 17 to 540 ng template DNA. One Taq standard reaction buffer was used. The PCR reaction was performed in a DNA Engine Peltier Thermal Cycler (Bio-Rad) under the following conditions:

94°C	30 sec	}	30 cycles
94°C	30 sec		
60°C	30 sec		
68°C	2 min		
68°C	10 min		
4°C	∞		

The PCR products were analysed by agarose gel electrophoresis and quantified by NanoDrop. The bands of the PCR products were compared to the expected length of the *hfq* genes, which were retrieved from NCBI.

2.3.3 Purification of the PCR product

The GenElute PCR Clean-UP Kit from Sigma-Aldrich was used to purify PCR product. The instructions from the kit manufacturer were followed. A GenElute plasmid mini spin column was inserted in a collection tube and added 0.5 ml of Column Preparation Solution before being centrifuged at 13,000 rpm for one minute. The eluate was discarded. 470 µl of Binding Solution was added to 94 µl of PCR product and mixed by pipetting the solution up and down. The solution was transferred to the binding column, and the tubes were centrifuged at 13,000 rpm for one minute before the eluate was discarded. The binding column was added 0.5 ml of Wash Solution and the tubes were centrifuged at 13,000 rpm for one minute. The eluate was discarded and the tubes were centrifuged for two minutes to remove any excess ethanol from the Wash Solution. The binding column was transferred to a fresh collection tube and 50 µl of sterile water was added. The tubes were left at room temperature for one minute before they were centrifuged at 13,000 rpm for one minute to elute the DNA. Total DNA in the purified PCR-product was quantified by NanoDrop.

2.4 Cloning

2.4.1 pET-21b(+) vector

The pET-21b(+) vector (Novagen) was used for ligation and transformation (Appendix C, table C.5). Insertion of the gene was done between the *HindIII* and *NdeI* site. This resulted in a fusion gene with six histidine residues added to the C-terminus of the recombinant proteins. The pET-21b(+) vector contains a sequence coding for ampicillin resistance which was later used as a selective marker. In addition to the selectable marker, the vector contains essential features, such as an origin of replication (ORI) and a multicloning site (MCS). Isopropylthio- β -galactoside (IPTG) was used to induce expression of the cloned gene.

2.4.2 Digestion of PCR-product and vector

The purified PCR products were digested with restriction enzymes *NdeI* and *HindIII* (New England Biolabs). For the *hfq* genes, digestion mix was performed in 30 μ l total reaction mixtures using 0.5-0.9 μ g PCR product and 20 Units of each restriction enzyme. The vector was digested in 20 μ l total reaction mixture using 0.1 μ g PCR product and 10 Units of each restriction enzyme. NEBuffer was used as recommended by the manufacturer. The samples were incubated in a water bath at 37°C for three hours, then kept overnight at 4°C. Following the overnight incubation the tubes were heated in a water bath holding 60°C for two minutes. The digested products were purified as described in section 2.3.3, but this time 150 μ l of Binding Solution was added to 30 μ l of cut PCR product, and 100 μ l of Binding Solution was added to 20 μ l of digested vector prior to loading onto the column. The DNA was analysed by agarose gel electrophoresis and total DNA was quantified by NanoDrop.

2.4.3 Ligation

The ligation of the digested PCR-product into the digested vector was performed in 25 μ l total reaction mixture with 10X ligation buffer using 5-7 ng PCR-product, 36 ng vector and 400 Units T4 DNA ligase. A control was included in which no insert was added. The ligation mix was left at room temperature overnight.

2.4.4 Transformation for sequencing

By transforming the plasmids into competent cells, sufficient copies could be made for sequencing and confirmation of the plasmid insert. After ligation the plasmid was transformed into One Shot TOP10 Chemically Competent Cells (Invitrogen). These cells are modified *E. coli*, which have been chemically manipulated so that they are capable of transformation. The instructions from the kit manufacturer were followed.

After thawing the competent cells on ice, five μl of each ligation mixture (section 2.4.3) was added and mixed gently with the cells. After incubation on ice for 30 minutes the cells were heat shocked for 30 seconds by placing the vials in a water bath at 42°C. The vials were then left on ice for two minutes before 250 μl of room tempered Super Optimal broth with Catabolic repression (SOC) medium was added. The vials were incubated at 37°C for 60 minutes before being plated onto Petri dishes containing Luria Bertani agar with ampicillin (LB+Amp₁₀₀) (Appendix B, B.7). The transformation mix was plated using four different volumes: 1 μl , 10 μl , 100 μl and 139 μl . The plates were incubated at 37°C overnight.

The successful transformation was detected with the help of a selectable marker. This selective feature is based on a gene in the pET-21b(+) vector coding for ampicillin resistance. Only cells that contain the plasmid survive the antibiotic treatment, and will grow on plates with ampicillin. Four colonies were picked from each clone and plated onto fresh LB+Amp₁₀₀ plates before being incubated at 37°C overnight. The following day single colonies were transferred to separate sterile culture tubes containing five ml LB+Amp₁₀₀ and incubated at 37°C overnight for extraction of plasmids.

2.4.5 Plasmid clone purification

Plasmid DNA from recombinant *E. coli* was isolated by using GenElute Plasmid Miniprep Clean-UP Kit (Sigma-Aldrich). The instructions from the kit manufacturer were followed. 1.5 ml of overnight culture of recombinant *E. coli* was centrifuged at 13,000 rpm for one minute. The pellet was resuspended in 200 μl of Resuspension Solution and mixed by pipetting the solution up and down. The cells were lysed by addition of 200 μl of Lysis Solution, gently inversion eight times and incubation on the bench for four minutes. This treatment made the solution clear and viscous. The cell debris was precipitated by adding 350 μl of Binding

Solution and mixed by gently inverting the tube six times. Cell debris was pelleted by centrifugation at 13,000 rpm for 10 minutes. GenElute Miniprep Binding Columns were inserted into collection tubes and 500 µl of Column Preparation Solution were added. The tubes were centrifuged at 13,000 rpm for one minute and the liquid was discarded. The lysate was then transferred to the prepared Binding Column and centrifuged at 13,000 rpm for one minute and the liquid was discarded. The column was then washed by adding 750 µl of Wash Solution 2 and centrifuged at 13,000 rpm for one minute. The liquid was discarded and the centrifugation step was repeated, but this time for two minutes. The Binding Column was transferred to a fresh tube and added 100 µl of sterile water. The tubes were left at room temperature for one minute before they were centrifuged at 13,000 rpm for one minute to elute the DNA. The purified plasmids were analysed by agarose gel electrophoresis, using NEB supercoiled ladder. Total DNA was quantified by NanoDrop.

2.4.6 Confirmation of plasmid inserts

2.4.6.1 Amplification of the *hfq* gene

To amplify the *hfq* gene, a PCR reaction was performed as described in section 2.3.2 using 5-6 ng plasmid template. The amplicons were analysed by agarose gel electrophoresis, and total DNA was quantified by NanoDrop.

2.4.6.2 Estimation of insert length

The length of the insert was estimated by digesting with restriction enzymes *NdeI* and *HindIII* and analysis by agarose gel electrophoresis. Digestion was performed in 10 µl total reaction mixtures with one µl 10X buffer, 9-12 ng plasmid DNA and 16 Units *HindIII* and *NdeI*. The digestion mix was incubated at 37°C for 60 minutes and successful digestion was verified by agarose gel electrophoresis on a 0.8% agarose gel, run at 65V for 70 minutes.

2.4.6.3 Big Dye sequencing

The correct insertion of the *hfq* genes was confirmed by Sanger sequencing using a Big Dye reaction kit (Life Technologies). In this reaction dideoxynucleotides (ddNTPs) are added, these are deoxynucleotides (dNTPs) that lacks an –OH group. Due to the missing –OH group elongation will terminate when a ddNTP is added to a strand. Each ddNTP also has a specific fluorescent molecule attached to its end. In this way each synthesized strand is terminated by a base which can be recognized by its fluorescent molecule. As both dNTPs and ddNTPs are added in the reaction the elongation is terminated at random sites of each strand. When the reaction is run for a certain amount of time strands of all possible lengths is synthesized, each with a fluorescent molecule at its end. The strands are separated on a capillary gel electrophoresis based on its lengths. The end base of each strand can then be recognized by their specific dye by a laser, and the result is shown in a chromatogram. The reaction was performed in 10 µl total reaction mixtures, using 520-570 ng PCR-product, 0.32 µM *hfq* primers (forward or reverse), one µl Big Dye and one µl of Sequencing Buffer. The sequencing reaction was performed in a DNA Engine Peltier Thermal Cycler (Bio-Rad) under the following conditions:

96°C	5 min		
96°C	10 sec	}	30 cycles
50°C	5 sec		
60°C	4 min		
60°C	10 min		
4°C	∞		

Following amplification, 10 µl of Milli-Q water was added to each tube, making a total volume of 20 µl. The sequencing reactions were analysed by the Sequencing Facility at the University of Bergen (www.seqlab.uib.no) where automated Sanger DNA sequencing is performed using a capillary-based Applied Biosystem 3730XL Analyzer. The sequences were later compared to the sequence of the *hfq* gene which was obtained by using the search engine Basic Local Alignment Search Tool (BLAST) from the NCBI online database (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

2.5 Overexpression of cloned genes

Plasmids with the *hfq* gene were transformed into One Shot BL21 (DE3) plysS ultracompetent cells (Invitrogen) for expression. This was done to produce recombinant His-tagged Hfq proteins and to check the protein solubility. A soluble protein could be pulled out of clone lysates using the attached His-tag.

2.5.1 Transformation for gene expression

Confirmed plasmids with the *hfq* gene were transformed into One Shot BL21 (DE3) plysS ultracompetent cells (Invitrogen), following the instructions from the kit manufacturer. Two μl of the ligation mix (from section 2.4.3) were added to a vial of competent cells, carefully mixed and incubated on ice for 30 minutes. The cells were then heat shocked at 42°C for 30 seconds, and put back on ice. Each vial was added 250 μl of room tempered SOC medium and incubated at 37°C for 60 minutes with shaking (185 rpm). Finally, volumes of 1, 10 and 100 μl of the transformation mix were plated onto separate agar plates with LB containing Ampicillin and Chloramphenicol (LB+Amp₁₀₀+Cam₃₄) (Appendix B, table B.8) and incubated at 37°C overnight. Singles colonies were transferred to fresh LB+Amp₁₀₀+Cam₃₄ agar plates, and incubated at 37°C overnight. New single colonies were then transferred to Erlenmeyer flasks containing 10 ml LB+Amp₁₀₀+Cam₃₄ and incubated at 37°C with shaking (180 rpm) overnight.

2.5.2 Induction of gene expression

The expression of genes occurs at the greatest rate when bacteria are in a mid-log phase. Bacterial growth was monitored by density measurements. By using a Spectronic 21 (Milton Roy Company) the absorption of the cells at 600 nm (OD₆₀₀) was measured regularly, first with 30 minute intervals, and when closing up to the mid-log phase each 10 minutes. 2.6 ml of overnight cultured transformants of BL21 (DE3) plysS from section 2.5.1 were transferred to 80 ml of LB+Amp₁₀₀+Cam₃₄, and incubated at 37°C with shaking (185 rpm). This was done with transformants with *hfq* insert from all three species, and an un-induced control transformant with an insert from *T. africanus*. When OD₆₀₀ was approximately 0.4, gene expression was induced by adding IPTG to a final concentration of 0.5 mM to all the cultures except the control. The cultures were kept in a shaking incubator at 37°C for three hours

before a new OD₆₀₀ measurement was made. The growth of the induced cells was compared to the growth in the control culture. Little or no growth would indicate overexpression of the cloned gene. Two ml of each culture were transferred to separate Eppendorf tubes and centrifuged at 13,000 rpm for five minutes. The pellet was later analysed by SDS-PAGE. The rest of the culture was transferred to separate Falcon tubes, centrifuged at 13,000 rpm for 15 minutes and kept in a freezer.

2.5.3 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE is a method that allows separation of proteins according to molecular weight. SDS ($\text{CH}_3\text{-(CH}_2\text{)}_{10}\text{-CHOSO}_3^-\text{Na}^+$) is an anionic detergent. It binds to amino acids and denatures proteins, and also gives it a negative charge. The binding is relative; the larger the protein, the more SDS binding, resulting in a greater negative charge. The separation is performed in a polyacrylamide gel by electrophoresis and the gel is stained with Coomassie Brilliant Blue afterwards. Coomassie binds to proteins by hydrophobic and electrostatic interactions, and gives a clear blue colour so that the bands can be visualized. The intensity of the blue colour will vary according to the amount of protein. The size of the proteins can be estimated by comparing to a commercial protein MW ladder.

Two polyacrylamide gels were made according to Appendix B, table B.9, a 4% stacking-gel and a 20% separation gel. The purpose of the stacking gel is to allow the proteins to easily penetrate to the beginning of the gel, and start migrating at the same position in the gel. The acrylamide is what makes the gel solid, and the higher the percentage of acrylamide in the gel the slower the proteins migrate through. The percentage of acrylamide in the gel is decided based on the size of the proteins that are examined. To be able to separate small proteins, a high percentage of acrylamide is used. EcoGene (www.ecogene.org) was used to calculate the actual size of the Hfq proteins, and these values were used to compare to the proteins in the gel to the standard. As the Hfq protein is approximately 9kDa, a 20% separation gel was used. Eppendorf tubes containing the cell pellets (from section 2.5.2) were mixed with 80 µl of sample buffer (Appendix B, table B.10). The tubes were then put in a heating block at 95°C for seven minutes and spun at 13,000 rpm for 30 seconds. A broad range standard was loaded in the first well, and six µl of each sample were loaded in the

subsequent wells of the gel. The gel was run at 190V for 45 minutes in a Model 1000/500 Power Supply (Bio-Rad). The gel was stained with a Coomassie staining solution (Appendix B, table B.12), and left at constant shaking (60 rpm) for 90 minutes. The gel was rinsed in distilled water before it was decoloured in a destaining solution (Appendix B, table B.13), and left at constant shaking for 90 minutes. The results were visualized in a transilluminator (Molecular Imager ChemiDoc xRST with Image lab software, Bio-Rad).

2.6 Testing the solubility of the Hfq protein

Soluble proteins should be found in the supernatant of a centrifuged sample of lysed cells. The solubility of the Hfq protein was tested by analysing the samples of supernatant and pellet from cells with expressed Hfq by SDS-PAGE.

Falcon tubes containing cell pellets were mixed with three ml HisTALON xTractor Buffer. A small amount of DNaseI was added, and the sample was incubated on ice with intermittent mixing for 15 minutes. The cells were lysed by sonication in a Sonifier 250 (Branson). Thereafter the sample was centrifuged at 13,000 rpm at 10°C for 25 minutes. The supernatant was separated from the pellet and both the supernatant and the pellet were used for SDS-PAGE. The gel was run at 190V for 110 minutes. Afterwards it was stained with Coomassie.

2.7 Antibody production

Antibodies are a part of the immune system of higher vertebrates. When vertebrates are exposed to a foreign molecule (antigen, in this case Hfq), the B-cells will produce antibodies that are specific and bind to the antigens. This is a helpful property for researchers when a specific protein, in this case Hfq, is to be detected.

A fresh gel was made to isolate the insoluble Hfq protein and cut it out of the gel. This was done according to section 2.5.3. Approximately 500 mg of protein in the gel was cut out using a scalpel, and sent to BioGenes (www.biogenes.de) for antibody production. The immunization was done in two rabbits, and a standard immunization protocol was followed:

1st day: Pre-immune serum (1.5 ml per rabbit) and first immunization

7th day: Second immunization

14th day: Third immunization

28th day: Bleeding (20 ml per rabbit)

The pre-immune serum and Hfq antiserum were shipped to Bergen on ice.

2.7.1 Western blotting

Western blotting is a method that allows the detection of specific proteins by the use of antibodies. In this procedure, the denatured proteins from an SDS-PAGE are transferred from the polyacrylamide gel to a nitrocellulose membrane by electrotransfer. Specific proteins can then be detected by incubating the nitrocellulose membrane with antibodies. For the bound antibodies to be detectable, a secondary antibody with a reporter group is needed which binds to the primary antibody. In this study goat anti-rabbit IgG conjugated to the enzyme horseradish peroxidase (HRP) was used. The reporter group (HRP) allows detection of the protein by oxidizing a substrate in the staining solution (Appendix B, table B.21) which results in a permanent purple colour on the membrane where the protein is found.

An SDS-PAGE was performed according to section 2.5.3. The proteins in the polyacrylamide gel were transferred to a nitrocellulose membrane (Trans-Blot Transfer Medium pure Nitrocellulose Membrane, 0.45 µm, Bio-Rad) by electrotransfer. The samples were run in duplicates to allow one gel to be stained and one gel to be blotted. By staining the gel after blotting the efficiency of protein transfer could be assessed. The instructions given by the supplier (Bio-Rad) were followed, and the recipes for the solutions are found in Appendix B. The gel and membrane were assembled like a sandwich, with a sponge, three filter papers, gel, membrane, three filter papers and a sponge. All components were incubated in 1X blotting buffer for five minutes prior to assembly. The sandwich was placed in the Trans blot cell with the membrane facing the anode, and gel facing the cathode. An ice block was placed in the cell as a cooling element, and the cell was filled with 1X blotting buffer. Due to the negative charge the proteins migrate towards the anode, from the gel to the membrane. The blotting was run at 100V for approximately two hours and afterwards the membrane

was incubated in 1X TBS for five minutes. All the proceeding steps were performed in room temperature at shaking (60 rpm). The membrane was incubated in blocking buffer for 60 minutes to prevent unspecific binding of antibodies. The membrane was then washed in TTBS for ten minutes and incubated in the antibody solution for two hours or overnight. The membrane was then washed two times in TTBS for five minutes and then incubated in the conjugate solution, containing the secondary antibody that binds the primary antibody, for 90 minutes. To wash off excess secondary antibodies the membrane was washed three times five minutes; first in TTBS two times and then in TBS. The membrane was incubated in staining solution for colour development. The reaction was stopped after 30 minutes maximum by washing the membrane in Milli-Q water. The membrane was dried on a filter paper and stored between two filter papers at 4°C.

2.8 Identification of Hfq-bound RNAs

The presence of Hfq proteins could be confirmed by analysing cell lysate by Western blot. This was first done by using the expression clone to confirm the specificity and quality of the antibodies. The experiment was thereafter done on cultures of *T. africanus*. As it is possible that Hfq is only expressed under certain conditions, cell lysate from *T. africanus* cultivated at near threshold and optimal temperatures was used when trying to identify the protein. Samples that had been immunoprecipitated were analysed by Western blot before RNA-extraction and cDNA-synthesis was performed.

2.8.1 Estimation of exponential phase

One ml of *T. africanus* cells from a dense culture was transferred to a 50 ml serum bottle containing the appropriate medium. OD₆₀₀ was measured regularly to estimate exponential phase.

2.8.2 Cultivation of *T. africanus* at threshold temperatures

T. africanus was cultivated at 70°C in 50 ml medium for 6.5 hours (estimated exponential phase). One culture was then moved to a heating cabinet holding 37°C, and the other was moved to 77°C (lowest and highest threshold temperatures, respectively) and incubated for 3.5 hours. The cultures were spun down at 13,000 rpm for 30 minutes. For whole cell samples the sample was mixed with sample buffer (1:1) before an SDS-PAGE and Western blotting were performed (according to section 2.5.3 and 2.7.1).

2.8.3 Co-immunoprecipitation (co-IP)

Co-IP was done to precipitate Hfq in complex with RNAs from a *T. africanus* cell lysate. The precipitation was assisted by Sepharose A beads. Co-IP was first done only with *T. africanus* cell lysate and Hfq antiserum, and secondly with Sepharose A beads added.

Sepharose A beads were prepared by allowing 0.1 gram of the beads to swell in 20 ml Milli-Q water. The beads were allowed to settle by gravity before the supernatant was pipetted out. This was repeated three times before 20 µl of 10X TBS were added to 200 µl of beads in Milli-Q water, and the solution was ready to use.

A dense culture of *T. africanus* was pelleted and washed in 1X TBS, and the pellet was resuspended in 200 µl 1X TBS. The cells were lysed by vortexing for 90 seconds with glass beads added in a volume ratio of 1:4. Microscopy was done to confirm a successful lysis. The sample was spun down at 13,000 rpm for three minutes and the supernatant was transferred to a fresh tube before it was respun. The supernatant was again transferred to a fresh tube and was ready for use. The pellet was expected to contain cell debris, and the supernatant was used for further analysis. Precipitation was done with and without Sepharose A beads in a pilot experiment. Reaction mixtures were prepared as described in Table 9 and 10, and were left to precipitate at 37°C for four hours before being spun down at 13,000 rpm for six minutes. After centrifugation 10 µl of supernatant was mixed with 10 µl of SDS-PAGE sample buffer while the pellet was resuspended in 20 µl of SDS-PAGE sample buffer. Both the supernatant and the pellet were analysed by SDS-PAGE and Western blot (as described in section 2.5.3 and 2.7.1).

Table 9: Reaction mixtures used for co-IP.

Component	Sample 1	Sample 2 (control)
<i>T. africanus</i> supernatant	18 μ l	18 μ l
10X TBS	2 μ l	2 μ l
Hfq antiserum	2 μ l	
Pre-immune serum		2 μ l
Total volume	22 μ l	22 μ l

Table 10: Reaction mixtures used for co-IP with Sepharose A.

Component	Sample 1	Sample 2 (control)
<i>T. africanus</i> supernatant	100 μ l	100 μ l
Sepharose A	15 μ l	15 μ l
Hfq antiserum	8 μ l	
Pre-immune serum		8 μ l
Total volume	123 μ l	123 μ l

2.8.4 Isolation of RNA

The GeneRacer Kit (Invitrogen) was used to amplify cDNA from total RNA. A slightly modified procedure given by the supplier was used. Four serum bottles containing freshly grown cultures of *T. africanus* were placed in an ethanol bath holding approximately -80°C , and left there until the temperature in the bottles were $<10^{\circ}\text{C}$. The bottles were opened and the culture was poured into six pre-cooled (-20°C) Falcon tubes (45 ml in each). The tubes were centrifuged at 13,000 rpm for 20 minutes at 4°C . This work was conducted rapidly to keep the cultures at a low temperature, and prevent exposure to large amounts of oxygen. All of the proceeding work was done on ice.

After centrifugation the supernatant was discarded and the pellets were resuspended in five ml pre-cooled 1X TBS and collected in one tube. The centrifugation step was then repeated, and following centrifugation the pellet was resuspended in one ml pre cooled 1X TBS and 25 μ l of RNasin corresponding to one unit per μ l. The cells were lysed by vortexing with glass beads, and a successful lysing was confirmed by microscopy. The sample was spun down at 13,000 rpm for 2X five minutes and transferred to a new Eppendorf tube between the centrifugation steps. A co-IP was conducted using the reagents described in Table 11. Incubation of the co-IP mix was done in a shaking cabinet holding 37°C and 50 rpm for four hours.

Following incubation, the samples were washed in a solution of 1X TBS and RNasin (RNasin; 1/2 unit per μl). The sample was washed in 200 μl and the control in 50 μl , and spun down for 30 seconds at 300 rpm for the beads to fall out. The supernatant was discarded, and the washing step was repeated four times. After the last wash the samples were spun down for three minutes at 10,000 rpm.

Table 11: Reaction mixtures used for co-IP.

Component	Sample 1	Sample 2 (control)
Lysed cells of <i>T. africanus</i>	900 μl	200 μl
Antiserum	72 μl	
Pre-immune serum		16 μl
Sepharose A	135 μl	30 μl
Total volume	1107 μl	246 μl

A phenol extraction was done to release the binding between the beads, antibody, protein and RNA, hereby allowing precipitation of only the RNA. For each sample 90 μl diethylpyrocarbonate treated water (DEPC water) and 100 μl of phenol:chloroform were added. Samples were then vortexed vigorously for 30 seconds and centrifuged at room temperature for five minutes at 13,000 rpm. The aqueous phase was transferred to a new tube before adding two μl of 10 mg/ml mussel glycogen and 10 μl 3 M sodium acetate (pH 5.2). Samples were mixed well before 220 μl of 99.8% ethanol was added. Samples were vortexed for 10 seconds and stored at -20°C .

Samples were centrifuged at 13,000 rpm at 4°C for 20 minutes to precipitate the RNA. The supernatant was discarded and 500 μl of 70% ethanol (350 μl 99.8% ethanol and 150 μl DEPC treated water) was added to the pellet. The samples were inverted a few times before centrifugation at 13,000 rpm at 4°C for two minutes. The supernatant was discarded and the centrifugation step was repeated to enable removal of ethanol remnant via pipetting. Remaining ethanol was removed by air-drying the samples in a fume hood for approximately two minutes at room temperature. The pellet was resuspended in 10 μl DEPC water from which one μl was used for determining the RNA concentration with NanoDrop, and the remaining sample was added one μl RNase OUT and left on ice.

2.8.5 cDNA synthesis

For the polyadenylation of RNA a Poly(A) Tailing Kit (Ambion, Life Technologies) was used. A slightly modified procedure given by the supplier was used.

Table 12: Reaction mixture used for the polyadenylation of RNA using *E. coli* Poly(A) Polymerase I (E-PAP) (Ambion, Life Technologies).

Component	Sample 1 (with antiserum)	Sample 2 (control with pre-immune serum)
RNA-extract	10 μ l	10 μ l
Nuclease-free water	18 μ l	18 μ l
5X E-PAP Buffer	10 μ l	10 μ l
25 mM MnCl ₂	5 μ l	5 μ l
10 mM ATP	5 μ l	5 μ l
E-PAP	3 μ l (6 units)	3 μ l (6 units)
Total	51 μ l	51 μ l

The components in Table 12 were mixed in the listed order, and incubated in a heating block at 37°C for one hour. After incubation, a phenol extraction followed by an RNA precipitation was conducted according to section 2.8.4. The samples were stored at -20°C for two days to allow precipitation before the RNA was pelleted and washed.

For the 5' ligation of RNA oligo to RNA, seven μ l of RNA sample was transferred to a pre-aliquoted, lyophilized GeneRacer RNA Oligo (0.25 μ g), and mixed by pipetting up and down. The samples were centrifuged briefly before being incubated in a heating block at 65°C for five minutes to relax the secondary structure of the RNA. They were then left on ice for two minutes and centrifuged briefly. The reagents listed in Table 13 were added, and the samples were mixed by pipetting before being centrifuged briefly and incubated in a water bath at 37°C for one hour.

Table 13: Reaction mixtures used for ligation of RNA oligo to RNA (Invitrogen).

Component	Amount
Sample	6 μ l
10X Ligase Buffer	1 μ l
10 mM ATP	1 μ l
RNaseOut (40 U/ μ l)	1 μ l
T4 RNA ligase (5 U/ μ l)	1 μ l
Total volume	10 μ l

The samples were centrifuged briefly and placed on ice before a new phenol extraction and RNA precipitation was conducted (section 2.8.4). They were stored at -20°C for two days to allow precipitation before the RNA was pelleted and washed. The pellet was resuspended in 11 μ l DEPC water, and one μ l was analysed by agarose gel electrophoresis to determine the content of RNA.

For reverse transcription of RNA one μ l of GeneRacer oligo dT primer and one μ l of dNTP Mix was added to the ligated RNA samples (10 μ l). To relax secondary structure of the RNA the samples were incubated at 64°C for five minutes prior to reverse transcription. The samples were chilled on ice for two minutes and centrifuged briefly before the reagents listed in Table 14 were added.

Table 14: Reaction mixture used for reverse transcription (Invitrogen).

Component	Amount
5X RT Buffer	4 μ l
Cloned AMV RT (15 U/ μ l)	1 μ l
Sterile water	2 μ l
RNaseOut (40U/ μ l)	1 μ l
Total	20 μ l

The samples were mixed by pipetting up and down, and incubated at 45°C for one hour. The Cloned AMV RT was inactivated by incubating the samples at 85°C for 15 minutes. The samples were centrifuged briefly, and two μ l were used directly for amplification. Remains of the samples were stored at -20°C. For amplifying cDNA a PCR was conducted with the components listed in Table 15.

Table 15: Composition of the PCR mix used for amplifying the cDNA.

	Sample 1: antiserum	Sample 2: pre-immune serum	Sample 3: negative control
GeneRacer 5' Primer, 10 μ M	3 μ l	3 μ l	3 μ l
GeneRacer 3' Primer, 10 μ M	3 μ l	3 μ l	3 μ l
RT Template	2 μ l	2 μ l	
5X PCR Buffer	10 μ l	10 μ l	10 μ l
dNTP Solution (10 μ M)	5 μ l	5 μ l	5 μ l
<i>Taq</i> DNA polymerase (5 U/ μ l)	0.5 μ l	0.5 μ l	0.5 μ l
Sterile water	26.5 μ l	26.5 μ l	28.5 μ l
Total	50 μ l	50 μ l	50 μ l

The components were mixed by pipetting and PCR conducted under the following conditions:

94°C 2 min
94°C 30 sec } 30 cycles
68°C 90 sec }
68°C 10 min
4°C ∞

A second amplification was done with the same components and PCR conducted under the following conditions:

94°C 2 min
94°C 30 sec } 30 cycles
73°C 30 sec }
68°C 30 sec }
68°C 10 min
4°C ∞

Following PCR, the amplicons were analysed by agarose gel electrophoresis.

2.8.6 Sequencing of cDNA

Bands that were considered interesting were excised from the gel, and excess gel was trimmed away to minimize the amount of agarose. The gel slices were weighed, and a gel extraction was done using a GenElute Gel Extraction kit (Sigma-Aldrich), following the procedure provided by the supplier. For dissolving the gel slices, Gel Solubilizing Solution was added; 153 μ l to 51 mg of the antiserum sample, and 291 μ l to 97 mg of the pre-immune serum sample. The gel mixtures were incubated at 55°C for 10 minutes, with intermittent mixing. Binding columns were prepared by adding 500 μ l of Column Preparation Solution to each binding column. The binding columns were centrifuged at 13,000 rpm for one minute and the flow-through liquid was discarded. The solubilized gel solution was loaded into the prepared binding column and centrifuged at 13,000 rpm for one minute before the flow-through liquid was discarded. The binding column was then added 700 μ l of Wash Solution, and the centrifugation step was repeated two times, discarding the flow-through liquid between the centrifugations. To elute the cDNA, the binding column was transferred to a fresh tube and added 50 μ l of Elution buffer to the center of the membrane. This was incubated for one minute and centrifuged at 13,000 rpm for one minute. The cDNA was analysed by agarose gel electrophoresis and NanoDrop. The cDNA was sequenced as described in section 2.4.6.3, using 6-8 ng template.

2.8.7 Sequencing of single cDNAs

Using the gel-extracted cDNA from section 2.8.6 as a template a PCR was run using the composition described in Table 15, under the conditions described for the second amplification in section 2.8.5. The PCR-products were analysed by agarose gel electrophoresis. Cloning was performed in chemically competent *E. coli*, using the TOPO TA Cloning Kit with pCR4 (Invitrogen). The instructions from the kit manufacturer were followed.

Reagents were added in the following order, and mixed by gently pipetting up and down; 0.5 μ l of PCR-product, 1 μ l salt solution, 3.5 μ l milli-Q water, 1 μ l vector. The mix was incubated at room temperature for five minutes, placed on ice, and 2 μ l of this mix was added to a vial of One Shot Competent cells, carefully mixed and incubated on ice for 20 minutes. Cells

were then heat-shocked at 42°C for 30 seconds, transferred to ice and left for two minutes. Room-tempered SOC-medium (250 µl) was added, and the vials were horizontally incubated with shaking (200 rpm) at 37°C for one hour. For blue-white screening, agar plates were prepared by spreading 40 µl X-gal on their surface, and warming them at 37°C for 30 minutes. Pre-warmed selective plates were added five, 10 and 30 µl respectively, of transformation mix, and incubated at 37°C overnight. Following incubation, 10 white colonies were picked and put directly to the PCR-mastermix.

A PCR reaction was performed in 25 µl total reaction mixtures, using 0.2 µM of M13 primers (forward and reverse), one µl BSA (2%), 0.2 mM dNTPs and 0.6 Units One Taq Polymerase, and 5µl of One Taq standard buffer. The PCR-reaction was run under the following conditions:

94°C	1 min		
94°C	30 sec	}	30 cycles
55°C	45 sec		
72°C	1 min		
72°C	6 min		
4°C	∞		

The PCR-products were analysed by agarose gel electrophoresis. ExoSap (Affymetrix UK) was used for cleaning up the PCR-products by adding 2 µl of ExoSap to 5 µl of PCR-product. The mix was incubated at 37°C for 15 minutes to degrade remaining primers and nucleotides. ExoSap was then inactivated by an incubation at 80°C for 15 minutes. The DNA concentration was measured by NanoDrop. A sequencing reaction was performed in 10 µl total reaction mixtures, using 0.5 µl template, 3.2 pmol of M13 primers (forward or reverse), one µl sequencing buffer, 5.9 µl MQ and one µl BigDye. The sequencing reaction was run under the conditions described in section 2.4.6.3. The reaction tubes were added 10 µl Milli-Q water, and delivered for sequencing.

2.9 Ouchterlony double immunodiffusion

An Ouchterlony double immunodiffusion test (Ouchterlony 1958) was done to verify the cross reaction between antigens from *T. africanus* and antibodies in the antiserum, pre-immune serum or the secondary antibody used for western blot and immunoprecipitation.

Two Petri dishes were filled with 20 ml 1% agar made with 1X TBS (Appendix B, table B.22). Wells were cut in the gel using a small glass cylinder made up of a sawed-off Durham tube. Supernatant from lysed *T. africanus* cells was pipetted into the centre well, and the surrounding wells were filled with either antiserum, secondary antibodies or pre-immune serum. The Petri dishes were surrounded by parafilm and left in a plastic bag with a wet piece of paper to prevent loss of moisture. The Petri dishes were left on the bench, at room temperature until a precipitation line was observed. They were stained in Coomassie, and decoloured in destaining solution as described in section 2.5.3.

2.10 Proteomics analysis

A cell pellet of *T. africanus* resuspended in 1X TBS was lysed using glass beads and a combination of vortexing and sonication. Lysis was confirmed by light microscopy. The lysate was centrifuged at 13,000 rpm in a biofuge for 10 min. The supernatant was carefully removed and transferred to a new tube. The sample was in-solution digested with trypsin and analysed by a NanoLC-MSMS (Orbitrap velos pro) at PROBE (the Proteomics unit at the University of Bergen; <http://www.uib.no/en/rg/probe>). The raw data was further analysed in Peptide Shaker (<http://www.uib.no/en/rg/probe/65218/peptideshaker>).

2.11 Bioinformatics

2.11.1 Phylogenetic tree

BLAST was used to obtain the sequence accession numbers of the 16S rRNA- and *hfq* genes from Thermotogales species and *Paenibacillus polymyxa*. Multiple sequence alignments based on 16S rRNA and *hfq* were made in ClustalX (Larkin, Blackshields et al. 2007). Based on these alignments phylogenetic trees were made in Mega 6 (Molecular Evolutionary Genetics Analysis) (Tamura, Stecher et al. 2013). The Neighbour Joining algorithm was used which gives a distance matrix where the distance between each taxa represents evolutionary relatedness. The bootstrap function was used to assess the quality of the tree. By using the bootstrap function one hundred replicates of the phylogenetic tree were made and the likelihood for the same branching to occur in all resulting trees is calculated. A high bootstrap value indicates a likely branching.

2.11.2 Multiple sequence alignment

A structure-based multiple sequence alignment of Hfq from species of Thermotogales and model species was made in ClustalX. The positions of the α -helix and β -sheets of the Hfq protein were decided by using *E. coli hfq* as a template. Highly conserved regions were marked. Amino acids shown to be involved in RNA-binding were marked based on findings in Hfq from *Staphylococcus aureus* (Schumacher, Pearson et al. 2002).

2.11.3 Prediction of protein structure

SWISS-MODEL is a protein structure homology-modelling server which was used for predicting the structure of TaHfq (<http://swissmodel.expasy.org>). The crystal structure of *Pseudomonas aeruginosa* Hfq has been determined with a resolution of 1.60 Å (Nikulin, Stolboushkina et al. 2005), and was used as a template when modelling TaHfq. The model was modified in Pymol (DeLano 2002). The RNA-binding site was determined based on the RNA-binding sites that has been identified in *Staphylococcus aureus* Hfq (Schumacher, Pearson et al. 2002).

3. Results

3.1 Cloning

3.1.1 Preparation of genomic DNA

K. olearia, *P. mobilis* and *T. africanus* were cultured in serum bottles holding 50 ml medium under anaerobic conditions (section 2.1). They were incubated in heating cabinets at 65°C, 55°C and 70°C, respectively. When the cultures were turbid (after two days to one week of incubation), genomic DNA was extracted from the three strains using the CTAB method. The quantity and quality of the DNA was assessed by NanoDrop (Appendix A, table A.1) and agarose gel electrophoresis (Fig. 10). Three distinct bands can be seen in the upper part of the gel in lane 2, 3 and 4, verifying the presence of high molecular weight genomic DNA.

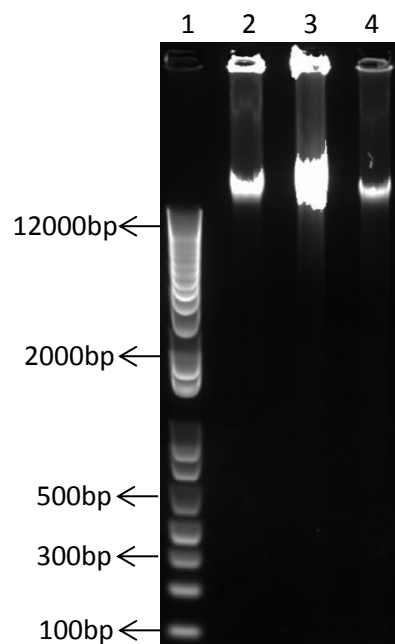


Figure 10: Agarose gel electrophoretic analysis of genomic DNA. Lane 1: DNA ladder. Lane 2: *K. olearia*. Lane 3: *T. africanus*. Lane 4: *P. mobilis*.

3.1.2 Gene amplification

Hfq genes were amplified and analysed by NanoDrop (Appendix A, table A.2) and agarose gel electrophoresis. The fragments are approximately 250-300 bp (Fig. 11). Compared to the expected length of the *hfq* genes; 240 bp for *K. olearia*, 243 bp for *P. mobilis* and 255 bp for *T. africanus*, this indicates successful amplification of the genes.

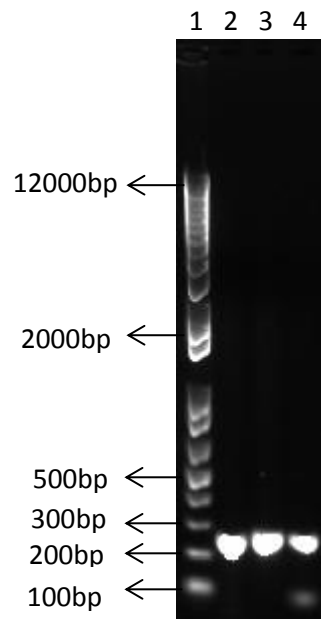


Figure 11: Agarose gel electrophoretic analysis of the amplified *hfq* genes. Lane 1: DNA ladder. Lane 2: *hfq* from *K. olearia*. Lane 3: *hfq* from *T. africanus*. Lane 4: *hfq* from *P. mobilis*.

3.1.3 Digestion of PCR-products and vector

To prepare the PCR products and vector for ligation they were digested with the restriction enzymes *Nde*I and *Hind*III, and the digested DNA was analysed by NanoDrop (Appendix A, table A.3) and agarose gel electrophoresis (Fig. 12). The vector was clearly linearized (Fig. 12b).

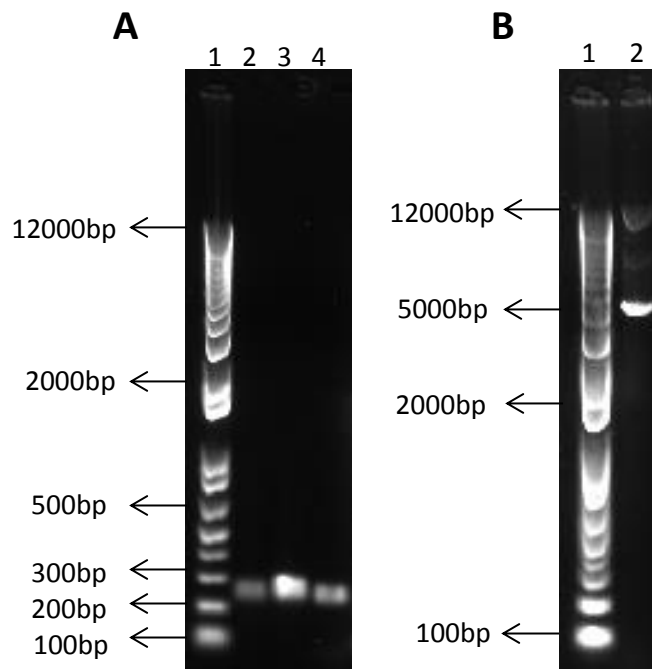


Figure 12: Agarose gel electrophoretic analysis of the digested *hfq* PCR products (**A**) and pET 21b (+) vector (**B**). **A:** Lane 1: DNA ladder. Lane 2: *hfq* from *K. olearia*. Lane 3: *hfq* from *T. africanus*. Lane 4: *hfq* from *P. mobilis*. **B:** Lane 1: DNA ladder. Lane 2: pET 21b(+) vector. pET 21b(+) is originally 5442 bp. 57 bp was cut (between the *Nde*I and *Hind*III restriction sites). Expected length after cutting was 5385 bp.

3.1.4 Ligation and transformation

Hfq from each of the bacterial strains was ligated into a pET-21b(+) vector, and the ligation mixture used for transformation of *E. coli*. For selection of transformants different volumes of the transformation mixes were plated on LB+Amp₁₀₀ plates as described in section 2.4.4. Following an overnight incubation, several hundred colonies were obtained. Four single and well separated colonies from each cloning series were picked and plated onto fresh LB+Amp₁₀₀ plates. Following incubation overnight, single colonies were transferred to culture tubes containing five ml LB+Amp₁₀₀. Ligation and transformation was first performed with *hfq* from *T. africanus* and *P. mobilis*. The plasmids were extracted and purified, and a successful result was verified by gel electrophoresis (Fig. 13) and NanoDrop (Appendix A, table A.4). The bands in lane 1 to 6 and in lane 8 are around 6000 bp (Fig. 13). Compared to the length of pET-21b (+) of 5385 bp without any insert - this implies a successful cloning. A successful cloning of *hfq* from *K. olearia* was also completed.

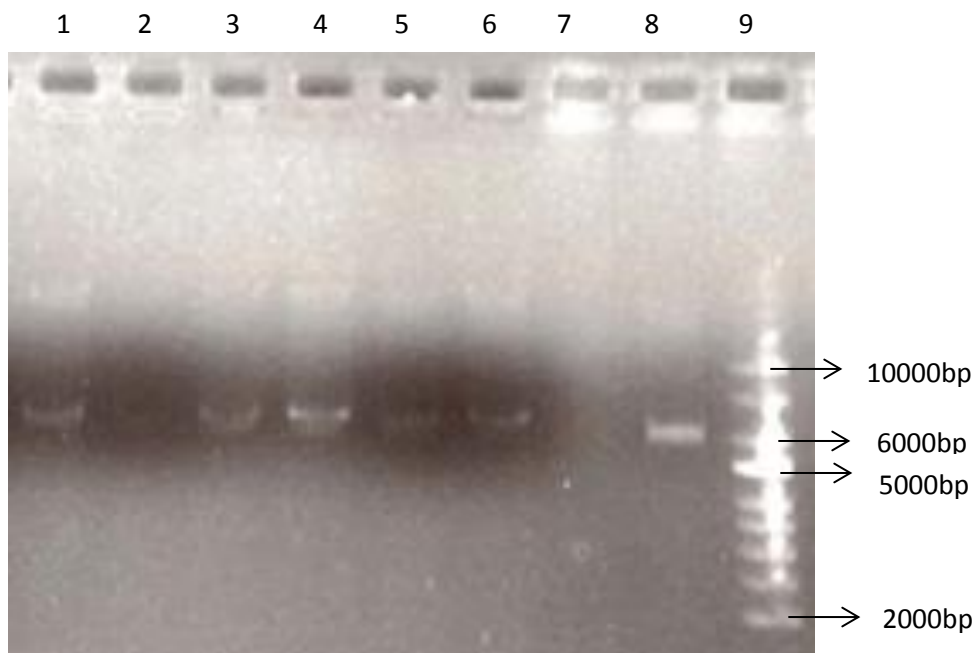


Figure 13: Agarose gel electrophoretic analysis of purified plasmids from transformants. Lane 9, Supercoiled DNA ladder; Lane 1-3, *T. africanus* clones; Lane 4-8, *P. mobilis* clones. Expected length of a plasmid with insert was 5619-5632 bp. Expected length for an unsuccessful clone was 5385 bp.

3.1.4.1 Verification of plasmid inserts

A PCR using the designed primers was performed for one clone from each series to confirm that the clones contained the inserts with the expected size. The size of the PCR products was estimated to be 250-300 base pairs (Fig. 14), which was the expected size of the inserts.

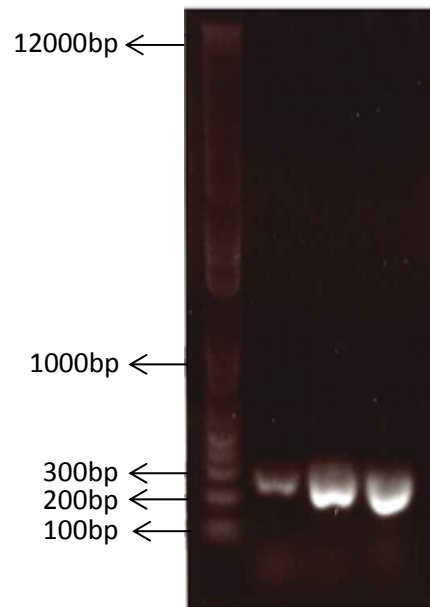


Figure 14: Agarose gel electrophoretic analysis of *hfq* amplicons from clones. In lane 1: DNA Ladder. Lane 2: *hfq* from *K. olearia*. Lane 3: *hfq* from *T. africanus*. Lane 4: *hfq* from *P. mobilis*. Expected length was between 250 and 270 bp.

To further confirm that the inserts had the correct sequence the inserts were sequenced. The results confirmed that *hfq* from all three strains had been correctly inserted and that no mutations had occurred during PCR and cloning (Appendix A, table A.5).

3.2 Overexpression

Plasmids containing the *hfq* gene were transformed into One Shot BL21 (DE3) plysS ultracompetent cells, and gene expression was induced at mid-log phase by adding IPTG to the cultures. All the cultures grew after induction, but not as much as the control (Appendix A, table A.6 and Fig. A.1). The result was analysed by SDS-PAGE (Fig. 15), where the approximate size of the proteins can be observed. The expected size of the Hfq protein was calculated in EcoGene (ecogene.org), and is 9.8 kDa (84 amino acids) for *T. africanus*, 9.3 kDa (80 amino acids) for *P. mobilis* and 9.3 kDa (79 amino acids) for *K. olearia*. A strong band is detected and encircled in the lane with the *T. africanus hfq* clone in the lower part in lane 3 (Fig. 15). This protein has the correct size range for the Hfq protein, indicating a successful expression. A similar band is not seen for the other two clones. The negative control (lane 5), which was the *T. africanus hfq* clone without IPTG induction, did not give a band in the Hfq size range as expected. As only expression of *TaHfq* appeared to be successful it was decided to proceed with *TaHfq* alone.

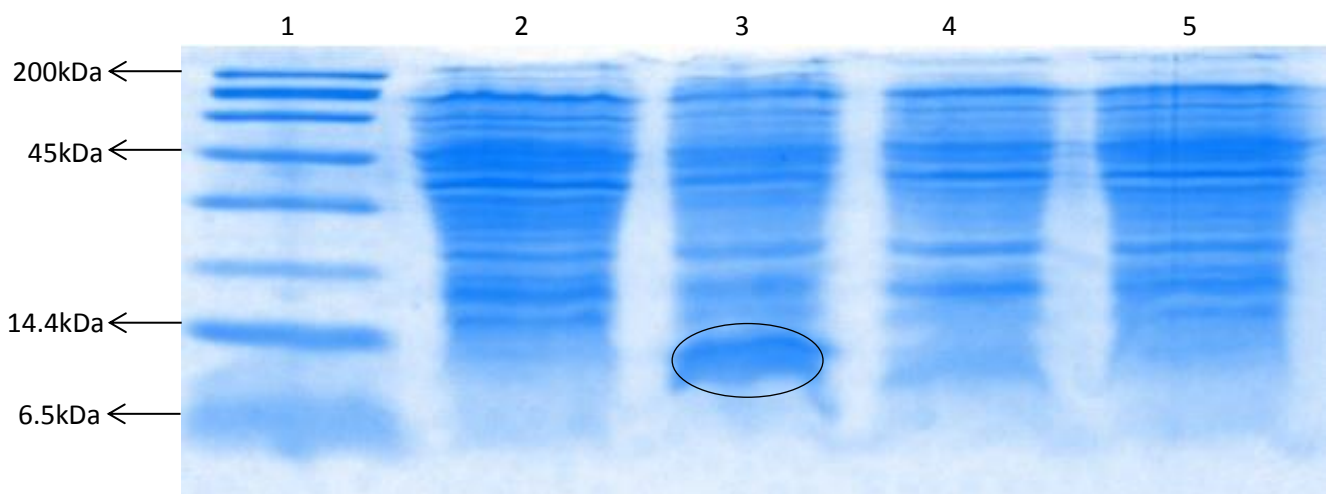


Figure 15: SDS-PAGE (20%) analysis of protein expression from clones with *hfq* gene insert from different strains; Lane 2: *K. olearia*. Lane 3: *T. africanus*. Lane 4: *P. mobilis*. Lane 5: Un-induced control *T. africanus*. Lane 1 contains MW marker (broad range). A strong band is encircled in the lower part in lane 3.

3.2.1 Testing the solubility of the *TaHfq* protein

Cells that had been induced for expression of the *TaHfq* protein were lysed and the lysate was centrifuged. Following separation of the supernatant and pellet, the two fractions were analysed by SDS-PAGE (Fig. 16) to check if Hfq was soluble (found in the supernatant) or insoluble (found in the pellet). A strong band representing a protein in the size range of ~10 kDa was detected in the pellet (lanes 3 and 5), while only a faint band was observed in the same region (or slightly above) for the supernatant. This shows that the recombinant *TaHfq* is insoluble. The band was subsequently cut from the gel (approximately 500 mg of protein), and used for production of antibodies as described in section 2.7.

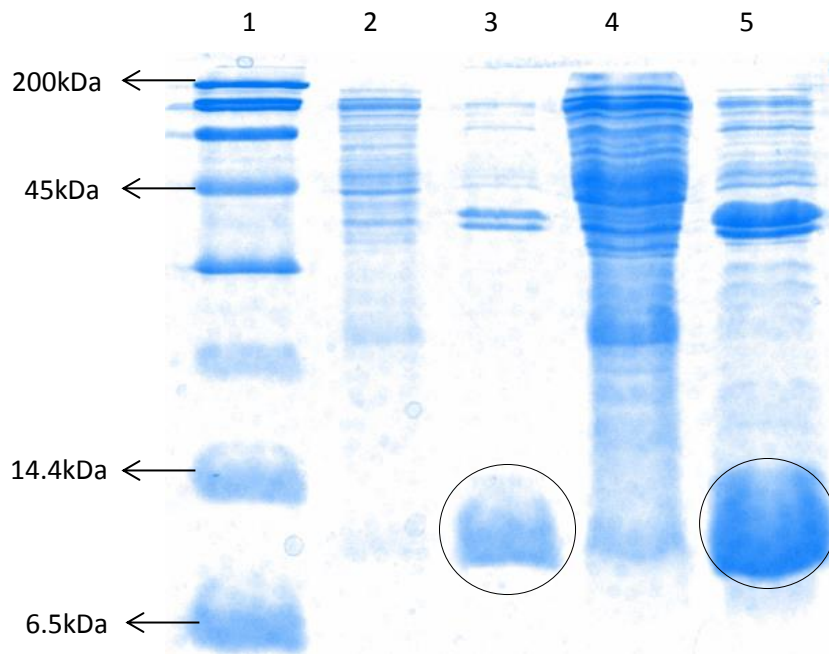


Figure 16: SDS-PAGE (20%) analysis of proteins from supernatant and pellet after expression and lysis of a clone with *T. africanus hfq*. Lane 1: Standard (broad range). Lane 2: Supernatant, 1 μl. Lane 3: Pellet, 1 μl. Lane 4: Supernatant, 5 μl. Lane 5: Pellet, 5 μl.

3.3 *T. africanus* growth curve

A freshly inoculated *T. africanus* culture was incubated at 70°C, and OD₆₀₀ was measured regularly for ten hours to estimate the time where the culture grew exponentially. The growth curve showed that the culture grew exponentially from 0-10 hours after inoculation (Appendix A, table A.7 and Fig. A.2). From this it was decided to harvest cells for RNA extraction following 6.5 hours of incubation, where the culture should be in the middle of exponential growth.

3.4 Western blot analysis

The rabbit antiserum raised against the *TaHfq* was first tested by Western blot, using an SDS-PAGE gel containing proteins from an induced clone and from a *T. africanus* whole cell lysate. Different dilutions of the primary antibody were used to test its specificity and efficiency, and as a dilution of 1:1000 gave the best result, this concentration was used in the rest of the experiments (gels not shown). The results in Fig. 17 showed that recombinant *TaHfq* gave a specific and strong signal (lane 2).

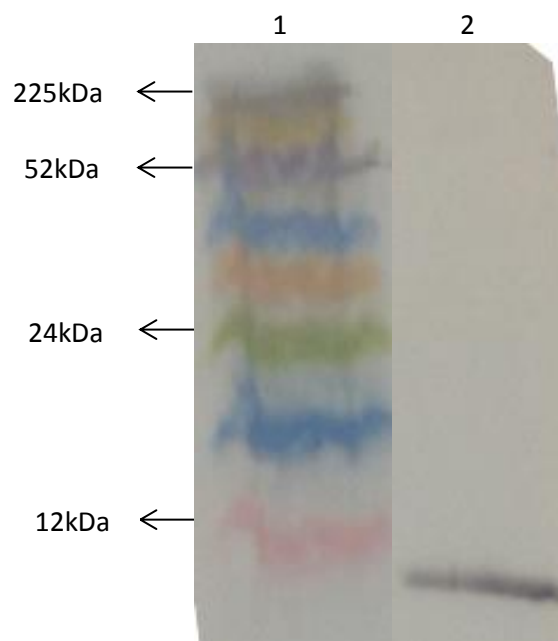


Figure 17: Western blot analysis assessing the specificity and efficiency of the anti-Hfq antiserum. Lane 1: Rainbow ladder. Lane 2: Whole cell lysate of *E. coli* expressing the *T. africanus hfq*.

Testing of a *T. africanus* lysate gave a signal corresponding to the correct size (Fig. 18, lanes 2 and 3), but cross-reaction with proteins in the upper part of the gel indicated that the antiserum was not completely specific. The Hfq protein in cell extracts from *T. africanus* appeared to be soluble as it was detected also after centrifugation of the lysate (Fig. 18, lane 3).

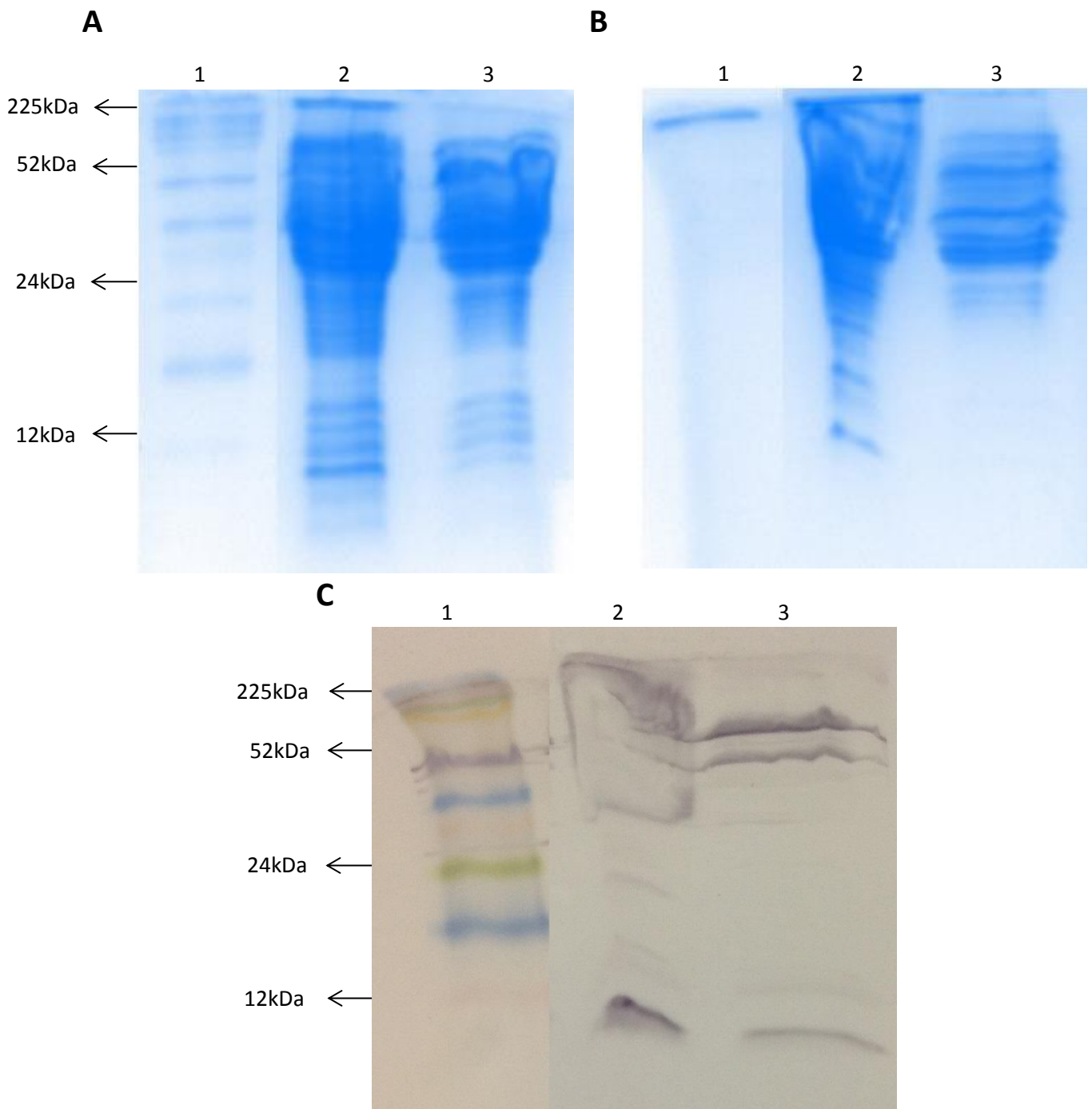


Figure 18: Western blot analysis of proteins from *T. africanus* cultures, whole cells and supernatant from lysed cells, grown at optimal temperature. **A:** Coomassie stained SDS-PAGE prior to protein transfer. **B:** Coomassie dyed SDS-PAGE after protein transfer. **C:** Western blot. Lane 1: Rainbow ladder. Lane 2: *T. africanus*, whole cell. Lane 3: *T. africanus*, cell lysate supernatant.

Since Hfq has been shown to be involved in regulation of bacterial stress responses, cultures of *T. africanus* grown at optimal and lower and upper threshold temperatures (37, 70 and 77°C, respectively) were tested. The result of heat shock to 77°C is shown in Fig. 19, lanes 2 and 3. This gave a weaker signal than that observed for cells grown at optimal temperature. The same was observed for cells grown at 37°C (gel not shown).

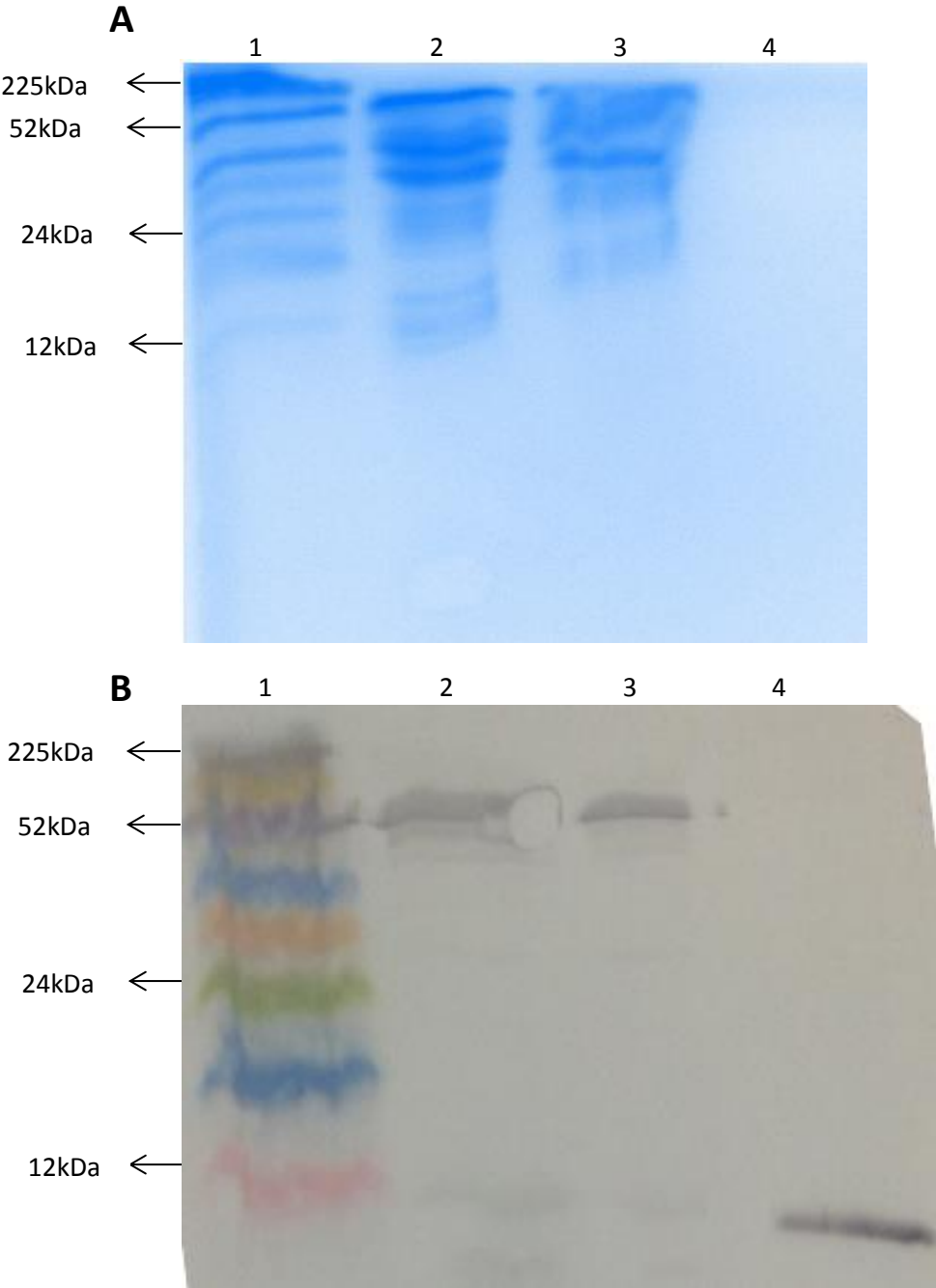


Figure 19: Western blot analysis of proteins from *T. africanus* cultures grown at upper threshold temperature (77°C). **A:** Coomassie stained SDS-PAGE prior to protein transfer. **B:** Western blot. Lane 1: Rainbow ladder. Lane 2 and 3: *T. africanus* heat shocked at 77°C. Lane 4: Control with *TaHfq* expressed in *E. coli*.

Despite an apparently low *hfq* expression in *T. africanus* and some problems with cross-reaction it was decided to proceed with immunoprecipitation experiments using anti-*TaHfq* antiserum and cell extracts from *T. africanus*. Fig. 20 shows a Western blot after pull-down of *TaHfq* using Sepharose A beads coated with anti-*TaHfq* antiserum (lanes 2 and 3) and including Sepharose A beads coated with pre-immune serum as a control (lane 4). An Hfq signal is clearly visible in lanes 2 and 3, but absent in lane 4. This demonstrates that the immunoprecipitation of Hfq works. However, both types of antisera resulted in significant cross-reactions with other proteins. A new pull-down experiment was performed in duplicates, where one of the membranes was only hybridized with the secondary conjugated antibodies. The results showed that the secondary antibody is the cause of the cross-reactions (Fig. 21), suggesting that the goat anti-rabbit IgG antiserum reacts with *T. africanus* antigens, but Hfq was not detected unless the primary antiserum was also used. Extracts from *K. olearia* and *P. mobilis* were also tested by Western blot, but the results were negative (gels not shown).

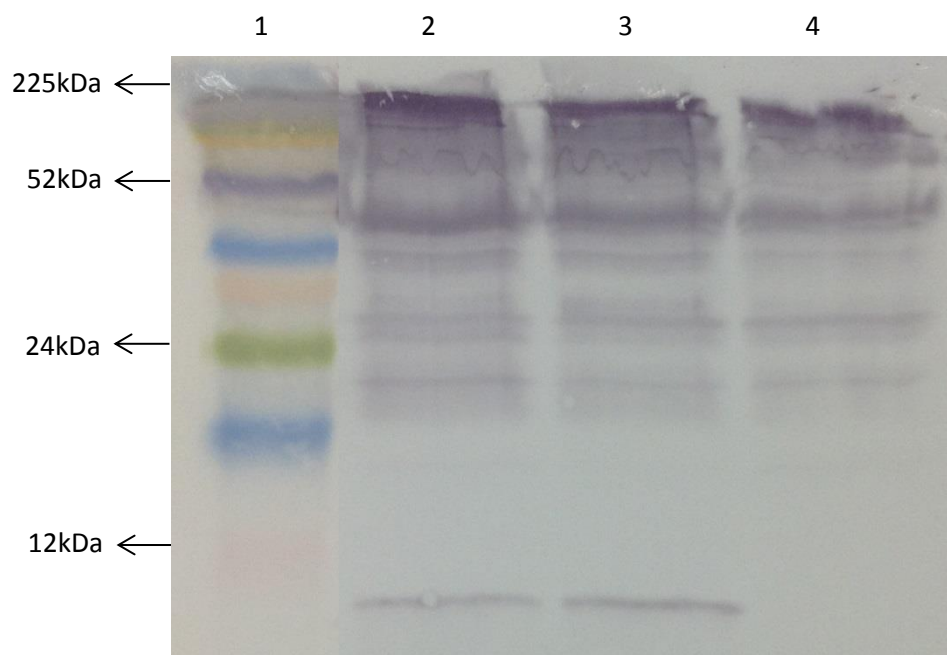


Figure 20: Western blot analysis of proteins from supernatant of lysed *T. africanus* cells immunoprecipitated with Sepharose A. Lane 1: Rainbow ladder. Lane 2 and 3: Immunoprecipitate from *T. africanus* lysate. Lane 4: Pre-immune serum precipitate from *T. africanus* lysate.

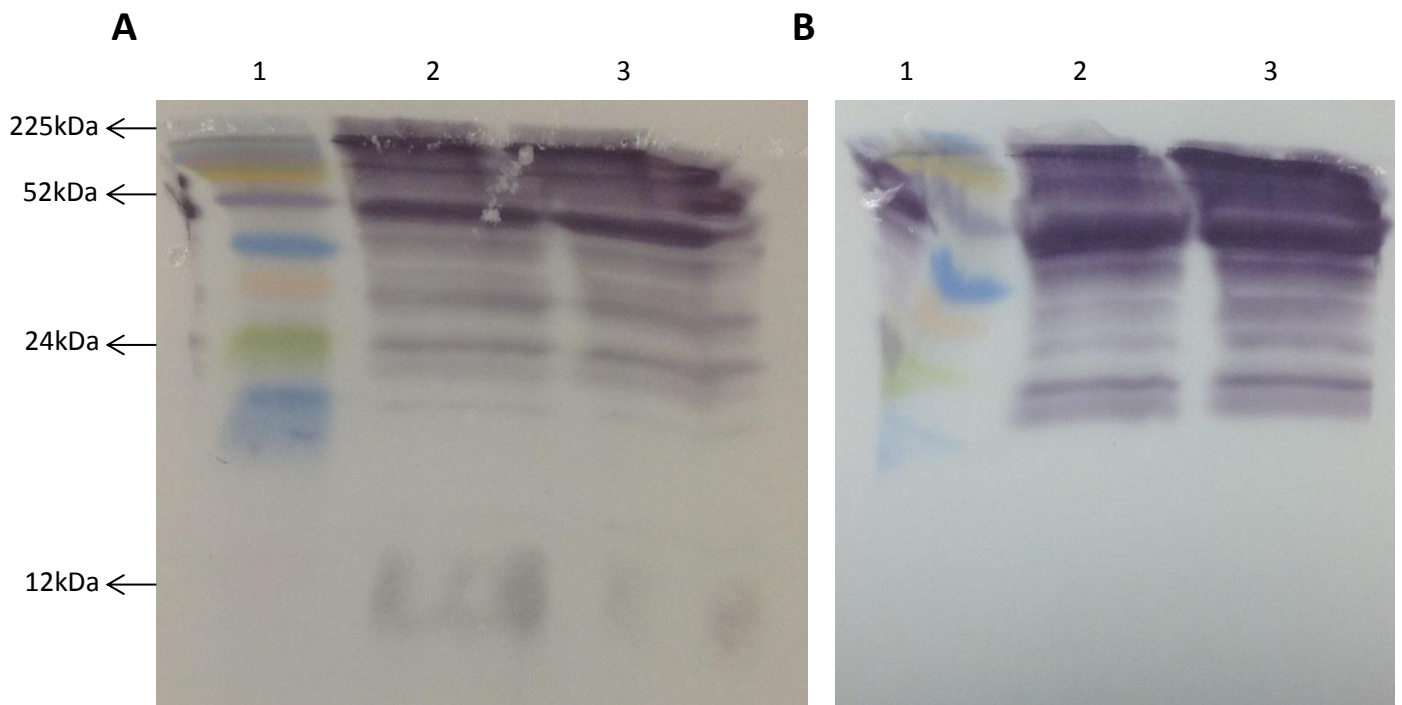


Figure 21: Western blot analysis for testing the cross reaction between *T. africanus* proteins and secondary antibody using a supernatant from lysed *T. africanus* cells immunoprecipitated with Sepharose A. **A)** Use of both primary and secondary antibody. **B)** Use of only secondary antibody. **A and B:** Lane 1: Rainbow ladder. Lane 2-3: *T. africanus* lysate precipitated with Sepharose A and Hfq antibody.

3.5 RNA work

3.5.1 Pull down of Hfq-RNA complexes and cDNA synthesis

Since the problems with cross-reaction was shown to be caused by the secondary antiserum, and the anti-*Ta*Hfq antiserum clearly detected *Ta*Hfq in crude *T. africanus* extracts as well as after immunoprecipitation, it was decided to proceed to co-immunoprecipitation experiments to pull out Hfq-binding RNAs from *T. africanus*. The secondary antiserum is not used for co-immunoprecipitation and should therefore not influence these experiments. Supernatant from lysed *T. africanus* cells was immunoprecipitated using Sepharose A beads coated with antiserum or pre-immune serum as a control. After incubation, the beads were spun down by centrifugation, hopefully in complex with antibodies, Hfq and Hfq-bound RNA. The pellet was washed, and a phenol extraction was done to release the binding between the beads, antibody, protein and RNA, and to purify the RNA. Following ethanol precipitation, purified RNA was analysed by agarose gel electrophoresis (Fig. 22A). Both samples (use of antiserum versus pre-immune serum) contained small RNAs in addition to a faint smear of larger RNAs. The RNA obtained by use of anti-*Ta*Hfq antiserum appeared to contain a larger amount of small RNAs, as would be expected if Hfq pull-down worked properly.

The RNAs were modified by the addition of a poly(A) tail to the 3' end and an RNA oligo to the 5' end, and reversely transcribed using the GeneRacer oligo dT primer and a Cloned AMV Reverse Transcriptase (section 2.8.5). Thereafter, the cDNA was PCR amplified and analysed by agarose gel electrophoresis (Fig. 22B). A strong band representing small cDNAs was observed for RNA pulled down with pre-immune serum (lanes 5 and 7), while a weaker band appeared for RNA pulled down with the anti-*Ta*Hfq antiserum (lanes 1 and 3). Also larger cDNA-PCR products were seen. The entire procedure was repeated to verify this result and the final cDNA-PCR product was analysed by electrophoresis using a 2% agarose gel (Fig. 23).

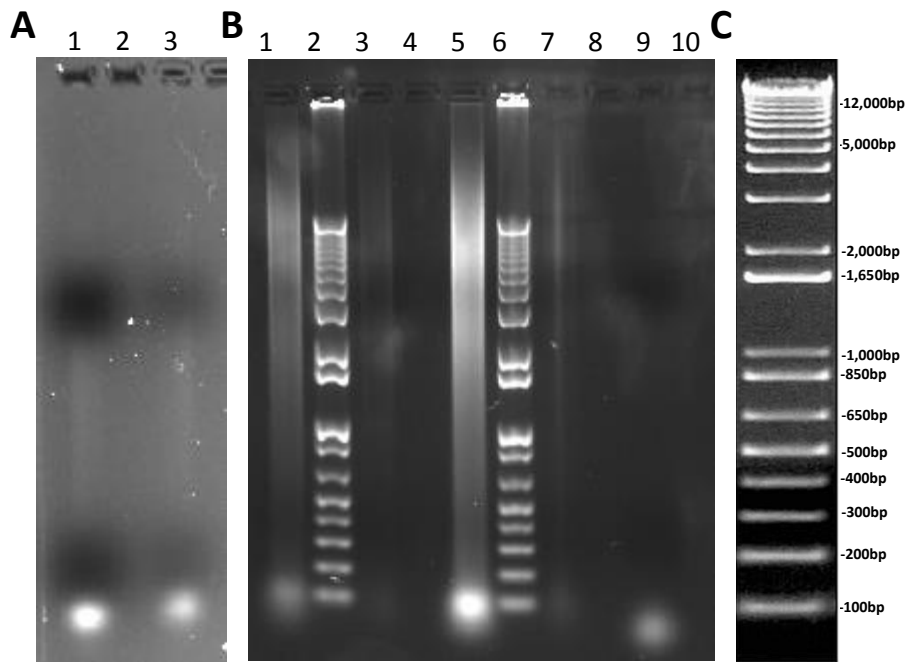


Figure 22: **A:** Agarose gel (1%) electrophoretic analysis of precipitated RNA. The gel was loaded with one μ l of the precipitated RNA. Lane 1: extract immunoprecipitated with antiserum. Lane 3: Extract immunoprecipitated with pre-immune serum as a control. **B:** Agarose gel (1%) electrophoretic analysis of amplified cDNA. Lane 1: cDNA from extract immunoprecipitated with antiserum. Lane 2: DNA Ladder. Lane 3: same as 1 but diluted 1:10. Lane 5: cDNA from extract immunoprecipitated with pre-immune serum. Lane 6: DNA Ladder. Lane 7: same as 5 but diluted 1:10. Lane 9: Control (not added template). **C:** DNA ladder.

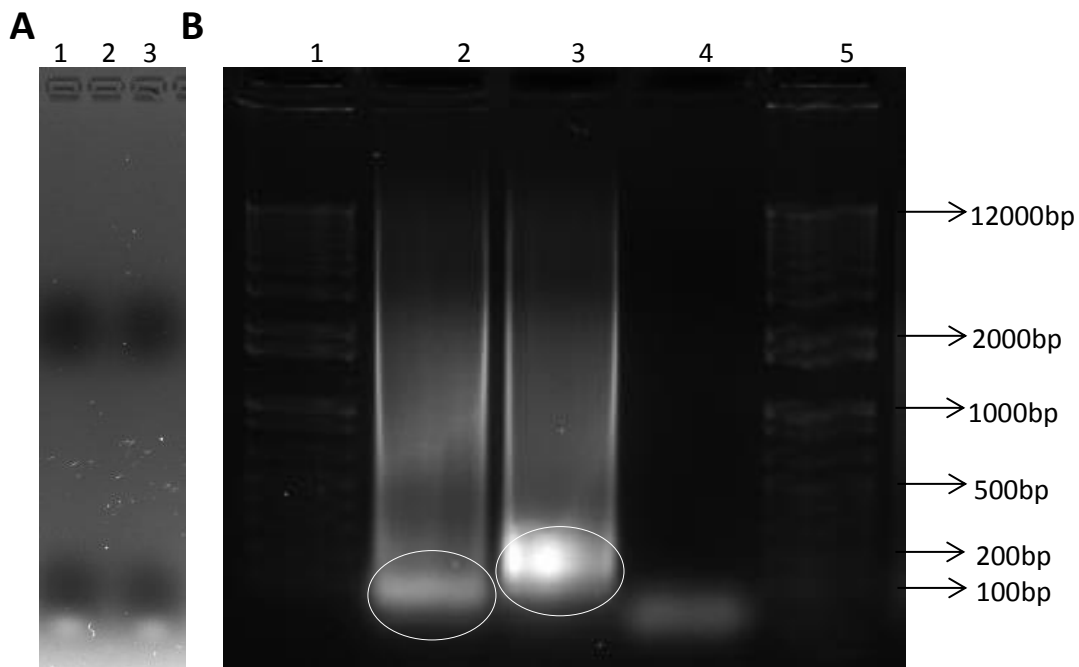


Figure 23: **A:** Agarose gel (1%) electrophoretic analysis of precipitated RNA. The gel was loaded with one μ l of the precipitated RNA. Lane 1: RNA sample immunoprecipitated with pre-immune serum. Lane 3: RNA sample immunoprecipitated with antiserum. **B:** Agarose gel (2%) electrophoretic analysis of amplified cDNA. The gel was loaded with five μ l of the cDNA. Lane 1: DNA Ladder. Lane 2: cDNA from extract immunoprecipitated with antiserum. Lane 3: cDNA from extract immunoprecipitated with pre-immune serum. Lane 4: Control PCR without template. Lane 5: DNA Ladder. The bands that were cut out of the gel and sequenced are encircled.

The second experiment gave a similar result: small PCR products of around 100 bp were found using both antiserum and pre-immune serum in the pull-down experiment (Fig. 22B and 23B). These products might represent amplicons of small RNAs, but because the control with pre-immune serum also gave a strong band it can be argued that the amplification is unspecific. The band in the control is, however, larger than for RNA pulled down with antiserum, which is difficult to explain but indicated that also the pre-immune serum has pulled down some RNAs that can function as a basis for cDNA synthesis. The smear of higher molecular weight products are probably a result of amplification of contaminating mRNAs.

3.6 Sequencing of cDNA

DNA from the encircled regions of the agarose gel shown in Fig. 23B was extracted, re-amplified by PCR and sequenced with the 5' oligo primer used for making the cDNA. This was done in order to reveal the nature of these products. The chromatograms from sequencing of these small cDNA fragments show that cDNA originating from the use of the pre-immune serum for pull down mostly contain poly(A) sequences (Fig. 24A) while the product from use of anti-Hfq antiserum contains a small poly(A) tail and a region yielding mixed sequence (Fig. 24B). The latter may represent cDNA originating from different sRNAs. No significant similarity was found when searching for similar sequences in BLAST.

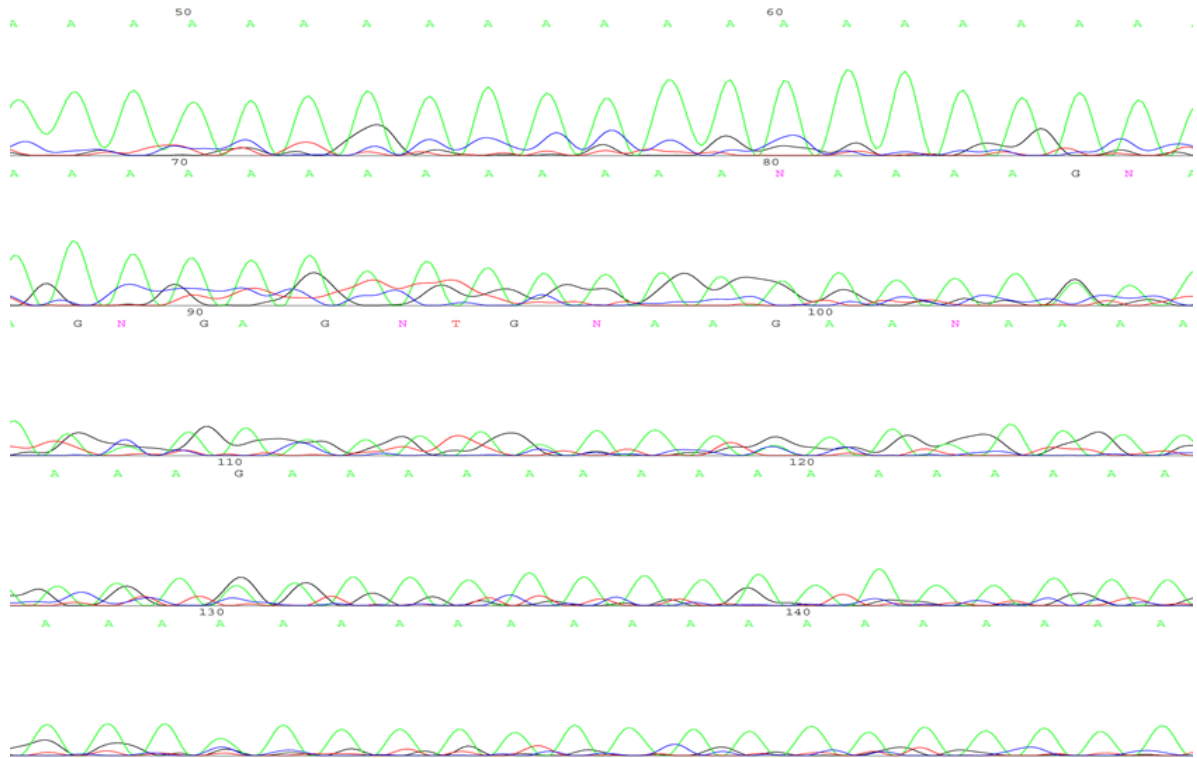
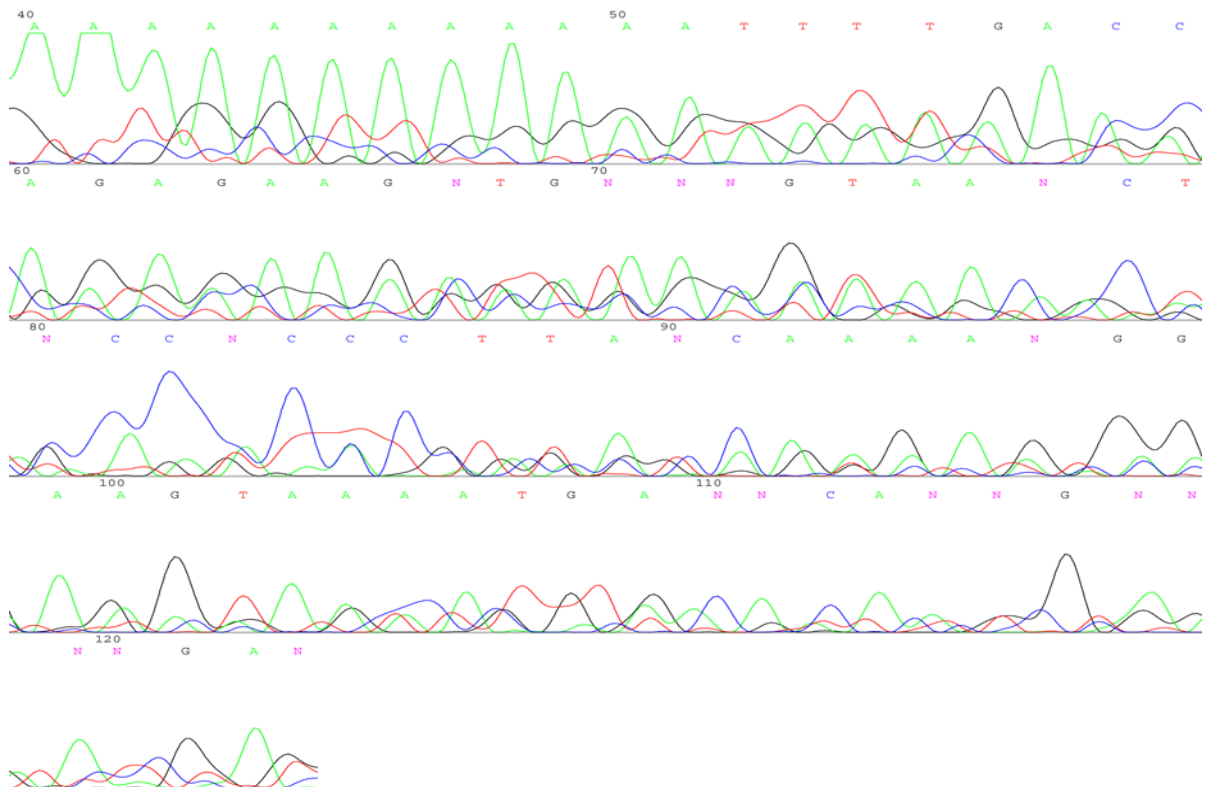
A**B**

Figure 24: A: A segment of the chromatogram from sequenced cDNA from RNAs precipitated with pre-immune serum. This sequence read mainly consisted of polyA. **B:** Chromatogram from sequenced cDNA from RNAs precipitated with antiserum. Bases 1-39 were removed as this sequence was unreadable.

To further characterize these products, the cDNA from the immune serum pull down was cloned into pCR4-TOPO and sequenced in both directions with vector primers M13rev and M13for. This should reveal sequences of isolated cDNA molecules. Analysis of 10 clones gave the same result, and showed that there was no cDNA between the polyA tail and the 5' oligonucleotide ligated to RNA (Fig. 25). The sequence of the 5' oligo and the dT polyA primer were identified on each side of the polyA/T region. It looks like polyadenylation has been carried out in the absence of an RNA 3' end, and that the 5' RNA oligonucleotide has been ligated directly onto the 5' end of polyA. There is no obvious explanation for this result, and due to time constraint it was not possible to repeat the experiment.

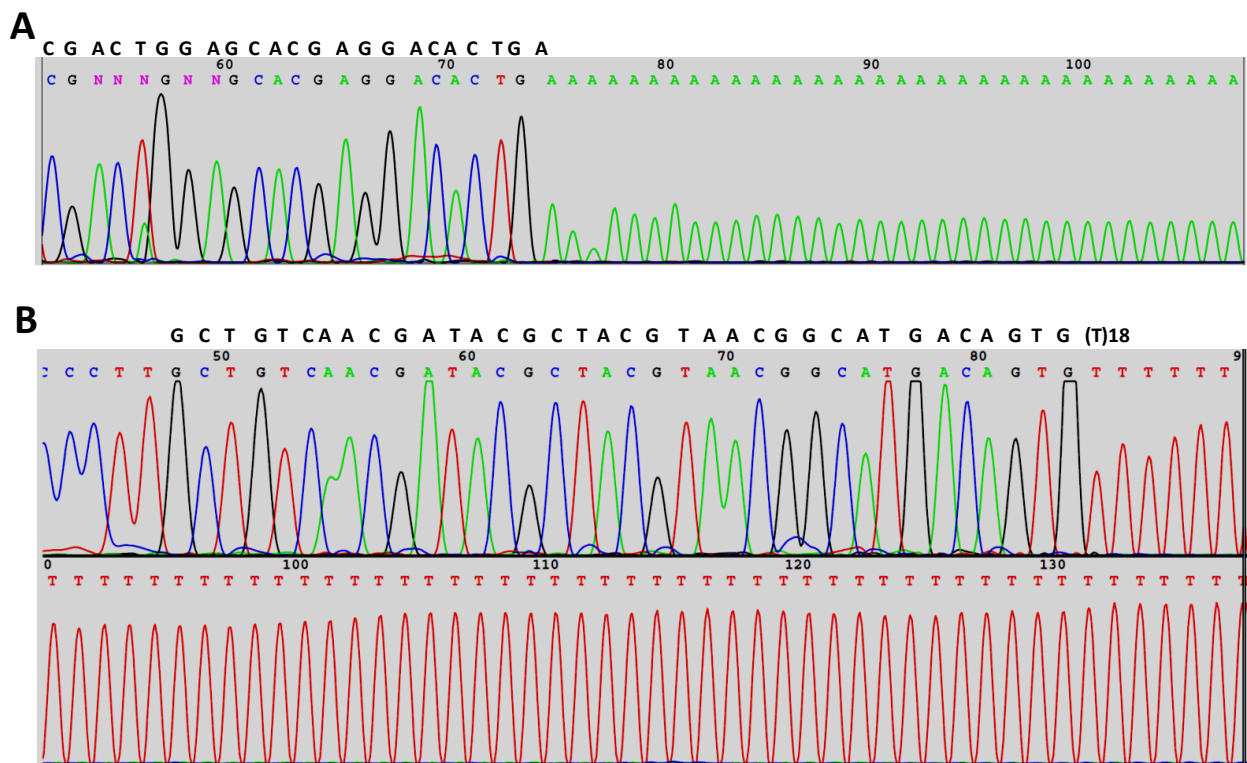


Figure 25: Chromatograms showing sections of the sequence reads of a cDNA clone sequenced with M13 forward primer (A) and M13 reverse primer (B). The GeneRacer 5' Primer and the GeneRacer Oligo dT Primer sequences are shown above (A) and (B), respectively, aligned with the corresponding sequence in the cloned cDNA product.

3.7 Proteomics

In order to verify expression of the Hfq protein in *T. africanus*, a crude cell extract was analysed using Orbitrap analysis. Out of several thousands of peptide spectra identified by Mascot search against the *T. africanus* genome, only one matched the Hfq protein (Appendix A, Fig. A.4), i.e. the amount of Hfq is just at the detection limit. This low expression level can explain the problems in obtaining cDNA representing sRNAs from *T. africanus*.

3.8 Bioinformatics

3.8.1 Phylogenetic tree based on Hfq and 16S rRNA

Juxtaposed phylogenetic trees based on multiple sequence alignments of 16S rRNA and Hfq reconstructs the evolutionary development of Thermotogales Hfq (Fig. 26). The branching order of the trees is mostly identical, except for the species *Thermotoga thermarum*.

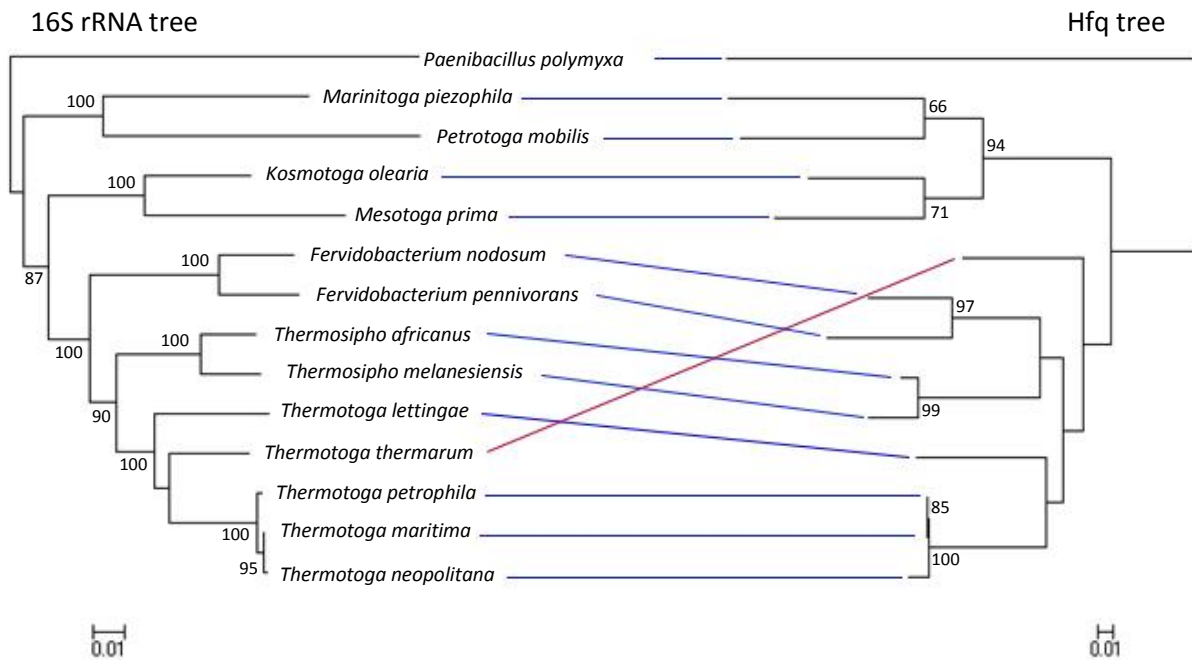


Figure 26: Phylogenetic tree based on 16S rRNA (left) and Hfq (right) sequences from the Thermotogae species. Multiple sequence alignment was made in ClustalX. The phylogenetic tree was made in Mega 6, and rooted with *Paenibacillus polymyxa*. The Neighbour Joining algorithm was used. The numbers at each node show the bootstrap value, which gives the frequency of this particular branching in a phylogenetic tree when repeating the process one hundred times. Bootstrap values below 66 are not included. The bars indicate the number of nucleotide (left) or amino acid (right) substitutions per site. A length of 0.01 represents 1% sequence change.

Accession number for the Hfq sequences: *Thermosipho africanus* YP_002333953, *Thermosipho melanesiensis* YP_001306984, *Thermotoga lettingae* YP_001470412, *Thermotoga maritima* YP_007976873, *Thermotoga petrophila* YP_001243997, *Thermotoga neapolitana* YP_002533686, *Fervidobacterium nodosum* YP_001409645, *Thermotoga thermarum* YP_004659805, *Fervidobacterium pennivorans* YP_005470618, *Kosmotoga olearia* YP_002940010, *Mesotoga prima* YP_006345778, *Marinitoga piezophila* YP_005096801, *Petrotoga mobilis* YP_001568408, *Paenibacillus polymyxa* YP_008912444. Accession numbers for the 16S rRNA sequences: *Thermosipho africanus* gi|217076133, *Thermotoga maritima* gi|568317838, *Thermotoga petrophila* gi|148269145, *Thermotoga neapolitana* gi|222098974, *Fervidobacterium pennivorans* gi|383785713, *Kosmotoga olearia* gi|239616411, *Mesotoga prima* gi|389842920, *Paenibacillus polymyxa* gi|386038690, *Thermosipho melanesiensis* gi|150019913, *Thermotoga lettingae* gi|157362870, *Fervidobacterium nodosum* gi|154248705, *Thermotoga thermarum* gi|338729812, *Marinitoga piezophila* gi|374339022, *Petrotoga mobilis* gi|160901491.

3.8.2 Multiple sequence alignment of Hfq from Thermotogales and model species

A multiple sequence alignment based on the secondary structure of Hfq emphasizes similarities and differences in the amino acid sequence of Hfq from Thermotogales and model species (Fig. 27). One hundred percent conserved residues are marked with an asterisk, and a red triangle marks amino acids involved in RNA-binding. The RNA-binding amino acids are 100% conserved across the Thermotogales.

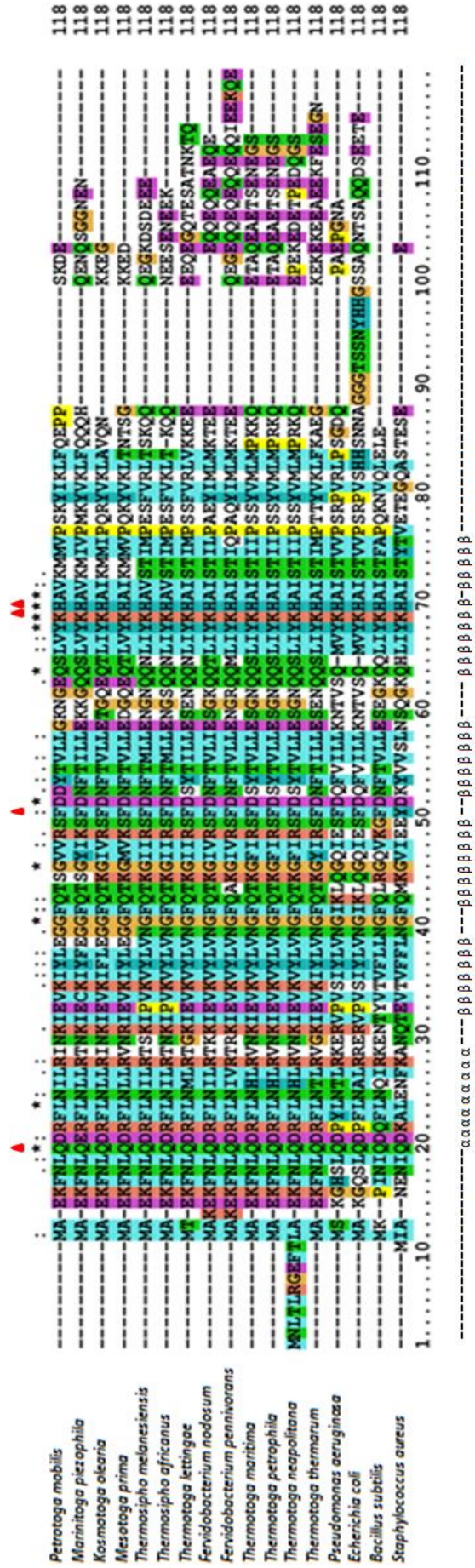


Figure 27: Multiple sequence alignment of Hfq from Thermotogales and model species, made in ClustalX. The species on the left corresponds to the sequence in the same line. The numbering at the bottom corresponds to the *Thermotoga neapolitana* sequence. Amino acids nr 20-29 makes the α -helix structure, and amino acids nr 33-39, 43-50, 53-59, 64-70 and 72-76 makes the five β sheets of one Hfq subunit (Sauter, Basquin et al. 2003). The colours represent different types of amino acids; light blue is non-polar aliphatic, pink is negatively charged, orange is positively charged, green is polar uncharged, yellow is hydrophobic. The asterisks on the top indicate that the amino acid is 100% conserved in the listed species. Two dots indicates a great degree of conservation, and one dot some degree of conservation. The red triangles marks amino acids which have been shown to be involved in RNA-binding in *S. aureus* (Schumacher, Pearson et al. 2002). The access numbers are listed below. *Thermosiphonia africana* YP_002333953, *Thermosiphonia melanesiensis* YP_001306984, *Thermotoga lettingae* YP_001470412, *Thermotoga maritima* YP_007976873, *Thermotoga petrophila* YP_001243997, *Thermotoga neapolitana* YP_002533686, *Fervidobacterium nodosum* YP_001409645, *Thermotoga thermarum* YP_004659805, *Fervidobacterium penninivorans* YP_005470618, *Kosmotoga olearia* YP_002940010, *Mesotoga prima* YP_006345778, *Marinitoga piezophila* YP_005096801, *Petrotoga mobilis* YP_001568408, *Pseudomonas aeruginosa* NP_253631.1, *Escherichia coli* YP_492314.1, *Bacillus subtilis* NP_389616, *Staphylococcus aureus* GI: 320140879.

3.8.3 Prediction of protein structure

The 3-dimensional structure of *TaHfq* was predicted using *Pseudomonas aeruginosa* Hfq as a template (Fig. 28). Tyrosine, Lysine and glutamine residues involved in RNA-binding, are conserved and shown as spheres. Fig. 29 shows the *Staphylococcus aureus* Hfq structure including a bound small RNA molecule (5'-AUUUUUG-3'). The RNA-binding site is highly conserved and the detailed uracil binding pocket is shown in Fig. 30.

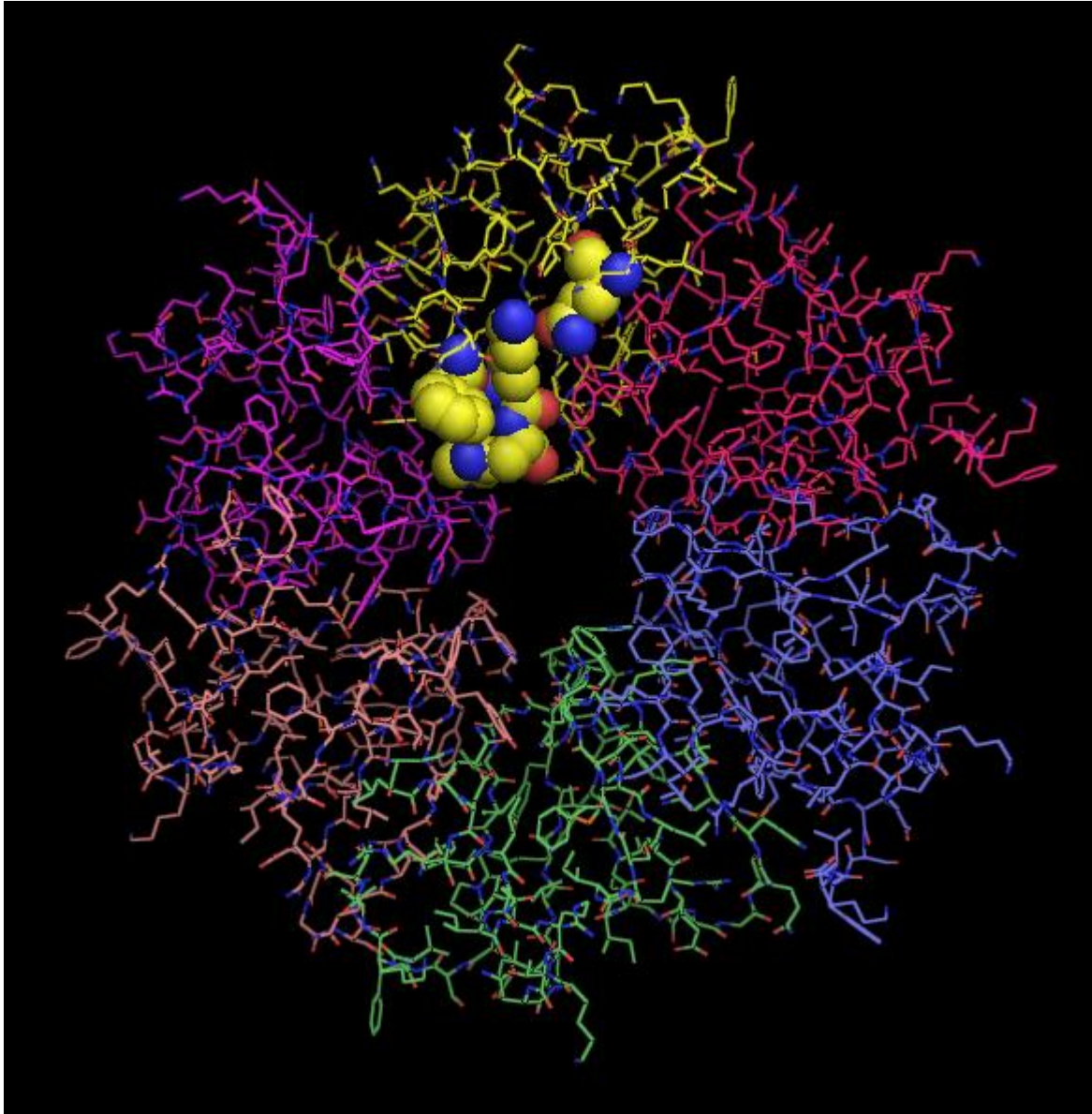


Figure 28: Prediction of the structure of *TaHfq* based on *Pseudomonas aeruginosa* Hfq (Accession number: 1u1S.1) which has been determined with a resolution of 1.60 Å (Nikulin, Stolboushina et al. 2005). Every subunit is given a different colour, and the amino acid residues binding to RNA are shown as spheres for one subunit where yellow represents carbon, red oxygen and blue nitrogen. The determination of the RNA-binding site was based on the RNA-binding sites that has been determined in *Staphylococcus aureus* Hfq (Schumacher, Pearson et al. 2002).

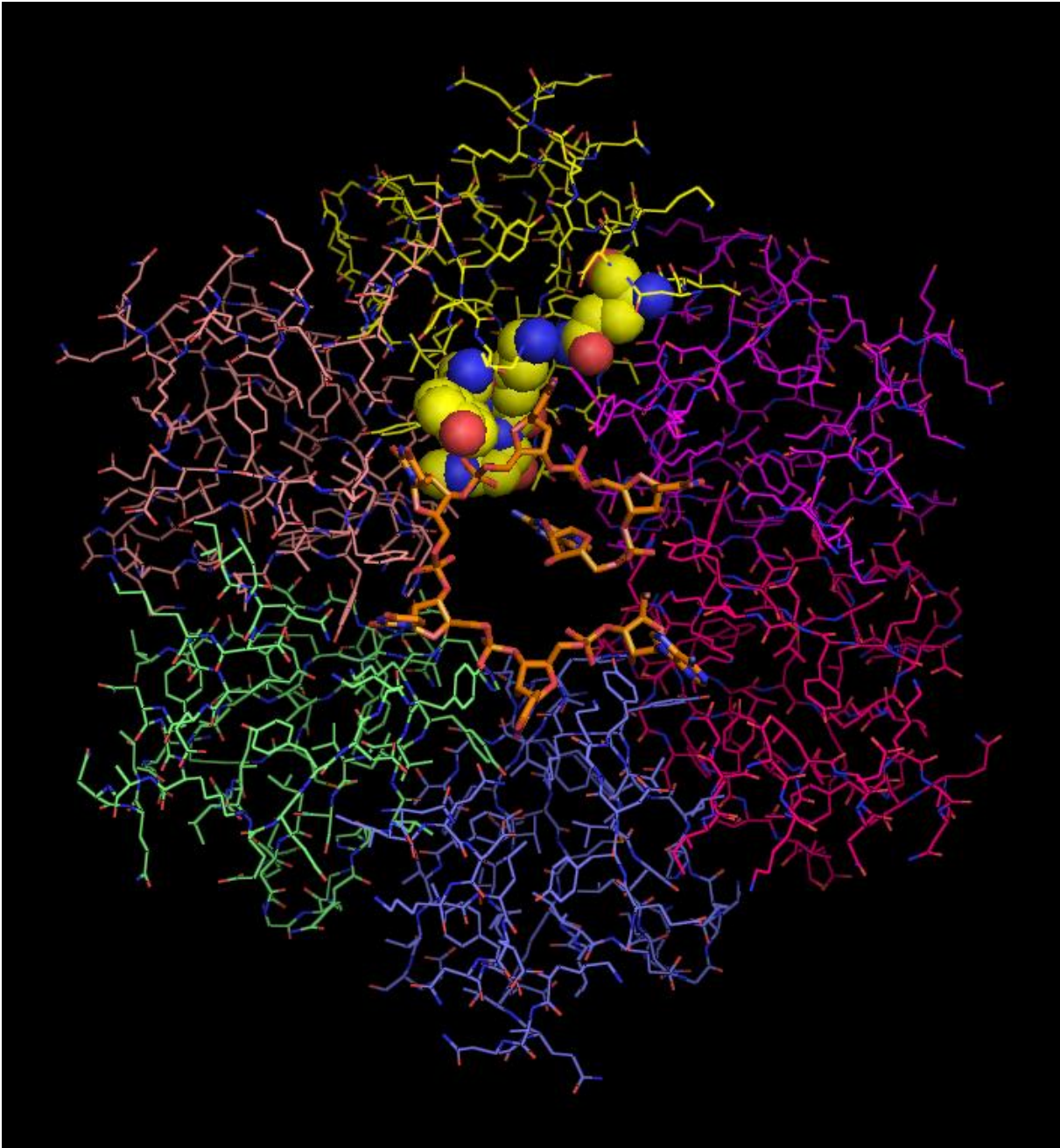


Figure 29: X-ray crystallographic structure of *Staphylococcus aureus* Hfq (PDB code: 1KQ2). The six subunits are given different colours. The amino acid residues binding to RNA has been determined by X-ray crystal structure (Schumacher, Pearson et al. 2002) and are shown in sphere format for one subunit, where yellow represents carbon, red oxygen and blue nitrogen. A bound RNA molecule (AU₅G) illustrating how binding can take place is shown in stick format.

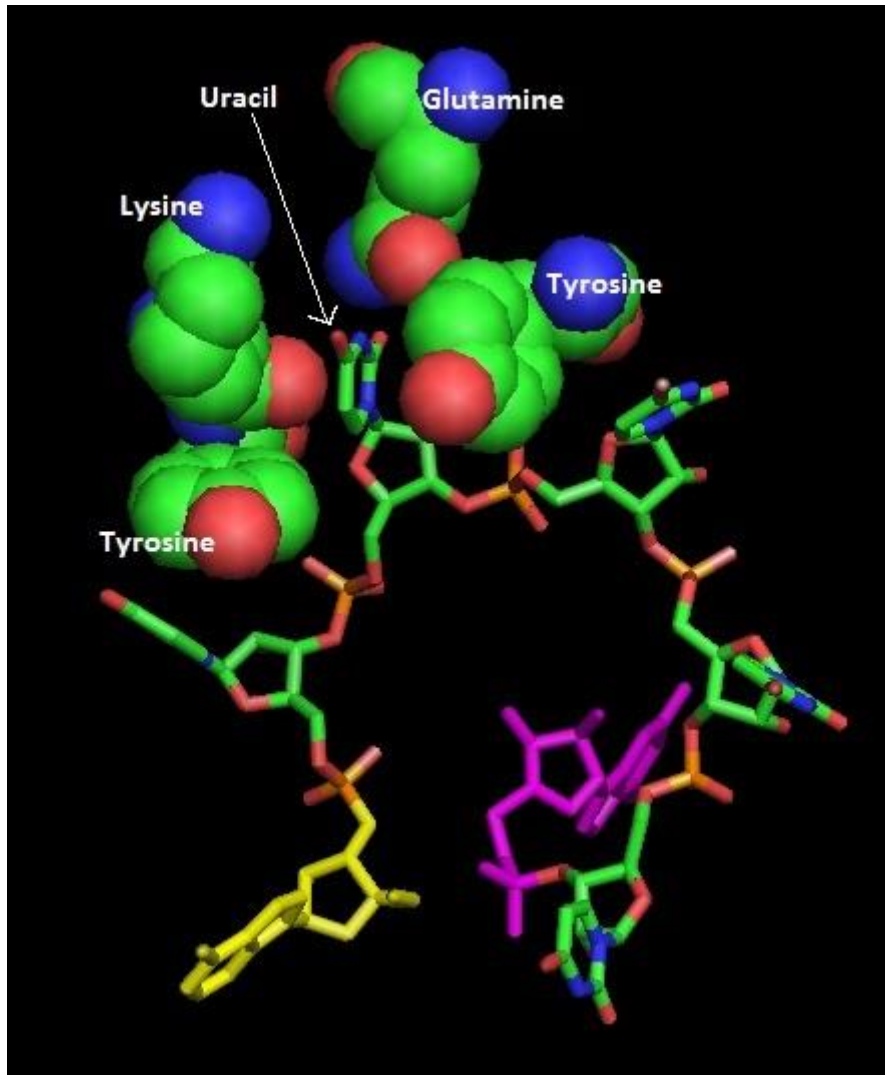


Figure 30: Structure of *Staphylococcus aureus* Hfq binding sites and a bound RNA molecule (AU₅G), as determined by X-ray crystal structure (Schumacher, Pearson et al. 2002). The amino acids involved in binding of one uracil are shown in sphere format. The bound RNA molecule (AU₅G) is shown in stick format where the yellow sticks represent the adenine base, purple is guanine and the green, red and blue sticks represents the uracils. In this model green represents carbon, blue is nitrogen, red is oxygen and orange is phosphate. Two subunits are involved in the binding of each base, tyrosine, lysine and glutamine from one subunit and tyrosine from the neighbouring subunit. The amino acids form a “pocket” where the base binds.

4. Discussion

Bacteria frequently use sRNAs as part of their mechanism positive and negative regulation of gene expression at the post-transcriptional level. Many of these sRNAs are dependent on the Hfq protein for their regulatory properties. This RNA chaperone is therefore a key component in the regulatory apparatus of many bacterial species. An increasing amount of studies have been made in the last couple of decades addressing Hfq-mediated regulation, but the detailed mechanisms of this regulation is still largely unknown. Studies have mostly been performed on pathogens and model species such as *E. coli*, *Salmonella enterica* and *S. aureus* (Vogel and Luisi 2011, Horstmann, Orans et al. 2012, Faner and Feig 2013). The fact that homologs of this protein have been found in Archaea and Eukarya suggests that a predecessor protein may have been present in the LUCA. Yet, no studies have been done on the role of Hfq in phylogenetically deeply branching bacteria. The aim of this study was to determine the role of Hfq in the phylogenetically deeply branching Thermotogales. Anti-TaHfq antiserum was successfully produced, and expression of *hfq* in *T. africanus* was confirmed by Western blotting. cDNA was made from an RNA-extract of co-immunoprecipitated RNA. However, sequencing of single low-molecular weight cDNA PCR products showed that these did not contain any cDNA derived from *T. africanus*, but only sequences related to oligonucleotides used for cDNA synthesis and polyA. It could not be determined whether Hfq is involved in gene regulation in *T. africanus*, and possible reasons for this are discussed below. The discussion is divided into two sections, with one part discussing the materials and methods used in this study, emphasizing what could have been done additionally or differently, and one part discussing the results.

4.1 Discussion of materials and methods

4.1.1 Gene expression in *E. coli*

P. mobilis hfq and *K. olearia hfq* were not successfully expressed in *E. coli*. This can be due to codon bias. Organisms belonging to different lineages often have different codon usage frequencies. It is possible that the codons commonly used in *E. coli* do not match the codons commonly used in Thermotogales species, as their phylogenetic relationship is very distant (Fig. 5). Codons that are used in *P. mobilis*, *K. olearia* and *T. africanus hfq* but are rarely

expressed in *E.coli* are listed in Table 16. A codon usage bias may have prevented efficient translation of *hfq* in *E. coli*. It is also possible that this has affected the solubility of the protein. The *E. coli* strain, Rosetta (DE3) *plysS*, could have been used to avoid these problems as it contains and expresses rare *E. coli* codons on a multicopy plasmid. However, the observation that *TaHfq* was successfully expressed in *E. coli* does not support this theory.

Table 16: Codons used in *P. mobilis*, *K. olearia* and *T. africanus hfq* that are rarely used in *E.coli*. The codons were found using Emboss (<http://emboss.bioinformatics.nl>). Codons that are rarely expressed in *E. coli* were obtained from Novy, Drott et al. (2001).

	Codon	AA	Fraction	Frequency	Number
<i>P. mobilis hfq</i>	AGA	R	0.333	12.346	1
	AGG	R	0.333	12.346	1
	CGA	R	0.333	12.346	1
	GGA	G	0.600	37.037	3
<i>K. olearia hfq</i>	AGA	R	0.250	12.500	1
	AGG	R	0.250	12.500	1
	GGA	G	0.200	12.500	1
<i>T. africanus hfq</i>	AGA	R	1.000	35.294	3
	GGA	G	0.333	11.765	1

4.1.2 Protein solubility

Analysis of the solubility of the *TaHfq* protein expressed in recombinant *E. coli* resulted in the detection of the Hfq protein in the pellet sample (Fig. 16), thus the collected protein was considered insoluble. As a result it was not possible to extract RNA by incubating cloned and tagged Hfq with cell extract and pull out complexes using a His-tag. The formation of insoluble aggregates is a frequent problem when overexpressing recombinant proteins in *E. coli* (Hannig and Makrides 1998), and there are several possible explanations for this;

- a) It is possible that Hfq folds differently in *E. coli* at 37°C than in the original cells because of the difference in temperature.
- b) Proteins may also fold incorrectly if translation occurs too quickly. Expression of the *hfq* gene was performed at 37°C. Lowering the temperature before inducing gene expression might have prevented an incorrect folding of the protein as folding occurs more slowly at reduced temperatures. The result might have been soluble Hfq proteins which could have been used for extracting Hfq-RNA-complexes by His-tag.

- c) Another possibility is that the protein is dependent on a chaperone that exists in *T. africanus* for correct folding. Hfq was detected in the Western blot with supernatant from lysed *T. africanus* cells (lane 3, Fig. 18). From this it was determined that Hfq is a soluble cytoplasmic protein in *T. africanus*.
- d) Incorrect folding of Hfq could be prevented by designing a synthetic *hfq* gene with optimal codon use for *E. coli*, and use this for overexpression.
- e) Use of the pETM-40 vector was considered as this contains the maltose binding protein gene designed for making MalT fusion gene constructs, which has often been used for producing soluble proteins. The maltose fusion protein is, however, large (42.5 kDa) compared to Hfq, and it may have prevented the correct folding of Hfq into a hexameric protein.

No attempts for alternative approaches were made because the anti-Hfq antiserum was specific and use of the co-IP strategy seemed promising. Time constraints were also a limiting factor.

4.1.3 RNA-extraction

An RNA-extraction was done after co-IP. The analysis by agarose gel electrophoresis showed that RNA had been extracted from both the sample and the control (Fig. 22A and 23A). This indicates that an unwanted effect had occurred prior to the cDNA synthesizing. Several measures could have been made to improve the RNA-extraction;

- a) A second control containing RNA of known size should have been included. This would have confirmed a successful polyadenylation and ligation of oligo to the RNAs, and ruled out some potential errors in the experiment such as contamination.
- b) A second co-IP could have been performed with the extracted RNAs before synthesizing cDNA. This could have decreased the precipitation of unspecific RNAs.
- c) Polyclonal antibodies were used in this study. Use of monoclonal antibodies could possibly improve the specificity.

Due to time constraint it was not possible to repeat and improve the experiment.

4.2 Discussion of results

4.2.1 Exclusion of *K. olearia* and *P. mobilis*

K. olearia and *P. mobilis* were excluded from further analysis because no expression of *hfq* was detected after induction with IPTG (Fig. 15). However, as there is a weak band in the correct size range in the protein sample from the clone with the *P. mobilis hfq* gene insertion it is possible that it was weakly expressed (Fig. 15, lane 4). This should have been tested further.

4.2.2 Weak transfer of proteins

A complete transfer was always observed for the protein ladder and small proteins during Western blotting. Most of the larger *T. africanus* proteins did not transfer to the membrane, and were left in the gel (Fig. 18). A longer transfer period at lower voltage was assessed to enhance the transfer, but without a positive result. The cause of the weak transfer was not determined, but as the low-molecular weight proteins seemed to transfer effectively the experiments were continued.

4.2.3 Cross reaction with secondary antibody

A significant degree of cross reaction was detected between the secondary antibody (goat anti-rabbit IgG conjugated to HRP) and proteins from *T. africanus* (Fig. 18-20). First it was thought that this cross reaction might have been Hfq proteins that had not been sufficiently denatured, and hence got stuck further up in the gel. Longer heating time (x2) in the sample buffer on the heat block prior to SDS-PAGE was assessed, but this gave the same result. Another possible explanation is that goats carry a *T. africanus*-like microorganism in their intestines and therefore possess cross-reacting antibodies. This is considered unlikely, as *T. africanus* thrive in high-temperature environments on sea floor or in hot springs, far from goat habitats. It may also be closer relatives of *T. africanus*, such as mesophilic Thermotogales that is present in goat intestine. All the cross reaction seemed to be caused by the secondary antibody, and not the primary antibody (Fig. 21). Only the primary antibody was used when pulling out the Hfq protein, hence the cross reaction was not

expected to have an impact on the further analysis, and no further examinations were carried out. The different degree of cross reaction is most likely due to different amounts of sample loaded in the SDS-PAGE.

4.2.4 RNA extraction, cDNA synthesis and sequencing

Following co-IP, RNA was only expected to be present in the sample with anti-Hfq antibodies (as illustrated in Fig. 4B). However, analysis of the precipitation by agarose gel electrophoresis confirmed the presence of RNAs also in the control that had been immunoprecipitated with pre-immune serum (Fig. 22A). A repetition of the experiment gave the same result (Fig. 23A). cDNA was synthesized from RNA in both the sample and the control to further analyse this result, and an agarose gel electrophoresis analysis of this cDNA showed slightly different profiles for the two (Fig. 22B and 23B), with the control containing larger RNAs than the sample. Hfq have also been shown to bind mRNAs (Vogel and Luisi 2011), and some of the larger fragments are possibly mRNAs of different sizes. The smallest fragments (approx. 100 bp) were further analysed by sequencing because the size coincided with the size of sRNAs (commonly 50-300 bp), which are of most interest in this study. These small fragments also gave the strongest band in the gel. Chromatograms from direct sequencing of the sample and control PCR products gave very different profiles (Fig. 24). The control almost exclusively consists of a long stretch of polyA, and possibly some cDNA from very small RNAs (Fig. 24A). The sample that had been immunoprecipitated with antiserum only contained a short poly(A) tail at its 3'-termini, and mixed sequence signals dominated the rest of the sequence read, which possibly could represent cDNA from several small RNAs. Sequencing of single cDNAs after cloning in a plasmid vector showed that the 3' polyadenylation, 5' ligation of oligo and cDNA synthesis was successful, but seemingly without any RNA-template (Fig. 25). Plasmid clones carrying cDNA PCR products were sequenced in both directions. After reading through the poly-A and the complementary poly-T regions, the polymerase seems to have slipped during the extension, yielding totally unreadable signals (personal communication, D. Turcu, sequencing lab, University of Bergen).

4.2.5 Possible reasons for unsuccessful RNA precipitation

Bands were seen in the control sample after the RNA extraction (Fig. 22A and 23A). This indicates that an unexpected effect had occurred prior to cDNA synthesizing. There are several possible reasons for this;

- a) The discrepancy may have been caused by operator errors in the lab. Yet, the obtained RNAs are not likely to be contamination as no significant similarities to the RNAs were found in BLAST searches.
- b) RNA transcripts are short-lived, and thus may be difficult to detect in these kinds of experiments.
- c) Sepharose A beads bind IgG antibodies non-specifically (Fig. 4B). Polyclonal antibodies were used in this study, hence all IgG antibodies from rabbit antiserum binds to the beads and precipitate.
- d) Out of several thousand peptide spectra identified by search against the *T. africanus* genome, only one matched the Hfq protein (Appendix A, Fig. A.4), i.e. just at the detection limit. This shows that the *hfq* gene is very poorly expressed in *T. africanus* under standard growth conditions.

4.2.6 Bioinformatics

The phylogenetic tree based on 16S rRNA and Hfq for the Thermotogae species suggest a lateral gene transfer of Hfq to *Thermotoga lettingae* (Fig. 26). The branching order of the trees is mostly identical, except for *T. thermarum*. This indicates that *hfq* is mostly vertically transferred and not prone to lateral gene transfer as has frequently been observed among the Thermotogales (Nelson, Clayton et al. 1999, Nesbø, Baptiste et al. 2009). *T. thermarum*, however, seems to have received its *hfq* gene from a more deeply branching Thermotogae as indicated by the lack of congruence with the 16S rRNA tree for this particular species. The bootstrap value for the branching of *T. thermarum hfq* is, however, below 66, indicating weak support for its branching point.

The homology model of *TaHfq* (Fig. 28) indicated a similar 3D structure to Hfq structures determined by X-ray crystallography, including the presence of a similar RNA-binding pocket. The highly conserved motifs and structure of this protein suggest that it has a role in gene

expression in Thermotogales. Conservation of *hfq* in all the genome-sequenced Thermotogales representatives is also in support of an important functional role for the Hfq protein. The multiple sequence alignment based on the secondary structure of Hfq reveals extra c-terminal amino acids with little degree of conservation (Fig. 27). This small region might act as a protein stabilizing factor.

5. Conclusions

Hfq from *T. africanus* was expressed in *E. coli* but were insoluble. *Hfq* from the other strains were not expressed, but could perhaps have been successfully expressed by taking measures discussed in section 4.1.2 and 4.1.3.

Hfq was detected in *T. africanus* cell lysates by specific anti-*TaHfq* antibodies raised in rabbits, but apparently in very low amounts. *Hfq* have in previous studies been shown to regulate expression of genes related to stress-responses. *Hfq* was detected in *T. africanus* cell lysates from cells grown at threshold temperatures, but not in large amounts. Alternative stressful cultivating conditions might have resulted in increased detection of *Hfq*.

The secondary antiserum (goat anti-rabbit) used in Western blot experiments cross-reacted strongly with *T. africanus*, causing problems in interpreting the blots. This was supported by the Ouchterlony double immunodiffusion experiment where a precipitation line was made between *T. africanus* cell lysate and the secondary antibody (Appendix A, Fig. A.3). This confirmed a cross-reaction between *T. africanus* proteins and the secondary antibody (Goat anti-rabbit).

Attempts to purify and make cDNA of sRNAs interacting *in vivo* with *TaHfq* were unsuccessful. This may be a result of technical problems in RNA capture or cDNA synthesis. Based on proteomics analysis of a *T. africanus* cell extract the expression level seems to be extremely low, and this makes the *Hfq*-RNA-complex more difficult to detect.

A homology model of *TaHfq* indicated a similar 3D structure as *Hfq* structures determined by X-ray crystallography, including the presence of a similar RNA-binding pocket (Fig. 28). A multiple sequence alignment comparing Thermotogales *Hfq* to model species showed a considerable deal of conservation in the amino acid sequences (Fig. 27). This conservation of both amino acid sequence and protein structure suggests that the protein is functional.

Whilst no definitive conclusions regarding the function of *Hfq* in Thermotogales could be drawn from the study, results may be indicative of a low basal expression of *hfq*, and further work should include cultivation of Thermotogales under stressful conditions, and attempts to solubilize the recombinant *Hfq*. This study may provide a valuable basis for further studies by identifying various methodological challenges.

6. Further work

- a) Threshold conditions such as amino acid starvation, high/low pH, NaCl or oxygen content during cultivation might have led to a greater expression of Hfq, assuming that Hfq is involved in gene regulation of stress-response.
- b) Expression of *Thermotogales hfq* in Rosetta strain to increase expression and solubility of Hfq proteins.
- c) Pyrosequencing of cDNA following co-IP to allow sequencing of a large number of fragments.
- d) Providing monoclonal antibodies (only anti-Hfq) to improve the specificity of the co-IP.
- e) RT-qPCR analysis to determine the expression level of *hfq* under various growth conditions.
- f) Complete transcriptomics analysis of *T. africanus* to identify and quantify all transcripts including *hfq* and sRNAs.
- g) Enrichment of sRNA by PAGE prior to Hfq binding to increase the amount of Hfq-interacting sRNAs in the sample, thereby lowering unspecific binding of Hfq.
- h) Expression in *T. africanus* of recombinant His-tagged Hfq, and pull out His-tagged Hfq-RNA complexes by nickel columns.
- i) Construction of a *hfq* knock-out mutant to study the physiological effects and global transcription patterns. This will require development of a genetic tool for *T. africanus* based on e.g. a suicide plasmid.

7. References

- Aiba, H. "Mechanism of Rna Silencing by Hfq-Binding Small RNAs." *Current Opinion in Microbiology* 10, no. 2 (2007): 134-139.
- Ausubel, M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith and K. Struhl. *Current Protocols in Molecular Biology*, Vol.1-2. John Wiley and Sons, Inc., Media PA, 1988.
- Ben Hania, W., A. Postec, T. Aullo, A. Ranchou-Peyruse, G. Erauso, C. Brochier-Armanet, M. Hamdi, B. Ollivier, S. Saint-Laurent, M. Magot and M. L. Fardeau. "*Mesotoga infera* Sp. Nov., a Mesophilic Member of the Order Thermotogales, Isolated from an Underground Gas Storage Aquifer." *International journal of systematic and evolutionary microbiology* 63, no. Pt 8 (2013): 3003-8.
- Bhandari, V. and R. S. Gupta. "Molecular Signatures for the Phylum (Class) Thermotogae and a Proposal for Its Division into Three Orders (Thermotogales, Kosmotogales Ord. Nov and Petrotogales Ord. Nov.) Containing Four Families (Thermotogaceae, Fervidobacteriaceae Fam. Nov., Kosmotogaceae Fam. Nov and Petrotogaceae Fam. Nov.) and a New Genus *Pseudothermotoga* Gen. Nov with Five New Combinations." *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology* 105, no. 1 (2013): 143-168.
- Caron, M. P., D. A. Lafontaine and E. Masse. "Small RNA-Mediated Regulation at the Level of Transcript Stability." *RNA Biology* 7, no. 2 (2010): 140-144.
- Chao, Y. J. and J. Vogel. "The Role of Hfq in Bacterial Pathogens." *Current Opinion in Microbiology* 13, no. 1 (2010): 24-33.
- Davey, M. E., W. A. Wood, R. Key, K. Nakamura and D. Stahl. "Isolation of Three Species of *Geotoga* and *Petrotoga*: Two New Genera, Representing a New Lineage in the Bacterial Line of Descent Distantly Related to the "Thermotogales"." *International Journal of Systematic Bacteriology* 16, (1993): 191-200.
- Davey, M. E., W. A. Wood, R. Key, K. Nakamura and D. Stahl. "Isolation of Three Species of *Geotoga* and *Petrotoga*: Two New Genera, Representing a New Lineage in the Bacterial Line of Descent Distantly Related to the "Thermotogales"." *Systematic and Applied Microbiology* 16, (1993): 191-200.
- DeLano, W.L. "The PyMol Molecular Graphics System" (2002) DeLano Scientific, San Carlos, CA, USA. <http://www.PyMOL.org>
- Dipippo, J. L., C. L. Nesbø, H. Dahle, W. F. Doolittle, N. K. Birkland and K. M. Noll. "*Kosmotoga olearia* Gen. Nov., Sp. Nov., a Thermophilic, Anaerobic Heterotroph Isolated from an Oil Production Fluid." *International journal of systematic and evolutionary microbiology* 59, no. 12 (2009): 2991-3000.

- Faner, M. A. and A. L. Feig. "Identifying and Characterizing Hfq-RNA Interactions." *Methods* 63, no. 2 (2013): 144-59.
- Franze de Fernandez, M. T., L. Eoyang and J. T. August. "Factor Fraction Required for the Synthesis of Bacteriophage Qbeta-RNA." *Nature* 219, no. 5154 (1968): 588-90.
- Hammer, B. K. and B. L. Bassler. "Regulatory Small RNAs Circumvent the Conventional Quorum Sensing Pathway in Pandemic *Vibrio cholerae*." *Proceedings of the National Academy of Sciences of the United States of America* 104, no. 27 (2007): 11145-9.
- Hania, B. W., R. Godbane, A. Postec, M. Hamdi, B. Ollivier and M. L. Fardeau. "*Defluviitoga tunisiensis* Gen. Nov, Sp. Nov., a Thermophilic Bacterium Isolated from a Mesothermic and Anaerobic Whey Digester." *International Journal of Systematic Evolutionary Microbiology* 62, (2012): 1377-1382.
- Hannig, G. and S. C. Makrides. "Strategies for Optimizing Heterologous Protein Expression in *Escherichia coli*." *Trends in Biotechnology* 16, no. 2 (1998): 54-60.
- Horstmann, N., J. Orans, P. Valentin-Hansen, S. A. Shelburne, 3rd and R. G. Brennan. "Structural Mechanism of *Staphylococcus aureus* Hfq Binding to an RNA a-Tract." *Nucleic acids research* 40, no. 21 (2012): 11023-35.
- Huber, R. and M. Hannig. "Thermotogales." In *The Prokaryotes*, edited by M. Dworkin, S. Falkow, E. Rosenberg, K. Schleifer and E. Stackebrandt, 899-922: Springer New York, 2006.
- Huber, R., T. A. Langworthy, H. König, M. Tkomm, C. R. Woese, U. B. Sleytr and K. O. Stetter. "*Thermotoga maritima* Sp. Nov. Represents a New Genus of Unique Extremely Thermophilic Eubacteria Growing up to 90°C." *Archives of Microbiology* 144, (1986): 324-333.
- Huber, R., C. R. Woese, T. A. Langworthy, H. Fricke and K. O. Stetter. "*Thermosiphon africanus* Gen. Nov., Represents a New Genus of Thermophilic Eubacteria within the "Thermotogales". " *Systematic and Applied Microbiology* 12, no. 1 (1989): 32-37.
- Hungate, R. E. "The Anaerobic Mesophilic Cellulolytic Bacteria." *Bacteriological Reviews* 14, no. 1 (1950): 1-49.
- Jayasinghearachchi, H. S. and B. Lal. "*Oceanotoga teriensis* Gen. Nov., Sp. Nov., a Thermophilic Bacterium Isolated from Offshore Oil-Producing Wells." *International Journal of Systematic and Evolutionary Microbiology* 61, (2011): 554-560.
- Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson and D. G. Higgins. "Clustal W and Clustal X Version 2.0." *Bioinformatics* 23, no. 21 (2007): 2947-2948.

- Lien, T., M. Madsen, F. A. Rainey and Birkeland N. K. "*Petrotoga mobilis* Sp. Nov., from a North Sea Oil-Production Well." *International Journal of Systemic Bacteriology* 48, (1998): 1007-1013.
- Link, T. M., P. Valentin-Hansen and R. G. Brennan. "Structure of *Escherichia coli* Hfq Bound to Polyriboadenylate RNA." *Proceedings of the National Academy of Sciences of the United States of America* 106, no. 46 (2009): 19286-19291.
- Liu, J. M. and A. Camilli. "A Broadening World of Bacterial Small RNAs." *Current Opinion in Microbiology* 13, no. 1 (2010): 18-23.
- Liu, Y., N. Wu, J. Dong, Y. P. Gao, X. Zhang, C. H. Mu, N. S. Shao and G. A. Yang. "Hfq Is a Global Regulator That Controls the Pathogenicity of *Staphylococcus aureus*." *Plos One* 5, no. 9 (2010).
- Madigan, M. T., J. M. Martinko, D. Stahl and D. P. Clark. *Brock Biology of Microorganisms*. Vol. 13. USA: Pearson Education, 2012.
- Murina, V. N. and A. D. Nikulin. "RNA-Binding Sm-Like Proteins of Bacteria and Archaea. Similarity and Difference in Structure and Function." *Biochemistry (Mosc)* 76, no. 13 (2011): 1434-49.
- Nelson, David L., Michael M. Cox and Albert L. Lehninger. *Lehninger Principles of Biochemistry*. 5th ed. New York: Freeman, 2008.
- Nelson, K. E., R. A. Clayton, S. R. Gill, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, J. D. Peterson, W. C. Nelson, K. A. Ketchum, L. McDonald, T. R. Utterback, J. A. Malek, K. D. Linher, M. M. Garrett, A. M. Stewart, M. D. Cotton, M. S. Pratt, C. A. Phillips, D. Richardson, J. Heidelberg, G. G. Sutton, R. D. Fleischmann, J. A. Eisen, O. White, S. L. Salzberg, H. O. Smith, J. C. Venter and C. M. Fraser. "Evidence for Lateral Gene Transfer between Archaea and Bacteria from Genome Sequence of *Thermotoga maritima*." *Nature* 399, (1999): 323-329.
- Nesbø, C. L., E. Baptiste, B. Curtis, H. Dahle, P. Lopez, D. Macleod, M. Dlutek, S. Bowman, O. Zhaxybayeva, N. K. Birkeland and W. F. Doolittle. "The Genome of *Thermosiphon africanus* Tcf52b: Lateral Genetic Connections to the Firmicutes and Archaea." *Journal of bacteriology* 191, no. 6 (2009): 1974-8.
- Nesbø, C. L., R. Kumaraswamy, K. Dlutek, F. W. Doolittle and J. Foght. "Searching for Mesophilic Thermotogales Bacteria: "Mesotogas" in the Wild." *Applied and Environmental Microbiology* 76, no. 14 (2010): 4896-4900.
- Nesbø, C.L., D.M. Bradnan, A. Adebisuyi, M. Dlutek, A. K. Petrus, J. Foght, W. F. Doolittle and K.M. Noll. "*Mesotoga prima* Gen. Nov., Sp. Nov., the First Described Mesophilic Species of the Thermotogales. Extremophiles." *International Journal of Systematic and Evolutionary Microbiology* 16, (2012): 387-393.

- Nikulin, A., E. Stolboushkina, A. Perederina, I. Vassilieva, U. Blaesi, I. Moll, G. Kachalova, S. Yokoyama, D. Vassilyev, M. Garber and S. Nikonov. "Structure of *Pseudomonas aeruginosa* Hfq Protein." *Acta crystallographica. Section D, Biological crystallography* 61, no. Pt 2 (2005): 141-6.
- Notman, D. D., N. Kurata and E. M. Tan. "Profiles of Antinuclear Antibodies in Systemic Rheumatic Diseases." *Annals of Internal Medicine* 83, no. 4 (1975): 464-469.
- Novy, R., D. Drott, K. Yaeger and R. Mierendorf. "Overcoming the Codon Bias of *E. coli* for Enhanced Protein Expression." *inNovations*, no. 12 (2001): 1-3.
- Ouchterlony, O. "Diffusion-in-Gel Methods for Immunological Analysis." *Prog Allergy* 5, (1958): 1-78.
- Patel, B. K. C., H. W. Morgan and R. M. Daniel. "*Ferrihydrobacterium nodosum* Gen. Nov. And Spec. Nov., a New Chemoorganotrophic, Caldoactive, Anaerobic Bacterium." *Archives of Microbiology* 141, no. 1 (1985): 63-69.
- Peng, Y., T. J. Soper and S. A. Woodson. "Positional Effects of Aan Motifs in Rpos Regulation by sRNAs and Hfq." *Journal of Molecular Biology* 426, no. 2 (2014): 275-285.
- Ravot, G., B. Olliver, B. K. Patel, M. Magot and J. Garcia. "Emended Description of *Thermosiphon africanus* as a Carbohydrate-Fermenting Species Using Thiosulfate as an Electron Acceptor." *International Journal of Systematic and Evolutionary Bacteriology* 46, no. 1 (1996): 321-323.
- Reysenbach, A. L., Y. Liu, A. R. Lindgren, I. D. Wagner, C. D. Sislak, A. Mets and S. Schouten. "*Mesoaciditoga lauensis* Gen. Nov., Sp. Nov., a Moderately Thermoacidophilic Member of the Order Thermotogales from a Deep-Sea Hydrothermal Vent." *International Journal of Systematic and Evolutionary Microbiology* 63, (2013): 4724-4729.
- Sauter, C., J. Basquin and D. Suck. "Sm-Like Proteins in Eubacteria: The Crystal Structure of the Hfq Protein from *Escherichia coli*." *Nucleic acids research* 31, no. 14 (2003): 4091-8.
- Schiano, C. A., L. E. Bellows and W. W. Lathem. "The Small RNA Chaperone Hfq Is Required for the Virulence of *Yersinia pseudotuberculosis*." *Infection and Immunity* 78, no. 5 (2010): 2034-2044.
- Schumacher, M. A., R. F. Pearson, T. Moller, P. Valentin-Hansen and R. G. Brennan. "Structures of the Pleiotropic Translational Regulator Hfq and an Hfq-RNA Complex: A Bacterial Sm-Like Protein." *The EMBO journal* 21, no. 13 (2002): 3546-56.
- Sobrero, P. and C. Valverde. "The Bacterial Protein Hfq: Much More Than a Mere RNA-Binding Factor." *Critical Reviews in Microbiology* 38, no. 4 (2012): 276-299.

- Soper, T., P. Mandin, N. Majdalani, S. Gottesman and S. A. Woodson. "Positive Regulation by Small RNAs and the Role of Hfq." *Proceedings of the National Academy of Sciences of the United States of America* 107, no. 21 (2010): 9602-9607.
- Storz, G. and D. Haas. "A Guide to Small RNAs in Microorganisms." *Current Opinion in Microbiology* 10, no. 2 (2007): 93-95.
- Storz, G., J. Vogel and K. M. Wassarman. "Regulation by Small RNAs in Bacteria: Expanding Frontiers." *Molecular cell* 43, no. 6 (2011): 880-91.
- Tamura, Koichiro, Glen Stecher, Daniel Peterson, Alan Filipski and Sudhir Kumar. "Mega6: Molecular Evolutionary Genetics Analysis Version 6.0." *Molecular Biology and Evolution*, (2013).
- Urban, J. H. and J. Vogel. "Two Seemingly Homologous Noncoding RNAs Act Hierarchically to Activate Glms mRNA Translation." *PLoS Biology* 6, no. 3 (2008): 64.
- Valentin-Hansen, P., M. Eriksen and C. Udesen. "The Bacterial Sm-Like Protein Hfq: A Key Player in RNA Transactions." *Molecular Microbiology* 51, no. 6 (2004): 1525-1533.
- Valverde, Patricio Sobrero and Claudio. "The Bacterial Protein Hfq: More Than a Mere RNA-Binding Factor." *Critical Reviews in Microbiology* 38, no. 4 (2012): 276-299.
- Vogel, J. and B. F. Luisi. "Hfq and Its Constellation of RNA." *Nature Reviews Microbiology* 9, no. 8 (2011): 578-589.
- Waters, L. S. and G. Storz. "Regulatory RNAs in Bacteria." *Cell* 136, no. 4 (2009): 615-28.
- Wery, N., F. Lesongeur, P. Pignet, V. Derennes, M. A. Cambon-Bonavita, A. Godfroy and G. Barbier. "*Marinitoga camini* Gen. Nov., Sp. Nov., a Rod-Shaped Bacterium Belonging to the Order Thermotogales, Isolated from a Deep-Sea Hydrothermal Vent." *International Journal of Systematic and Evolutionary Microbiology* 51, (2001): 495-504.
- Widdel, F., G.-W. Kohring and F. Mayer. "Studies on Dissimilatory Sulfate-Reducing Bacteria That Decompose Fatty Acids 3. Characterization of the Filamentous Gliding *Desulfonema limicola* Gen. Nov. Sp. Nov., and *Desulfonema magnum* Sp. Nov." *Archives of Microbiology* 134, no. 4 (1983): 286-294.
- Widdel, F. and N. Pfenning. "Studies on Dissimilatory Sulfate-Reducing Bacteria That Decompose Fatty Acids. 1. Isolation of New Sulfate-Reducing Bacteria Enriched with Acetate from Saline Environments. Description of *Desulfobacter postgatei* Gen. Nov., Sp. Nov." *Archive of Microbiology* 129, no. 5 (1981): 395-400.

8. Appendices

8.1 Appendix A: Results

Table A.1: NanoDrop measurements of genomic DNA extracted from *K. olearia*, *T. africanus* and *P. mobilis*. OD_{260}/OD_{280} is the ratio of the optical density of the samples at 260 nm and 280 nm.

	ng/ μ l	OD_{260}/OD_{280}
<i>K. olearia</i>	57	1.78
<i>T. africanus</i>	180.1	1.76
<i>P. mobilis</i>	64.6	1.85

Table A.2: NanoDrop of gene amplification from *K. olearia*, *T. africanus* and *P. mobilis*.

	ng/ μ l	OD_{260}/OD_{280}
<i>K. olearia</i>	40.4	1.77
<i>T. africanus</i>	43.6	1.81
<i>P. mobilis</i>	27.8	1.73

Table A.3: NanoDrop of the cut *hfq* gene from *K. olearia*, *T. africanus* and *P. mobilis* and vector.

	ng/ μ l	OD_{260}/OD_{280}
<i>K. olearia</i>	12.9	1.45
<i>T. africanus</i>	17.7	1.75
<i>P. mobilis</i>	9.4	1.53
pEt 21b+	3.6	1.14

Table A.4: NanoDrop of the plasmids containing *hfq* genes from *K. olearia*, *T. africanus* and *P. mobilis*, extracted from the clones. K represents the *K. olearia hfq*, A represents the *T. africanus hfq* and M represents the *P. mobilis hfq*.

	ng/ μ l	OD_{260}/OD_{280}
K1	25.7	1.57
K2	7.5	1.75
K3	5.8	1.67
K4	12.8	1.67
K5	4.6	1.34
A6	5.1	1.68
A7	2.3	1.61
A8	4.4	1.72
M11	5.8	1.39
M12	4.6	2.18
M13	3.9	2.24
M14	3.3	2.85
M15	4.1	1.52

Table A.5: Sequenced PCR products (section 3.1.4.2). Primer areas are highlighted in grey.

<i>Kosmotoga olearia</i> <i>hfq</i> PCR product	atggctgaaaaattcaatCTTCAGGATCGTTTCCTGAatcTcttgaggATCAAcaagatcgaggt taagatttccttgaaggtgggttcaaacgaaggaatcggtcgttcctttgataatttacggttctccttg gaccggccaggaacagacactgattacaagcacgcgataaagatgatGAtccccagagatacgttAa GCTGGCTGTCCagaataagaaggagggaagcttgatcatat
<i>Thermosipho</i> <i>africanus hfq</i> PCR product	tatcatatggcagaaaaatttaatttacaagacagATTTTTAaACAttCtaagaacta caaaaatacca gttaaagtttacttggttaatggtttcaaactaaaggaattataagatcctttgataattttacaatgctttg gaaaatggtagccaacaaaatttaattataaacatgcagttagtagacaattatgccagaaagttttgtcaaa ctaaCTaAGCAACAAaacgaagaatctgaaaatgaagaaaagaagcttatata
<i>Petrotoga mobilis hfq</i> PCR product	NNNNNNTTTNNGANATTTTTNAGNNNNACANATAGAAGTAAAAATCTATTTAG AAGGGGGATTTCAAACAAGTGGAGTGGTGAGATCTTTTGATGATTATACTGTTTT GTTGGGAAAAAACGGTGAACAATCACTGGTTTATAAACATGCCGTAAAAATGATG GTACCTTCGAAATATATTAACCTTTTCCAAGAACCACCGTCTAAAGATGAGAAGCT TATATA

Table A.6: The density of the cells was measured at different time intervals to get a final OD₆₀₀ of approximately 0.40, the time where the gene expression was induced. Clones with *hfq* from *K. olearia* were incubated for 130 minutes while clones with *hfq* from *T. africanus* and *P. mobilis* were incubated for 145 minutes.

Time intervals (minutes), OD ₆₀₀	<i>T. africanus hfq</i> clones control	<i>K. olearia hfq</i> clones	<i>T. africanus hfq</i> clones	<i>P. mobilis hfq</i> clones
0	0.04	0.04	0.04	0.04
30		0.05		
60				0.11
90			0.26	
110		0.31		
120		0.37	0.34	0.32
130	0.42	0.43	0.38	0.38
145			0.40	0.40
After induction	0.91	0.68	0.85	0.68

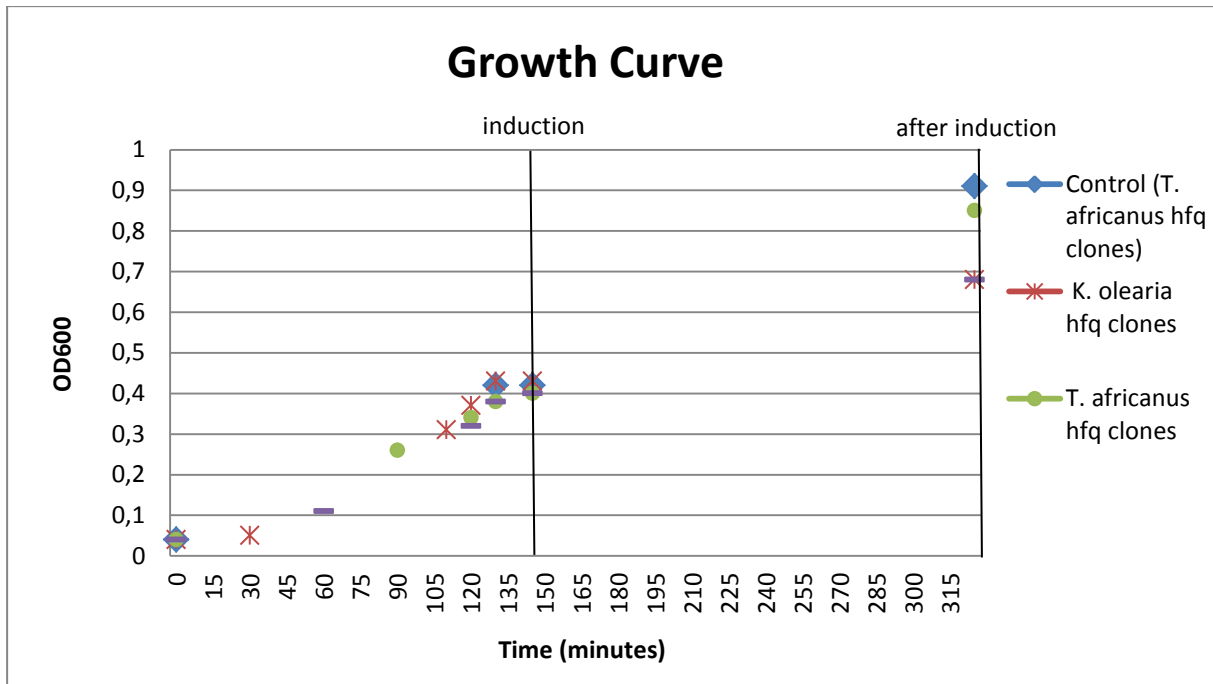


Figure A.1: The density of the cultures was measured at different time intervals to get a final OD₆₀₀ of approximately 0.40, the time where gene expression was induced. Clones with *hfq* from *K. olearia* were incubated for 130 minutes while clones with *hfq* from *T. africanus* and *P. mobilis* were incubated for 145 minutes.

Table A.7: OD₆₀₀ of *T. africanus* cultured at 70°C.

Time intervals (hours)	<i>T. africanus</i> absorbance at OD ₆₀₀
2	0.0073
4	0.0180
6	0.0434
8	0.0725
10	0.1088

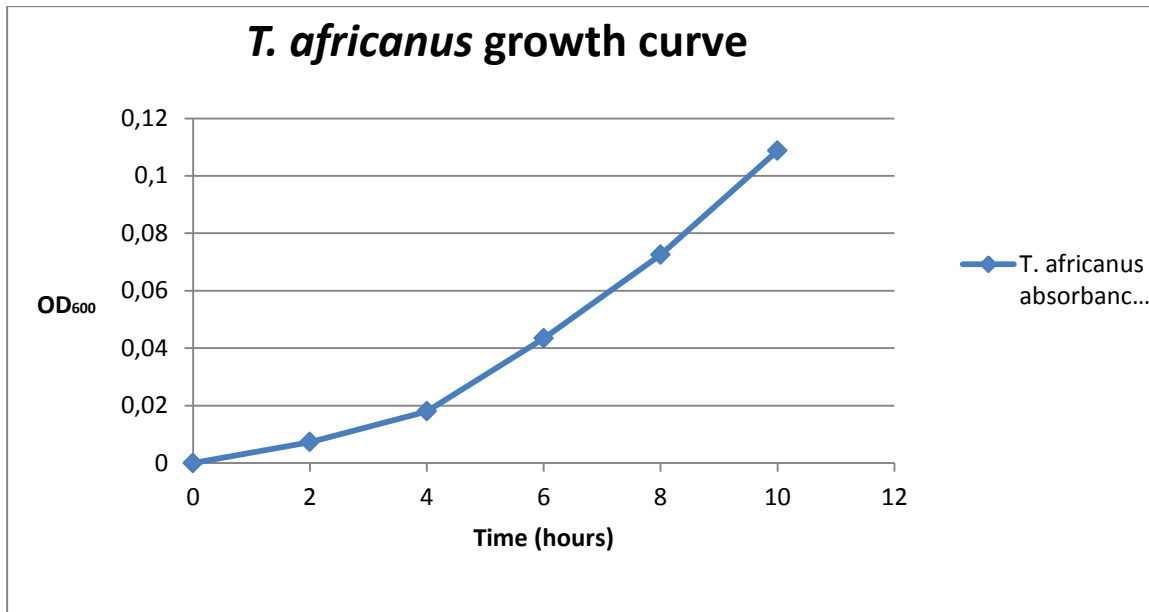


Figure A.2: OD₆₀₀ of *T. africanus* cultured at 70°C.

Table A.8: NanoDrop of DNA following gel extraction described in Section 2.12.

	ng/ μ l	OD ₂₆₀ /OD ₂₈₀
Anti	3.2	3.14
Pre	8.2	2.32

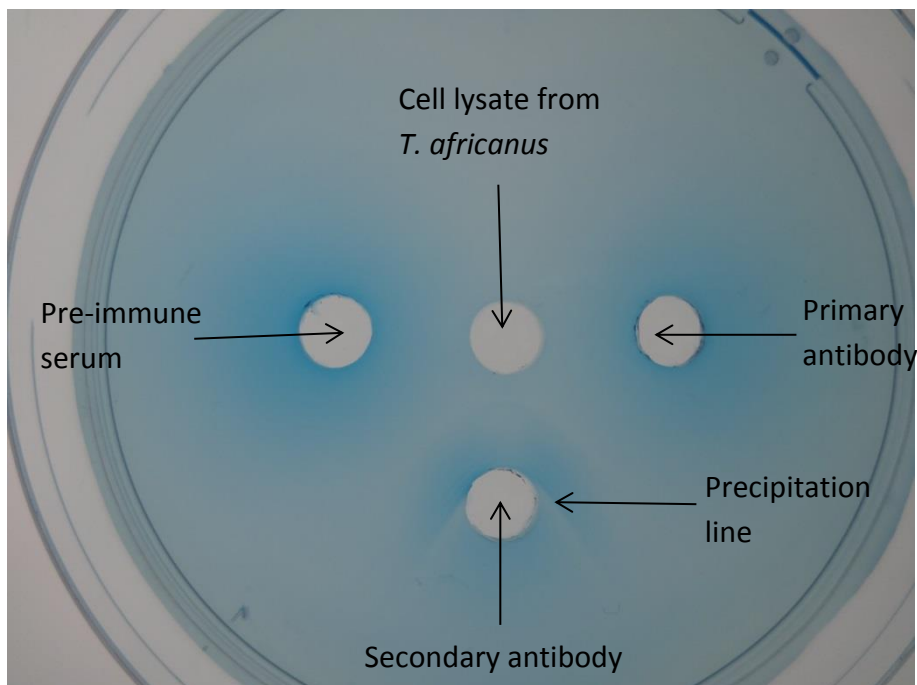


Figure A.3: Ouchterlony double immunodiffusion. In the center well; cell lysate from *T. africanus*, in the left well; pre-immune serum, in the bottom well; secondary antibody (goat anti-rabbit), in the right well; Primary antibody (rabbit anti-*TaHfq*). A precipitation line is formed between *T. africanus* cell lysate and the secondary antibody.

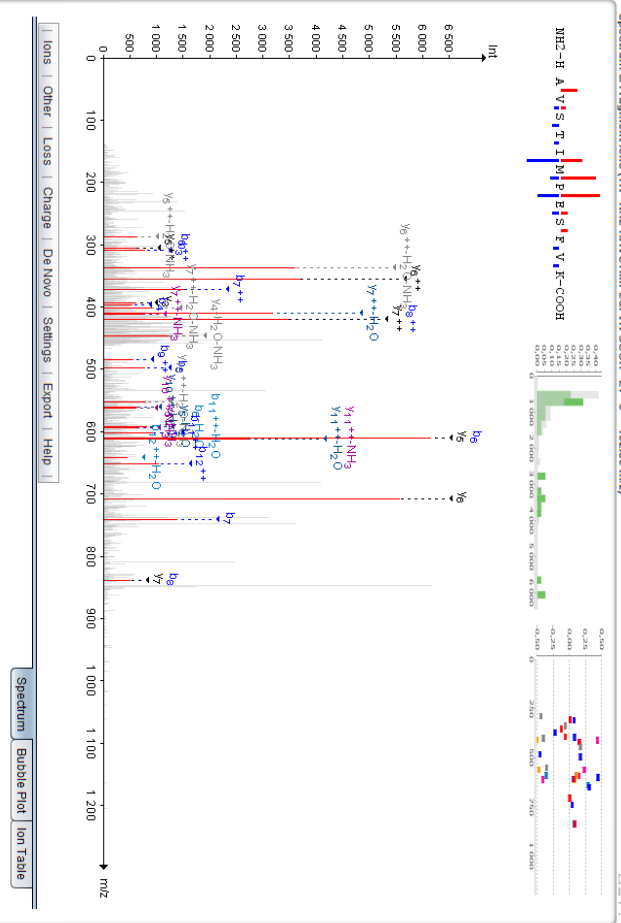
Accession	Description	Chr	Coverage	#Peptides	#Spectra	MS2 Quant	MW	Confidence
B7IHN6	Chromosome partition protein Snc (B7IHN6_THEAB)		18.98	21	24	0.00	134.00	89
B7IDK6	Sygal domain protein (B7IDK6_THEAB)		48.18	19	17	0.01	23.21	89
B7IJS6	Phosphotransferase/glycyl-protein-lyase (B7IJS2..)		5.88	2	2	0.00	20.00	89
B7IJS6	Phosphate transferase protein (B7IJS2_THEAB50_THEAB)		18.02	3	4	0.00	85.00	43
K2E009	HD core domain protein (K2E009_THEAB)		17.96	6	6	0.00	48.74	43
B7IEU6	RNAP sigma factor (B7IEU6_THEAB)		15.48	1	1	0.00	51.29	43
B7IGV3	Cisprf-associated protein (Csf11 family) (B7IGV3_THEAB)		23.21	23	27	0.00	138.83	44
ENSEMBL_ENSBTAP0000000735..	ENSEMBL_ENSBTAP0000000735 (Bos taurus) similar t..		3.30	36	37	0.00	192.87	43
B7IDJ9	Beta-galactosidase (B7IDJ9_THEAB)		66.96	4	4	0.00	128.82	38
B7ICV4	Unc119acetylase protein (B7ICV4_THEAB)		16.92	19	23	0.01	24.86	38
Q28107	SMISS-PROT_Q28107 (Bos taurus) Coagulation..		20.87	32	34	0.00	248.83	40
B7IEJ2	Alcohol dehydrogenase zinc-binding (B7IEJ2_THEAB)..		17.32	7	7	0.00	37.27	40
Q88VZ3	Q88VZ3 SWISS-PROT_Q88VZ3 Tax. Id=950a Gene_Sim.		12.42	15	16	0.00	103.94	38
B7ICQZ	DNA polymerase III, alpha subunit, Gam-positive type (B..		23.44	27	27	0.00	282.23	38
B7IEE9	Cisprf-associated protein, Csm2 family, putative (B7IEE9..		24.67	20	22	0.00	156.65	37

Peptides (11)

PI	Sequence	Start	#Spectra	Confidence
1	NH2-HAVSTIPESPFK-COOH	58	1	98

Peptide-Spectrum Matches (11)

ID	Sequence	Charge	Mass Error	Confidence
1	NH2-HAVSTIPESPFK-COOH	3	0.37	99



Protein Sequence Coverage (15.48%, 0.0% confident, 15.48% doubtful, 0.0% not validated, -100.0% possible - 94 AA)

Figure A.4: Result of Proteomics analysis of a cell extract of *T. africanus* as displayed with PeptideShaker. As shown, only 1 spectrum corresponding to the Hfq protein was identified.

8.2 Appendix B: Solutions and buffers

CTAB

Table B.1: TE Buffer. PH was adjusted to 7.5 with NaOH, and the solution was autoclaved prior to use.

TE buffer	
Tris	0.363 g
EDTA	111.6 mg
Milli-Q water	300 ml

Gel electrophoresis

Table B.2: Composition of the 1Kb+ ladder.

1 Kb Plus Ladder	
Standard (Invitrogen)	10 μ l
1xTAE	70 μ l
Gel loading Dye blue (6X)	20 μ l

Table B.3: Composition of the 6X loading buffer, the buffer was sterile filtrated prior to use.

6X loading buffer	
Bromophenol blue	0.05 g
Xylene cyanol FF	0.05 g
Sucrose	8 g
1xTAE	20 μ l

Table B.4: Supercoiled DNA ladder

Supercoiled DNA ladder	
Supercoiled DNA ladder (New England Biolab)	20 μ l
1xTAE buffer	60 μ l
Gel loading Dye blue (6x)	20 μ l

Table B.5: 50X TAE buffer

50X TAE buffer	
Tris base	242 g
Glacial Acetic Acid	57.1 ml
EDTA (0.5 M)	100 ml
Milli-Q water	Up to one litre

Table B.6: 1X TAE buffer

1X TAE buffer	
50X TAE buffer	20 ml
Milli-Q water	980 ml

Cultivation

Table B.7: LB(Luria-Bertani Medium) plates, final pH of 7.0 adjusted with NaOH. The medium was autoclaved. Ampicillin was added when the LB medium was approximately 55°C. LB+Amp₁₀₀ was then poured on Petri dishes. The dishes were left for solidification with the lid 1/5 (to avoid moist) off in a fume hood.

LB+Amp₁₀₀ (Luria-Bertani Medium)*	
Trypton	10 g
Yeast extract	5 g
NaCl	10 g
Agar*	15 g
Ampicillin	1 ml 100 mg/ml
Milli-Q water	1000 ml

*For liquid medium the agar was not added.

Table B.8: LB+Amp₁₀₀ og LB+Amp₁₀₀+Cam₃₄. LB(Luria-Bertani Medium) plates, final pH of 7.0 adjusted with NaOH. The medium was autoclaved. Ampicillin and Chloramphenicol was added when the LB medium was approximately 55°C. The dishes were left for solidification with the lid 1/5 (to avoid moist) off in a fume hood. *For liquid medium the agar was not added.

LB+Amp₁₀₀ og LB+Amp₁₀₀+Cam₃₄	
Trypton	5 g
Yeast extract	2.5 g
NaCl	5 g
Agar*	7.5 g
Ampicillin	500 µl 100 mg/ml
Chloramphenicol	500 µl 34 mg/ml
Milli-Q water	500 ml

SDS PAGE

Table B.9: Composition of the two polyacrylamide gels that were used in the SDS PAGE.

	Separation gel (20%)	Stacking gel (4%)
Milli-Q water	335 µl	1525 µl
1,5M Tris-HCl, pH 8.8	1250 µl	
0,5 M Tris-HCl, pH 6.8		625 µl
10% w/v SDS	50 µl	25 µl
30% Acrylamide	3340 µl	325 µl
10% w/v APS	25 µl	12.5 µl
TEMED	2.5 µl	2.5 µl
Total	5000 µl	2500 µl

Table B.10: Sample buffer used to prepare samples for SDS PAGE.

Sample buffer	
Milli-Q water	4 ml
0,5 M Tris-HCl, pH 6.8	1 ml
86-88% glycerol	0.8 ml
10% (w/v) SDS	1.6 ml
2 β mercaptoethanol	0.4 ml
0,1% (w/v) bromophenol blue	0.2 ml
Total	8 ml

Table B.11: 10X Running buffer used in SDS-PAGE

10x running buffer	
Tris	30.3 g
Glycine	144.2 g
SDS	10.0 g
Milli-Q water	

Table B.12: Staining solution for SDS PAGE

Staining solution	
Coomassie Brilliant Blue G-250	1 ml
Ethanol	400 ml
Acetic acid	100 ml
Milli-Q water	499 ml

Table B.13: Destaining solution for SDS PAGE

Destaining solution	
Ethanol	100 ml
Acetic acid	200 ml
Milli-Q water	700 ml

WESTERN IMMUNOBLOT

Table B.14: 10X Blotting buffer

10x blotting buffer (TGS)	
Tris	7.57 g
Glycine	36 g
Milli-Q water	250 ml

Table B.15: 1X blotting buffer

1x blotting buffer	
10x blotting buffer (TGS)	100 ml
Methanol	200 ml
Milli-Q water	700 ml

Table B.16: 1X Tris buffered saline (TBS)

1x TBS	
10x blotting buffer (Tris buffered saline, Bio-Rad)	100 ml
Milli-Q water	900 ml

Table B.17: TTBS

TTBS	
10x blotting buffer (Tris buffered saline, Bio-Rad)	70 ml
Milli-Q water	630 ml
Tween-20	350 µl

Table B.18: Blocking buffer

Blocking buffer	
1x TBS	50 ml
Gelatine (3%)	1.5 g

Table B.19: Antibody solution

Antibody solution	
1x TTBS	50 ml
Gelatine	0.5 g (1%)
Antiserum	50 µl (1:1000)

Table B.20: Conjugate solution

Conjugate solution	
1x TTBS	50 ml
Gelatine	0.5 g (1%)
Goat Anti-Rabbit IgG (H+L) – HRP conjugate, Bio-Rad	16.5 µl (1:3000)

Table B.21: Staining solution

Staining solution for immunoblotting	
HRP color reagent B, Bio-Rad	150 µl
HRP color development buffer, Bio-Rad	25 ml (diluted 1:10)
HRP color reagent A, Bio-Rad	5 ml

Ouchterlony double immunodiffusion

Table B.22: Agar for Ouchterlony double immunodiffusion. The mix gave three Petri dishes with 20 ml agar in each dish.

1% agar for Ouchterlony double immunodiffusion	
1X TBS	60 ml
Agarose	0.6 g

8.3 Appendix C: Materials and methods

Table C.1: List of the suppliers of chemicals, reagents and gasses used in the laboratory work.

Chemical/reagent	Supplier
2-Mercaptoethanol	Bio-Rad
Sodium acetate 3 M	Invitrogen
Sodium chloride 5 M	Sigma-Aldrich
4-Aminobenzoic acid	Sigma-Aldrich
NEBuffer2 (10x)	New England Biolabs
Acrylamide/Bis-Acrylamide 30%	Bio-Rad
Acetic acid	Sigma-Aldrich
Agar	Fluka Analytical
Agarose	Lonza
Ammonium chloride	Merck
Ammonium iron(II) sulphate hexahydrate	Merck
Ammonium Persulphate APS	Bio-Rad
Ampicillin sodium salt	Sigma-Aldrich
Big Dye Sequencing buffer	Life technologies
BigDye	Life technologies
Boric acid	Merck
Bromophenol blue	Sigma-Aldrich
Butanol	Merck
Ca-D(+) pantothenate	Sigma-Aldrich
Calcium chloride dehydrate	Merck
Chloramphenicol	Sigma-Aldrich
Chloroform	Merck
Cobalt(II) chloride hexahydrate	Merck
Coomassie Brilliant Blue R-250	Bio-Rad
Copper(II) sulphate pentahydrate	Merck
CTAB	Sigma-Aldrich
Cysteine hydrochloric acid hydrate Cysteine-HCl x H ₂ O	Sigma-Aldrich
D(+)-biotin	Sigma-Aldrich
DEPC water	Invitrogen
Dipotassium hydrogenphosphate	Sigma-Aldrich
dNTPs (10 mM)	New England BioLabs
Ethanol	Sigma-Aldrich
Ethylenediaminetetraacetic acid EDTA	Sigma-Aldrich
Gelatine	Bio-Rad
GelRed	Biotium
Glucose	Sigma-Aldrich
Glycerol	Merck
Glycine	Sigma-Aldrich
HisTALON xTractor Buffer	Clontech
Hydrogen chloride	Sigma-Aldrich
Iron(II) sulphate heptahydrate	Sigma-Aldrich

Isoamyl alcohol	Merck
Isopropylthio-β-galactoside IPTG	Amersham GE Healthcare Life Sciences
Isopropyl alcohol	Sigma-Aldrich
Ligation buffer 10x with 10 mM ATP	New England Biolabs
Magnesium chloride hexahydrate	Sigma-Aldrich
Magnesium sulphate heptahydrate	Fluka Analytical
Maltose	Sigma-Aldrich
Manganese(II) sulphate monohydrate	Merck
Methanol	Sigma-Aldrich
Mussel glycogen	Invitrogen
N,N,N',N'-tetramethylethylenediamine TEMED	Bio-Rad
Nickel(II) chloride hexahydrate	Merck
Nicotinic acid	Merck
Nitrilotriacetic acid	Sigma-Aldrich
Nitrogen gas	YARA
One Taq Standard Reaction 5x Buffer	New England BioLabs
Peptone	Merck
PIPES 1,4-Piperazinediethanesulfonic acid	Sigma-Aldrich
Potassium aluminium sulphate dodecahydrate KAl	Merck
Potassium chloride	Sigma-Aldrich
Potassium dihydrogen phosphate	Merck
Pyridoxamine dihydrogenchloride	Sigma-Aldrich
Resazurin sodium salt 7-Hydroxy-3H-phenoxazin-3-one 10-oxide Dye content	Sigma-Aldrich
Sodium dodecyl sulphate SDS	Sigma-Aldrich
Sepharose A	Amersham GE Healthcare Life Sciences
Sodium bicarbonate	Merck
Sodium chloride	Sigma-Aldrich
Sodium dithionite	Merck
Sodium molybdate dihydrate	Kebo Lab
Sodium sulphide	Merck
Sodium sulphide nonahydrate	Merck
Sodium thiosulphate pentahydrate	Merck
Sucrose	Sigma-Aldrich
Thiamine dichloride	Sigma-Aldrich
Trizma base Tris(hydroxymethyl) aminomethane	Fluka Analytical

Tris buffered saline 10x	Bio-Rad
Trypton	Fluka Analytical
Tween-20	Bio-Rad
Xylene Cyanol FF	Sigma-Aldrich

Table C.2: Standards and ladders

Marker	Supplier
1 Kb Plus DNA ladder	Invitrogen
SDS-PAGE standard (Broad range)	Bio-Rad
Full-Range Rainbow Molecular Weight Marker	Amersham GE Healthcare Life Sciences
NEB Supercoiled DNA Ladder	New England Biolabs

Table C.3: Kits

Kit	Supplier	Protocol available from*:
GenEluteGel Extraction Kit	Sigma-Aldrich	http://www.sigmaaldrich.com/content/dam/sigmaaldrich/docs/Sigma/Bulletin/na1111bul.pdf
GenElute PCR Clean-UP Kit	Sigma-Aldrich	http://www.sigmaaldrich.com/content/dam/sigmaaldrich/docs/Sigma/Bulletin/na1020bul.pdf
GenElute Plasmid Miniprep Clean-UP Kit	Sigma-Aldrich	http://www.sigmaaldrich.com/content/dam/sigmaaldrich/docs/Sigma/Bulletin/pln70bul.pdf
GeneRacer Kit	Invitrogen, Life technologies	http://www.iitg.ernet.in/sachinku/Invitrogen_RACE%20kit.pdf
Immun-Blot AP Colorimetric Kit	Bio-Rad	http://www.bio-rad.com/webroot/web/pdf/lsr/literature/LIT171.pdf
One Shot TOP10 Chemically Competent Cells	Invitrogen, Life technologies	http://tools.lifetechnologies.com/content/sfs/manuals/oneshottop10_chemcomp_man.pdf
One Shot BL21(DE3) One Shot BL21(DE3)pLysS One Shot BL21(DE3)pLysE Competent Cells	Invitrogen, Life technologies	http://tools.lifetechnologies.com/content/sfs/manuals/oneshotbl21_man.pdf
Poly(A) Tailing Kit	Ambion, Life technologies	http://tools.lifetechnologies.com/content/sfs/manuals/bp_1350.pdf
TOPO TA Cloning kit for sequencing	Invitrogen, life technologies	http://tools.lifetechnologies.com/content/sfs/manuals/topotaseq_man.pdf
USB ExoSap-IT PCR Product Cleanup	Affymetrix UK	http://media.affymetrix.com/support/technical/usb/brief_proto/78200B.pdf

*Downloaded 08.11.2014.

Table C.4: Enzymes and polymerases

Enzyme / polymerase	Supplier
DNase1	Sigma-Aldrich
One Taq DNA polymerase, 5,000 U/ml	New England Biolabs
Recombinant RNasin Ribonuclease inhibitor, 40 U/ μ l	Promega
Proteinase K (20 mg/ml)	Sigma-Aldrich
RNaseA (20 mg/ml)	Sigma-Aldrich
T4 DNA ligase, 400,000 U/ml	New England Biolabs

Table C.5: The vector (pET-21b(+)) used for cloning of *hfq*.



pET-21a-d(+) Vectors

TB036 12/98

	Cat. No.
pET-21a DNA	69740-3
pET-21b DNA	69741-3
pET-21c DNA	69742-3
pET-21d DNA	69743-3

The pET-21a-d(+) vectors carry an N-terminal T7•Tag® sequence plus an optional C-terminal His•Tag® sequence. These vectors differ from pET-24a-d(+) only by their selectable marker (ampicillin vs. kanamycin resistance). Unique sites are shown on the circle map. Note that the sequence is numbered by the pBR322 convention, so the T7 expression region is reversed on the circular map. The cloning/expression region of the coding strand transcribed by T7 RNA polymerase is shown below. The f1 origin is oriented so that infection with helper phage will produce virions containing single-stranded DNA that corresponds to the coding strand. Therefore, single-stranded sequencing should be performed using the T7 terminator primer (Cat. No. 69337-3).

pET-21a(+) sequence landmarks

T7 promoter	311-327
T7 transcription start	310
T7•Tag coding sequence	207-239
Multiple cloning sites (<i>Bam</i> H I - <i>Xho</i> I)	158-203
His•Tag coding sequence	140-157
T7 terminator	26-72
<i>lac</i> I coding sequence	714-1793
pBR322 origin	3227
<i>bla</i> coding sequence	3988-4845
f1 origin	4977-5432

The maps for pET-21b(+), pET-21c(+) and pET-21d(+) are the same as pET-21a(+) (shown) with the following exceptions: pET-21b(+) is a 5442bp plasmid; subtract 1bp from each site beyond *Bam*H I at 198. pET-21c(+) is a 5441bp plasmid; subtract 2bp from each site beyond *Bam*H I at 198. pET-21d(+) is a 5440bp plasmid; the *Bam*H I site is in the same reading frame as in pET-21c(+). An *Nco* I site is substituted for the *Nde* I site with a net 1bp deletion at position 238 of pET-21c(+). As a result, *Nco* I cuts pET21d(+) at 234, and *Nhe* I cuts at 229. For the rest of the sites, subtract 3bp from each site beyond position 239 in pET-21a(+). *Nde* I does not cut pET-21d(+). Note also that *Sty* I is not unique in pET-21d(+).

