Graz University of Technology

Institute of Biochemistry

Bachelor Thesis:

In-silico identification, production and verification of the receptor binding protein of the *Paenibacillus* phage HB10c2

Felix Schweigkofler

01635655

Supervised by

Univ.-Prof. Dr. rer. nat. Peter Macheroux

January 2020

Acknowledgements

First and foremost I want to thank Prof. Dr. Peter Macheroux, who made this thesis as well as the iGEM 2019 project *Beeosensor* possible. Both allowed me to finally dive into real science, as the study had been rather theoretical until this point. Opportunities like this one really help with the transition from a student to a researcher.

Secondly, I want to express my gratefulness to the student team *iGEM NAWI Graz 2019* for opportunity to do my bachelor thesis on this topic related to the project *Beeosensor*. The award-winning project was first envisioned and planed by my team colleague Nikola Vinko, BSc, whom I want to thank for the support during the planning stage and in the laboratory. It was an elevating experience to work with him and the rest of my team.

My special thanks go to Dr. Marina Toplak, who has helped me very much with all my questions and problems and has shown more than any other person in the course of my studies how science looks like in the lab and thus in reality. I have learned a lot, explicit knowledge as well as – and even more importantly – implicit behavior in a lab and within scientific institutions. I am very thankful for the guidance and the insights I received.

Furthermore, I want to mention at least some of the other people, who have made this work possible: Prof. Karl Gruber for the *in silico* planning, Henrik Seyfried, BSc and Dr. Gustav Oberdorfer for the more detailed modeling of the RBP to give me confidence in my prior assumptions about the receptor binding protein, Birgit Grill, MSc for the Gibson cloning and Dr. Klaus Natter for the help with the Cell Counter. The rest of the biochemistry research group will be unnamed but not forgotten and I want to thank you for your patient help and advice with the many small problems.

Last and therefore the most I want to finally honor and mention my parents and sisters on an in a work of mine for all the things they have done to make it possible for me to live my life as I do. They have almost never expressed unfounded concerns and if they did it was out of their best intentions for me. Many have not been as lucky as I am by being blessed with this crazy and loving family who has given me so much.

Zusammenfassung

Bakteriophagen-Rezeptorbindeproteine (RBP) sind in die spezifische Erkennung und die Bindung von Bakterien über verschiedene Typen von Rezeptoren involviert. In der Phagen-Unterklasse *Siphoviridae* sind RBPs an der äußersten Spitze des Siphons in der „Baseplate“ verankert. Die „Baseplate“ besteht aus einer je nach Phagem variierenden Anzahl an Proteinen und ist für die Bindung des Phagen an den Wirt verantwortlich, um die Injektion der Phagen-DNA in das Zytoplasma zu ermöglichen.

Diese Proteine sind das Kernelement des Bindemechanismus der Phagen, dennoch ist in der Literatur nur *relativ* wenig Information über sie zu finden. Da die spezifische Erkennung von Bakterien ein großes Potential für die Whole-Cell-Detection und andere Anwendungen hat, ist es jedoch von großer Bedeutung mehr über diese Proteine zu erfahren.

Im Zuge dieser Arbeit wurde das Genprodukt gp17 (HB_00017) als wahrscheinlicher Kandidat für das RBP des *Paenibacillus* Phagen HB10c2 identifiziert. Dafür wurden das sequenzierte Genom mit dem Genom anderer Phagen verglichen und die Proteinkandidaten mit Phyre2 modelliert. Anschließend wurde das Protein heterolog in *Escherichia coli* produziert. Dieses schien aber unlöslich zu sein. Ein neu entworfener Test zur Identifizierung von RBPs, basierend auf einer Ni-NTA-Affinitätschromatographie, ist fehlgeschlagen. Ein Colony-Assay zur Identifizierung von RBPs ergab einen Hinweis darauf, dass das von HB_00017 codierte Protein tatsächlich das RBP von HB10c2 ist.

Abstract

Receptor binding proteins (RBP) of bacteriophages are involved in the specific recognition and binding of bacteria via all types of receptors. In the phage subclass *Siphoviridae* they are located at the most distal part of the tail, anchored in the baseplate. The baseplate consists of a varying number of proteins for each phage and is responsible for connecting the phage with the host bacterium to enable the phage to inject its DNA into the cytoplasm.

These proteins are key for the binding mechanism of phages, but *rather* limited information is available in literature. Since the specific recognition of bacterial cells has a great potential for whole cell detection and other applications, it would, however, be important to obtain a better understanding of these proteins (RBPs).

In the course of my studies, the gene product gp17 (HB_00017) has been identified as a likely candidate for being the RBP of the *Paenibacillus* phage HB10c2. The sequenced genome of HB10c2 were compared with the genome of other phages and putative protein candidates were modelled with Phyre2. Subsequently, the protein was produced heterologously in *Escherichia coli*. However, it seemed to be insoluble. A newly designed test for RBPs based on Ni-NTA affinity chromatography failed to confirm the interaction of gp17 with *Paenibacillus larvae*. A colony-assay for the identification of RBPs suggested that the protein encoded by HB_00017 could be the RBP of HB10c2.

Abbreviations

AA	acrylamide
AFB	american foulbrood
APS	ammonium persulfate
BHI	brain-heart-infusion
B-PER	bacteria protein extraction reagent
CSA	columbia sheep blood agar
CFU	colony forming units
dit	distal tail protein
DMSO	dimethyl sulfoxide
EDTA	ethylenediaminetetraacetic acid
fw	forward
HB_000xx	name of gene product of gene number xx in The *Paenibacillus larvae* phage HB10c2
IPTG	isopropyl β-d-1-thiogalactopyranoside
LB	Luria-Bertani broth
LBC	LB with chloramphenicol
LBK	LB with kanamycin
MYPGP	growth medium for *Paenibacillus larvae*
Ni-NTA	nickel-nitrilotriacetic acid-N,N bis(carboxymethyl) glycine
OD_{600}	optical density at a wavelength of 600 nm
ONC	overnight culture
PAGE	polyacrylamide gel electrophoresis
PBST	phosphate buffered saline with tween 20
PCR	polymerase chain reaction
PMSF	phenylmethylsulfonyl fluoride
RBP	receptor binding protein
rv	reverse
SDS	sodium-dodecyl-sulphate
TAE	tris, acetic acid and EDTA buffer
tal	tail associated lysin
tmp	tail length tape measure protein
T_M	melting temperature

VHH5	heavy chain antibody fragment obtained from llama
Y-PER	yeast protein extraction reagent

Internal nomenclature of genes and proteins

g17 is the gene number 17 in the genome of the *Paenibacillus larvae* phage HB10c2 and equivalent to HB_00017.

gp17 is the gene product of gene number 17. It is equivalent to p17 (protein 17), which is the name used for the protein when produced during this work.

tg17 (twist g17) is the result of the codon optimization of g17 for *Escherichia coli*. Its product when expressed is the same as gp17 and p17.

Table of contents

Introduction

1.1 American foulbrood (AFB) and *Paenibacillus larvae*

The American foulbrood was the most threatening bee disease in the first half of the 20th century and could only be reduced in the second half by state apiary programs which monitored and treated outbreaks [1]. This however does not mean the disease itself has lost its threatening potential. Beekeepers, beekeeping institutions and bee health researchers are still very concerned about the disease.

AFB is caused by the gram positive, endospore forming bacterium *Paenibacillus larvae* [2]–[4]. The spores of *P. larvae* are viable for more than 35 years and can be found everywhere in a beehive, including the honey [5]. When approximately 10 spores are fed to the honey bee larvae by worker bees via contaminated honey, they are activated in the larval gut [6,7]. Activated spores proliferate and thus digest and kill the larvae. The slime contains a very high number of bacteria turned into endospores, ready for infecting more larvae.

The disease is highly contagious and is spread by contaminated bees accidentally flying to other hives rather than to their own, robbery by healthy colonies, the beekeeper transferring bacteria from one hive to another [8]. Also, contaminated honey is a source of infection. Contaminated honey mostly stems from non-EU countries, where the disease is monitored less rigidly and where local honey bee subspecies might be more resistant to the AFB [9]. To prevent the rapid spread of the disease, drastic measures are enforced by law and might require the burning of the hives and colonies [10]. Beekeepers in Styria also said that some would burn the hives without noticing officials.

The disease can be identified visually and by examination of the honey bee brood, which is the first step of the legal diagnosis of AFB in Austria (§5 of the Bienenseuchengesetz (25th of May 1988)). Current detection methods for the bacterium include quick tests and a culture method. The quick tests are unreliable and have a high detection limit. Therefore, they are not used frequently by beekeepers. The culture method for analyzing the growth of *P. larvae* colonies is preferably done using selective media for gram positive bacteria and is much more reliable. The presence of *P. larvae* is confirmed by a microscopic examination and sometimes also by a PCR for the unequivocal identification [11]. The aim of the culture method is to identify possibly endangered hives as soon as possible to give the beekeeper the opportunity to take preemptive measures, which can prevent the outbreak of the disease.

As the effectiveness of antibiotics decreases and the disease still poses a threat to beekeeping and an economic threat to the beekeeping industry, other ways of controlling and treating AFB are being researched. Particularly promising in this aspect is the phage therapy, using *Paenibacillus* phages [12].

1.2 *Paenibacillus larvae* phages HB10c2

Paenibacillus larvae phages are a member of the *Siphoviridae* family. Electron microscopy has revealed that the phage consists of an elongated capsid, twice as long as wide, a siphon (tail), 1.5 times the size of the capsid, and a baseplate at the end of the siphon with twice its diameter. The total length is about 250 nm [13]. This corresponds to the B2 morphotype of *Siphoviridae* phages [14].

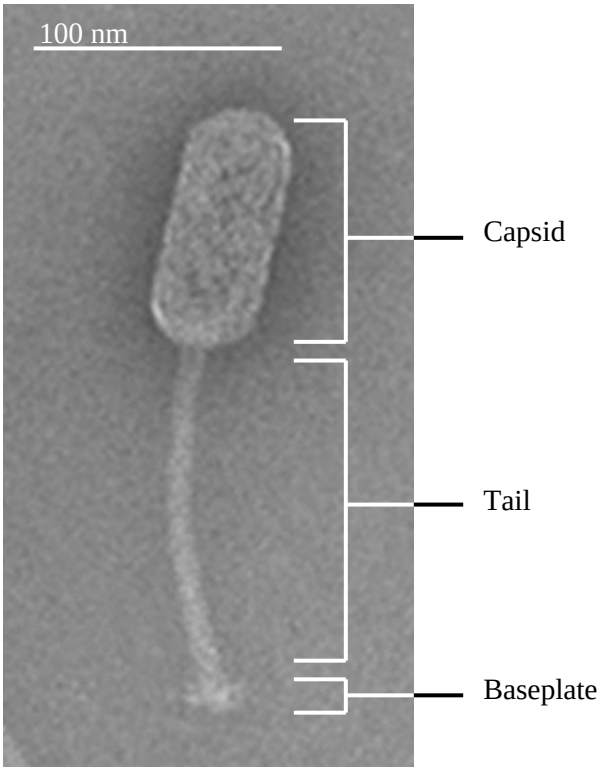

Figure 1: Detail of a transmission electron micrograph of HB10c2 performed by Hannes Beims [13]. The compact baseplate at the tip of the tail recognizes the host bacterium and attaches to it. The capsid contains the DNA, which is injected into the host cell trough the tail and the baseplate. The RBP is located at the tip of the baseplate. Modeling of neither part have been performed yet, therefore it is impossible to say how exactly the baseplate looks like, as they share their basic structure, but are still diverse in detail. The micrograph confirms the classification of HB10c2 as Siphovirus.

The host range of HB10c2 is rather broad with not only *P. larvae* strains susceptible to it, but also several other members of the *Bacillales* genus, most of which are *Paenibacillus* species. No gram-negative bacteria were shown to be lysed by HB10c2, so far. Even though its lytic activity on P. larvae is confirmed, HB10c2 does not seem to have a therapeutic effect on diseased bee larvae, infected by *P. larvae* ERIC I and ERIC II [13].

1.3 The use of *Paenibacillus larvae* phages

Paenibacillus larvae phages have been known for decades, with the first strain discovered in the 1950s [15]. Since that time many more phages have been isolated, with the most recent

being a strain from Portugal [16], a strain from Germany [13] (used in this research) and dozens of strains in the US [15]. All of them have been sequenced and thus may be used for *in silico* comparison.

The authors of the related papers highlight the possibilities of the phages for treating AFB, especially looking at the ban of antibiotics in apiculture in Europe (EC regulation 2377/90: Regulation on Maximum Residue Limits), and the increase in antibiotic resistance in *P. larvae* [12]. Also, it seems to be a practical therapy, compared to the high disadvantages of the most commonly applied methods of treating AFB such as burning the hives and the low effectiveness of alternative therapies, such as essential oils [2].

Paenibacillus phages can be used for a classical phage therapy and furthermore the use of a recombinantly produced phage lysin has been proposed [17,18]. Not all of the *Paenibacillus* phages are suitable as therapeutics, e.g. HB10c2 shows lytic activity against *P. larvae* in plaque assays, but lacks a toxin encoded by the Phage phiIBB_Pl23 [13]. The phage therapy also struggles with the high specificity of some phages, as many of the phages isolated in the USA could only attach to a few strains of *Paenibacillus larvae* [19].

1.4 Bacteriophages

Bacteriophages are different types of viruses, that infect only bacteria and archaea and are usually very specific for one species or even one strain [20]. As they are viruses they have no metabolism and rely on the metabolism of the host for their reproduction.

The life cycle of lytic phages contains the recognition of a proper host, attachment of the phage to the host, injection of the phage DNA into the cell and corruption of the host metabolism with the DNA products. The phages reassemble themselves and lyse the cell to set free new phages. The cycle of lysogenic phages includes the insertion of the phage DNA into the host DNA via characteristic recombination mechanisms, e.g. the cre/loxP-mechanism [21]. This way, the phage DNA is reproduced by physiological bacterial cell division. The phage DNA switches to the lytic cycle when triggered by rough circumstances which threaten the proliferation of the host [22].

The ecological relevance of bacteriophages can hardly be overestimated, as e.g. the oceans contain an estimated number of 10^{31} bacteriophages which are a relevant part of the ecological equilibrium [23]. The outstanding relevance of some phages to humans stems from their effective antagonism against bacteria used in the food industry as well as pathogenic bacteria. While the rise of antibiotics has led to a low interest in bacteriophages during the 20th century,

they are gaining more and more interest as therapeutics with the crisis of antibiotics since the beginning of the 21th century [24].

Alongside the more famous *therapeutic* potential of bacteriophages, research on the *diagnostic* use of phages is increasing. Here, the bacteriophages act as a very specific receptor for the bacterium. The methods may vary, but share the phage as recognition element for the bacterium. Receptor binding proteins of phages might be able to substitute the use of whole phages in some of these applications.

1.5 Receptor binding proteins of tailed bacteriophages

Receptor binding proteins are the core element required for host recognition by tailed bacteriophages. From now on "phage" stands for "tailed bacteriophage". Each phage has at least one type of receptor binding protein, but it is not necessarily limited to having only one type, as two types can broaden its host spectrum [25].

The mode of phage attachment to the host is remarkable, as almost every collision with the host leads to an irreversible attachment. *Siphoviridae*, a subsection of tailed bacteriophages, as well as other phages use a dual technique, when attaching to a bacterium. First they walk across the surface with weak interactions between the tail fibers and the surface of the bacterium. When the RBPs make contact with their respective receptors, they form a nearly irreversible bond with the bacterium [26].

1.6 Kinetics of RBP attachment

The kinetics of a protein attaching to a receptor are different to the kinetics of a phage attaching to the same receptor. This would be interesting to examine, but is beyond the means of this thesis, as the kinetics of a protein and a phage are very different.

Tailed bacteriophages use their fibers to walk across the surface of bacteria until the RBPs attach to their receptor [26]. The weak electrostatic interactions allowing for this walk likely stem, at least partially, from the RBPs of the phages, as they are located on the very tip of the baseplate [27]. As these interactions are weak, probably no special protein motif is necessary.

If however the interaction is caused by the RBPs in a way that is different from any normal interaction between soluble proteins and cells, many of the RBPs added to a solution of bacteria might attach randomly to the surface of any bacteria. While the phage can walk and eventually detach from any unspecific bacteria, as its comparatively huge size makes it easier to create the force necessary to loosen the weak bond, an RBP might stay attached. This is not true for cases,

where the RBP is immobilized onto some surface, but might play a role in other scenarios. The kinetics could therefore drastically influence the specificity of any test using soluble RBPs attaching to and highlighting bacteria.

1.7 Relevance of isolated RBPs

A single mutation in the RBP gene can lead to a drastic change of the affinity, as it was shown for the RBP of *Lactococcus lactis* phages. After just a single nucleotide replacement they were able to avoid the binding of VHH5, which is produced by the bacterium to irreversibly attach to the RBP of the phage and to thereby inactivate it [28]. Recombinant phages can infect new hosts, if their RBP has been exchanged with the RBP of a phage directed to this host [29]. This underlines the importance of understanding the structure of the baseplate of phages and most importantly the RBP.

The potential of phages to adapt to certain circumstances poses a threat if the phages are unwanted, but it also gives a tool for the improvement of phage applications, which are based on the specific and the strong binding of phages to their respective hosts. Understanding the mechanisms of phage assembly and RBP-binding is necessary to develop new approaches to guarantee the effectiveness of phages, even against attempts of the bacteria to develop resistance mechanisms.

Furthermore, the identification of the RBP of a phage is also one possible way of finding the corresponding receptor by determining the structure of the binding domain to postulate a fitting structure for the receptor.

1.8 Advantages of RBPs over whole phages as recognition element

As the recognition of the host bacterium is based on the RBP or RBPs of a phage, the rest of the phage is not necessary for binding. Once the RBP of a phage is known, it can be produced and used instead of the phage for the specific recognition in several applications. This approach has numerous advantages.

The production and purification of a protein is far easier and cheaper than the production and purification of bacteriophages and can be scaled up easily, making it even cheaper when producing larger amounts of protein. The production does not rely on potentially dangerous and pathogenic host bacteria, which might in addition be difficult and inefficient to cultivate.

Proteins are smaller than complete phages, which could be necessary for some applications. Proteins are generally more stable than large structures built from several proteins and can

endure harsher environments. They can be modified to enhance their stability, e.g. the receptor binding domain of the RBP of *Lactococcus lactis* phage bIL170 has been shown to be more stable than the rest of the protein [30].

Using proteins with a known function does not pose any biological threat to the environment and biosafety, while phages could potentially harm the ecosystem when applied in an unusually large amount. Phages can furthermore pose the threat of transducing dangerous DNA sequences and turning harmless bacteria into pathogens. This might be the case for HB10c2, which expresses a beta-lactamase-like protein [13].

1.9 *In silico* identification of receptor binding proteins

Phage genes are often roughly grouped in functional clusters and the same goes for *Siphoviridae*, the family of HB10c2 [13]. One cluster contains structural proteins and the rough order of the genes is conserved throughout many species.

The genes coding for the baseplate structure of the phages are located between the tail length tape measure protein (tmp) and host lysis genes. Two genes regularly found are the distal tail protein (Dit) and the tail-associated lysin protein (Tal), which build the core of the baseplate and are followed by more baseplate proteins, including the receptor binding protein (RBP) [31]. Many RBPs are trimers and share the same functional structure, with a region linked to the phage, a middle part and the binding site, which contains the specific loop regions [31–33].

1.10 The genome of HB10c2

The genome of HB10c2 is a linear and double stranded DNA, coding for 56 proteins, and has a length of 35,644 bp, with 50% of all known phages sharing a similarly sized genome [13] [34]. Other *Paenibacillus larvae* phages vary in length and sequence, e.g. Bacteriophage Diva (NCBI accession number KP296791) has 37246 bp and phiIBB_Pl23 (NCBI accession number KF010834.1) has 41294 bp, but the main character remains the same. Other *Paenibacillus* phages contain several genes not found in HB10c2 and have a slightly different sequence in the genes they share [13].

The genome of HB10c2 is clustered and consists of regions with different functions. Packaging genes are found on the 5'-end of the genome, followed by structural genes which form the biggest cluster ranging from gene HB_00003 to gene HB_00017, later followed by host lysis genes and genes for the lysogenic cycle, according to a comparison of the genomic structure of HB10c2 and phiIBB_Pl23. The gene HB_00014 has been identified as putative tmp and gene HB_00020 has been identified as the first of the putative host lysin genes. Several of the genes

and their putative function have been identified, while the genes coding for the baseplate including the receptor binding protein remained unknown, so far [13].

1. Materials

All the chemicals and media were purchased from

Sigma Aldrich (St. Louis, MO, USA), Roth (Karlsruhe, Germany), Merck (Darmstadt, Germany), Fluka (Buchs, Switzerland), Ottonordwald (Germany)

The Gene with the original codon sequence was amplified by PCR from the phage HB10c2 which was sent by Hannes Beims from LAVES (Niedersächsisches Landesamt für Verbraucherschutz und Lebensmittelsicherheit), Germany. The gene with the optimized codon sequence for expression in *Escherichia coli* was ordered from Twist Biosciences (San Francisco). All the primers were purchased from Sigma Aldrich.

2.1. Culture/Media

- LB-Medium
    - 5 g/L Yeast Extract
    - 10 g/L Trypton/Pepton
    - 5 g/L NaCl
- BHI-Medium
    - 37 g/L BHI (Brain heart infusion)
    - 3 g/L Yeast Extract
- MYPGP
    - 10 g/L **M**ueller Hinton Broth
    - 15 g/L **Y**east Extract
    - 3 g/L $K_2H\mathbf{P}O_4$
    - 1 g/L Sodium **p**yruvate
    - 2 g/L **G**lucose (add only after cooling down to 60 °C)
- CSA – Columbia sheep blood agar (purchased from Ottonordwald)
    - 23 g/L Peptospecial
    - 1 g/L Starch
    - 55 g/L Sodium chloride
    - 0.01 g/L Colistin sulphate
    - 0.015 g/L Nalidixic acid
    - 5% Sheep blood defibrinated
    - 14 g/L Agar 14.0

For agar plates: Add 15 g/L Agar (-Agar) after dissolving the other components but before autoclaving.

For Antibiotics: Add antibiotic after autoclaving when the temperature of the medium is not higher than 70 °C.

2.2. Buffer systems

- PBST: Phosphate buffered saline with tween 20 (137 mM NaCl, 2.7 mM KCl, 12 mM $H_2PO_4^-$) with 0.5% Polysorbat 20 (Tween 20), pH 7
- Gibson buffer: 25% [wt/vol] PEG-8000, 500 mM Tris-HCl pH 7.5, 50 mM MgCl2, 50 mM DTT, 1 mM each of the four dNTPs, 5 mM NAD
- Gibson Mastermix: 699 µL H_2O, 320 µL Gibson buffer, 0,64 µL 10 U/µL T5 exonuclease (Epicentre‡), 20 µL 2 U/µL Phusion DNA polymerase (NEB), 160 µL 40 U/µL Taq DNA ligase (NEB)
- Blocking solution: 5% skim milk (Roth) in PBST
- Binding Buffer: 50 mM NaH_2PO_4, 10 mM imidazole, pH 8
- Washing Buffer: 50 mM NaH_2PO_4, 20 mM imidazole, pH 8
- Elution Buffer: 50 mM NaH_2PO_4, 300 mM imidazole , pH 8
- Storage Buffer: 50 mM NaH_2PO_4, pH 8
- Running buffer for agarose-gel electrophoresis: TAE-buffer (50x): 2 M Tris/acetate, 0.05 M EDTA, pH 8.5

Buffers and solutions for SDS-polyacrylamide gel electrophoresis

Table 1: Pipetting instructions for the preparation of 2 SDS-polyacrylamide gels

	Separating gel (12.5 %)	**Stacking gel (5 %)**
acryl amide (40 %) / 0.8 % Bis[1]	3.12 mL	562.5 µL
1.5 mM Tris-HCl pH 8.8	3.75 mL	-
0.5 mM Tris-HCl pH 6.8	-	625 µL
20 % SDS	50 µL	25 µL
dH_2O	2.99 mL	3.69 mL
10 % APS (ammonium peroxodisulphate)	48 µL	25 µL
TEMED[2]	10µL	5µL

[1] *N, N'*-methylenebisacrylamide

[2] *N, N, N', N'*-tetramethylethylenediamine

- Running buffer for SDS-PAGE:
    - SDS-PAGE Buffer (5x): 15 g/L Tris, 71 g/L glycine, 2.5 g/L SDS, 1.68 g/L EDTA
- Staining solution:
    - 75 mL/L acetic acid, 500 mL/L EtOH, 2.5 g/L Brilliant Blue R
- Destaining solution:
    - 75 mL/L acetic acid, 200 mL/L EtOH
- Sample buffer (2x):
    - 1 mL Tris-HCl pH 6.8, 400 mg SDS, 300 mg DTT, 20 mg Bromphenol Blue, 2 mL glycerol, 10 mL H_2O

2.3. Agarose gel for electrophoresis

Preparing 1% agarose gel (small/large): 0.5/1.3 g of agarose were dissolved in 50/130 mL of 1x TAE buffer by heating in the microwave until no visible agarose powder remained. The gel was cooled to 60-70°C and 3/8 µL dye were added. The mixture was casted in the small/large chamber and the comb placed in its position. After approximately 20 minutes of polymerization, the gel was run at 100V for 35 min. "Gene Ruler" 1 kb DNA ladder (Thermo Scientific) was used as reference.
These gels were used as control gels and preparative gels

2.4. Standards

- DNA standard: Gene Ruler 1 kb DNA ladder (Thermo Scientific)
- Protein standard: PageRuler™ Prestained Protein Ladder (Thermo Scientific)

2.5. Kits

Miniprep:	GeneJET Plasmid miniprep Kit, #K0503 (Thermo Scientific)
PCR-CleanUp:	GeneJET PCR purification Kit, #K0701 (Thermo Scientific)
Gel purification:	Wizard® SV Gel and PCR Clean-Up System #0000225825 (Promega)
Blunt-end cloning:	CloneJET PCR Cloning Kit, #K1232 (Thermo Scientific)

2.6. Bacterial strains and phage

Table 2: Names of all bacterial strains and vector used in cloning and protein expression

Name	**Purchased from/ provided by**	**Used for**
E. coli TOP 10	Stratagene	Plasmid amplification
E. coli BL21 (DE3)	Stratagene	Protein production
E. coli Rosetta (DE3)	Stratagene	Protein production
Paenibacillus phage HB10c2	Hannes Beims, LAVES	Gene g17 amplification by PCR
Paenibacillus larvae ERIC II	Wolfgang Schuehly, University of Graz	Verification tests

2.7. Vectors

Table 3: Vectors used for the cloning and expression of g17 and tg17

Name	**Purchased from**	**Used in/for**
pETM11	Stock from institute	Gene expression
pJET1.2	Thermo Scientific	DNA storage and restriction digestion of the gene

2.8. DNA

Table 4: Genes coding for the gp17 of HB10c2 (HB_00017)

Name	**Sequence (see appendix)**	**Purchased from/ provided by**	**Description**
g17 (gene 17)	See sequence 2	Hannes Beims	Original sequence
tg17 (twist gene 17)	See sequence 3	Twist Bioscience, San Francisco	Codon optimized for the expression in *E. coli*

2.9. Primers

Primers were used for the amplification of g17 from the phage genome and for the preparation of tg17 for the Gibson cloning. All primers were ordered from Sigma Aldrich.

Table 5: Primers used in amplification PCR of g17 and tg17, restriction sites in small letters, pairing with template DNA underlined, bases inserted for necessary frameshift bold.

Name	**Sequence**	**T_M [°C] (SnapGene)**
g17 fw (NcoI)	ATATccatggCAGAAACCTACAGATTTTTTGACTCGACGG	65 °C
g17 rv (NotI)	ATATgcggccgcTACCTTTGCCCGTAAAATGAGGAGGGG	64 °C
Gibson fw (for tg17)	CCCCACTACTGAGAATCTTTATTTTCAGGGCGC**a**GCCGAAACCTATCGCTTTTTTG	58 °C
Gibson rv (for tg17)	TGGTGGTGGTGCTCGAGTGCTTAACGCTGACCATAAAAGC	58 °C

2.10. Enzymes

Table 6: Names of the enzymes used in protein mutagenesis and characterization

Enzyme	**Purchased from/ provided by**	**Used in**
Proteinase K	**thermo**	
DNA-polymerase Phusion®	Thermo Scientific, USA	Amplification
Restriction enzyme NcoI	NEB	Restriction digest
Restriction enzyme NotI	NEB	Restriction digest
Phosphatase cip	NEB	Restriction digest
Ligase T4	Thermo Scientific, USA	Blunt-end ligation, Sticky end ligation

2.11. Equipment

PCR devices:	2720 Thermal Cycler (Applied Biosystems) Gene Ampl® PCR System 9700
DNA electrophoresis equipment:	Sub-Cell® GT Agarose Gel Electrophoresis System (BIO-RAD)
Agarose gel documentation system:	Gel Doc 200 (BIO-RAD)
Electrophoresis power supply:	PowerPAc™ 300 Cell (BIO-RAD)
Protein electrophoresis equipment:	Mini-PROTEAN® 3 Cell (BIO-RAD)

Incubation shaker:	Infors HT Multitron Standard (Switzerland)
Centrifuges:	Heraeus Laborfuge 400R
	Sorvall® RC-6 Plus (Thermo Electron Corporation)
	Eppendorf Centrifuge 5810R
Sonication device:	Labsonic L (B. Braun Biotech International)
NanoDrop:	NanoDrop200 Spectrophotometer (Thermo Scientific)
Nickel-NTA	HP, 5mL (GE health care)
	FF, 1 mL (GE health care)
Spectrophotometer	Specord 205 (analytikjena)
Laminar flow fumehood:	Two 30 #EBS504019 (Faster)
Peristaltic pump:	Dynmax, RP-1 (Rainin Instruments)
pH-device:	Orion star A211 (Thermo Scientific)
Magnetic stirrer	MR 3001 K (Heidolph)
Various:	Centriprep 10kDa/30kDa MWCO (Millipore)
	Eppendorf Thermomixer comfort
	Eppendorf Thermomixer compact
	Vortex Genie 2 (Scientific Ind. Inc)
	Heating block TR-L 288 (Liebisch)
	Autoclave 5075 MLV (Tuttnauer)
	CertoClav EL (Certoclav®)

2.12. Other

Whatman filter paper

Nitrocellulose membrane

UV-lamp

3. Methods

3.1. In-silico

The genome sequence of HB10c2 (35644 bp, isolated in Germany, NCBI accession number NC_028758) and also many putative gene products are known, based on the comparison with the better characterized *Paenibacillus* phage phiIBB_Pl23 (41294 bp, isolated in Portugal, NCBI accession number NC_021865).

The sequence of the different genes in the first gene cluster (upstream) was blasted with NCBI using the blastn-tool [35]. As the RBP is known to be located in the cluster for structural genes, following the proteins tmp, Dit and Tal, only the proteins gp15, gp16 and gp17 were promising candidates [31,36]. The protein sequence of the gene products gp15 to gp19 were modelled with Phyre2 with the modelling mode set to intense [37].

Although only little protein data is available, modeling of the proteins gp15 and gp16 showed similarities to a Dit and a Tal protein respectively and gp 17 had similarities with the RBP of the *Lactococcal* phage 1358. The Following proteins gp18 and gp19 were only small and their structure did not resemble the typical structure of RBPs at all.

3.1.1. Ordering of tg17

The protein sequence was run through the GeneArt program of Thermo Scientific by Marina Toplak in order to create a DNA-sequence, which is codon-optimized for the expression in *E. coli.* Additionally, restriction sites for NcoI including the start codon and NotI downstream following the stop codon were added (see sequence 3). This sequence was ordered from Twist Bioscience (San Francisco).

3.1.2. Primer design

The primers g17 fw and g17rv were designed by Marina Toplak and Felix Schweigkofler. The fw-primer contains a NcoI site and the rv-primer contains a NotI site in order to allow the sticky-end ligation with pETM11.

The primers Gibson fw and Gibson rv were designed by Birgit Grill and Felix Schweigkofler. The primers are designed to amplify tg17 for the Gibson-ligation with the plasmid pETM11 and Gibson fw does not contain and therefore removes the NcoI site during the process of amplification.

3.2. DNA-preparation for the cloning

3.2.1. Phage-DNA isolation

The DNA of HB10c2 was isolated from a phage lysate with an unknown titer (10^5-10^8 pfu/mL), received from Nikola Vinko. The DNA was isolated using the phenol-chloroform-protocol provided by the Center of Phage Technology at the Texas A&M University [38].

3.2.2. g17-amplification from the whole-phage-genome

The primers (g17 fw + g17 rv) were suspended in ddH_2O according to the manufacturers instructions and then diluted 10fold with ddH_2O to a concentration of 10 µmol/L. The template used was the isolated DNA from the phage lysate. The quantity was chosen higher than 10 ng, as contamination with bacterial DNA was suspected.

Table 7: Pipetting instructions for the g17-amplification from the whole-phage-genome

Component	**Volume [µL]**
Q5 buffer (5x)	10
dNTPs (2 mM)	5
Primer g17 fw	5
Primer g17 rv	5
Template (15 ng/µL)	1
ddH_2O	23.5
Q5 DNA Polymerase	0.5
Total	50

Three PCR-mixes were prepared following this scheme. The samples were placed in the thermocycler and the PCR was run with 35 cycles following the program shown in Table 8.

Table 8: Temperature program for the g17-amplification

T [°C]	98	98	65	72	4
Time	2 min	20 s	30 s	10 min	∞
Step	Initial denaturation	Denaturation	Annealing	Elongation	Hold
Cycles		35			

Proper amplification of g17 was checked with an agarose gel and the PCR product was then purified with the GeneJET PCR purification Kit according to the manufacturers instructions. The purified PCR-product was stored at -20 °C.

3.2.3. Control/preparative gel electrophoresis

To check the products of the PCRs and the restriction digests and to purify specific DNA-fragments a 1% agarose gel was used as control gel or preparative gel respectively (see agarose gel for electrophoresis, 2.3). The polymerized gel was placed in an electrophoresis chamber and the chamber was filled with 1x TAE-buffer. Five parts of the DNA-sample were mixed with 1 part of purple loading dye and pipetted into a slot. For control gels 10 µL of DNA per sample (12 µl with loading dye) were used, for preparative gels not more than 24 µL of sample (30 µL with loading dye) were pipetted into one slot. 8 µL of the 1 kb gene ruler were used as DNA standard for all DNA-gels. The electrophoresis was run at 100 V and 400 mA for 35 min. After completing the run, the gel was examined under UV-light using the Gel Doc 200.

3.2.4. Plasmid production

The plasmid pETM11 was received in *E. coli* Top10 and with an undesired gene between the NcoI and NotI restriction sites. The cells containing the plasmid were plated out on LB-K and grown at 37 °C overnight. The next day a miniprep was carried out with the GeneJET Plasmid miniprep Kit (Thermo Scientific) according to the manufacturers instructions

3.2.5. Blunt-end cloning with pJET1.2

The amplified g17 as well as the ordered tg17 were blunt-end cloned into pJET1.2, in order to ensure a high quality of the restriction digest.

3.2.6. Restriction digest

The vector gained by miniprepping *E. coli* Top10, harboring the pETM11 plasmid, was digested with NcoI and NotI in order to remove the undesired gene and to linearize the backbone.
The purified PCR-product was digested with NcoI and NotI to obtain the desired insert with sticky ends.
The pJET1.2-vectors which had been ligated with g17 and tg17 were both digested with NcoI and NotI to obtain the desired insert with sticky ends. This was done to exclude possible errors, which could happen with the digest of a linear insert with only 4 bp 5’ and 3’ to the restriction sites.

Table 9: Pipetting instructions for the restriction digest

	Vector	**Insert**
Component	**Volume [µL]**	**Volume [µL]**
DNA-sample	41	42
Cutsmart Buffer 10x	5	5
NcoI	1.5	1.5
NotI	1.5	1.5
Phosphatase cip	1	0
Total	50	50

The mix was incubated at 37°C for 4 hours and then, the enzymes were inactivated by either heating them to 80°C for 2 min in the thermomixer or by directly adding the loading dye and loading them onto a preparative agarose gel.

3.2.7. tg17-amplification for Gibson-cloning

The primers (Gibson fw + Gibson rv) were suspended in ddH_2O according to the manufacturers instructions and then diluted 10fold with ddH_2O for a concentration of 10 µmol/L. As a template the original tg17 from Twist Biosciences was used.

Table 10: Pipetting instructions for the tg17-amplification for Gibson-cloning

Component	**Volume [µL]**
Q5 buffer (5x)	10
dNTPs (2 mM)	5
Primer Gibson fw	5
Primer Gibson rv	5
Template (10 ng/µL)	1
ddH_2O	23.5
Q5 DNA Polymerase	0.5
Total	50

Two PCR-mixes were created following this scheme and later pooled. The samples were placed in the thermocycler and the PCR was run with 35 cycles following the program shown in Table 11.

Table 11: Temperature program for the tg17-amplification for Gibson-cloning

T [°C]	98	98	63	72	4
Time	2 min	20 s	30 s	10 min	∞
Step	Initial denaturation	Denaturation	Annealing	Elongation	Hold
Cycles		35			

Proper amplification of g17 was checked with an agarose gel and the PCR product was then purified with the GeneJET PCR purification Kit according to the manufacturers instructions. The purified PCR-product was stored at -20 °C.

3.1. Ligation of the Gene with pETM11

3.1.1. Classical cloning

The digested inserts were sticky-end ligated with the cut pETM11 using a T4 ligase. The concentrations of both, vector and insert, were determined with nanodrop and the required volumes were calculated considering that 100 ng vector and a fivefold molar amount to insert should be used for each ligation. As the cut pET11 has roughly 5300 bp and the insert has roughly 1100 bp, amount of insert needed were 100 ng as well (Equation 1).

$$\frac{1100\ bp}{5300\ bp} * 5 * 100\ ng \approx 100\ ng$$

Equation 1

Table 12: Pipetting instructions for the Classical ligation, the volumes of vector and insert depend on their concentration

Component	**Volume [µL]**	**DNA [ng]**
Vector (pETM11)	x	100
Insert (g17 and tg17)	x	100
Ligase buffer (10x)	2	
Ligase T4	1	
ddH_2O	x	
Total	20	

The mix was incubated at room temperature for 1 h and then either the ligase was inactivated by heating the reactions to stopped by heating it up to 60 °C for 20 min or transformed into *E. coli* Top10 cells.

3.1.2. Gibson cloning

For Gibson cloning a prepared and ready-to-use cloning mix was used, provided by Birgit Grill (see: Gibson Mastermix). The volume of vector and insert were calculated the same way as for the classical cloning and both 100 ng of vector and insert were used. As insert the purified PCR-product from the Gibson-PCR (see classical cloning, 3.2.1.) was used.

Table 13: Pipetting instructions for the Gibson ligation, the volumes of vector and insert depend on their concentration

Component	**Volume [µL]**	**DNA [ng]**
Vector (pETM11)	x	100
Insert (g17 and tg17)	x	100
Gibson cloning mix	15	
ddH_2O	x	
Total	20	

The mix was incubated for 50 °C 1 h at and then either stopped by heating it up to 60 °C for 20 min or right away transformed into *E. coli* Top10 cells.

3.2. Heat-shock transformation of *E. coli* Top10, BL21 and Rosetta

For Plasmid amplification and protein production the recombinant plasmids were transformed into different *E. coli* strains. All of them had been made chemically competent with $CaCl_2$. The prepared competent cells were stored on -80 °C and thawed on ice. If the concentration of the DNA was known, roughly 100 ng were used to transform the competent cells. When the cells were transformed with the ligation product, the full 20µl of the ligation was used. The thawed cells were added to the prepared plasmid and mixed by gently pipetting up and down a few times. The Eppis were kept on ice for 15 to 30 minutes and then, heat-shocked for 2 min on 42 °C in the Thermomixer (Eppendorf Thermomixer comfort). Immediately, 900 µL LB-medium were added and for regeneration the mix was put on 37 °C (350 rpm) for 30 to 60 min. Subsequently, the cells were pelleted and resuspended in roughly a tenth of the original supernatant by pipetting. The resuspension was plated out for single colonies on LB-agar

containing 50 µg/mL kanamycin for pETM11-transformants, 100 µg/mL ampicillin for pJET1.2-transformants and 50 µg/mL kanamycin + 20 µg/mL chloramphenicol for Rosetta pETM11-transformants.

3.3. Verification of the construct

Upon transformation, single colonies were obtained. A number of those colonies was streaked out on LBK-Agar and inoculated on 37 °C over night. The next day, a miniprep was performed according to the manufacturers instructions. The concentration of the plasmid was determined with Nanodrop and a control restriction digest (see restriction digest, 3.2.6.) was done either with NcoI and NotI to cut the gene from the plasmid or with PstI, which has 3 internal cutting sites in tg17 and none in pETM11. If the results were or seemed positive, the plasmid was sequenced using the service of LGC Genomics.

3.4. Protein production in BL21 and Rosetta

3.4.1. Protein production

One colony of the *E. coli* BL21-transformants and one of the *E. coli* Roseatta-transformants respectively were used to inoculate 50 mL of LB-medium containing 50 µg/mL kanamycin and additional 20 µg/mL chloramphenicol in the case of Rosetta. The cultures were incubated at 37 °C and 150 rpm over night.

For both “ONCs” 3 Erlenmeyer flasks containing 800 mL of LB and the corresponding antibiotics (50 mg/L kanamycin and 50 mg/L kanamycin + 20 mg/L chloramphenicol, respectively) were prepared and inoculated with 12 mL of the ONC. The flasks were shaken at 37°C (140 rpm) until the OD_{600} was 0.75 and 1 respectively. The expression was induced by adding 0.1 mM IPTG and then the flasks were incubated on 20°C (precooled shaker, 140 rpm) for 19 hours.

The cells were harvested by centrifuging the cultures at 4°C (5000 rpm) for 10 min in 400 mL beakers with the rotor F10S 6x500y. The pellets of *E. coli* BL21 and *E. coli* Rosetta were gently resuspended in 25 mL sonication lysis buffer (50 mM NaH_2PO_4, 150 mM NaCl, 10 mM imidazole) each and transferred to two separate 50 mL falcon tubes and vortexed for 2 min. Subsequently, both solutions were transferred to separate sonication rosettes and lysed by ultrasonication with Labosonic L (Braun) with 120 W in two steps of 5 min each. The lysates were centrifuged at 18,000 rpm at 4 °C for 40 min at 4°C with the rotor F10S 8x50. Afterwards, the supernatants were each transferred to a 50 mL falcon tube.

3.4.2. Protein purification via Ni-NTA affinity chromatography

For protein purification a Ni-NTA-column with a volume of 1 mL was used. The system was operated with a peristaltic pump (Dynmax, Rainin Instruments) . With the speed “20” the tubes and the column were rinsed with the tenfold column-volume of water (no air must reach the column) and afterwards, with the tenfold column-volume of ultrasonication lysis buffer. The column was then loaded with cleaned lysate at a speed of “15”. Then, it was washed with the tenfold column-volume of washing buffer. The flow-through and the wash were collected separately. Subsequently, the proteins were eluted with elution buffer. The flow-through, the wash and each eluate fraction were collected separately in a falcon tube. The elution fractions collected were around 4 mL each.

From each elution fraction, the wash, the flow-through and the loaded supernatant 10 µL were taken and mixed with 200 µL biuret reagent. The decreasing color of the fractions showed when all protein had been eluted.

The column was cleaned by washing with the tenfold volume of water, 100mM EDTA and water respectively and was then prepared by rinsing it with the threefold column-volume of 100 mM $NiSO_4$.

3.4.3. Purification and concentration

Imidazole was removed from the protein solution by dialysis. The protein solution was transferred to a dialysis tube and incubated at 4 °C overnight with 2 L dialysis buffer (50 mM NaH_2PO_4 and 150 mM NaCl, pH 8).

To enrich the protein, it was centrifuged in a centricon (cut-off: 30 kDa) at 4 °C in a precooled centrifuge with 3500 rfc. The centricon was filled with the protein solution and centrifuged with another centricon as counterweight for 10 min each step. The counterweight was refilled with water after each step and the centricon was filled with protein solution consecutively, until 1.5 mL of the enriched protein solution remained in the upper part of the centricon. This solution was stored at 4 °C.

The protein concentration was determined with the Lambert-Beer-Law (Equation 2) by using the extinction at 280 nm (Nanodrop, program a280) and the mass of the protein, which was calculated with the Prodparam tool of Expasy [39].

$$c = \frac{E}{e * d}$$

Equation 2

3.5. SDS-polyacrylamide gel electrophoresis

The pellet, the supernatant, the flow-through, the wash and the fractions 2, 3 and 5 of both *E. coli* Rosetta and BL21 were analyzed by SDS-PAGE.

3.5.1. Preparation of SDS-polyacrylamide gels

The SDS-polyacrylamide gels was prepared by first pipetting the ingredients of the separating gel listed in Table 1 into a disposable tube. As soon as TEMED and APS are added, working quickly is necessary, as the acrylamide starts to polymerize. 5 mL of the gel were pipetted into the slit between the mounted glass plates, so that it reached the upper line of the plates. Butanol was pipetted on top of it until it overflowed to guarantee an evener surface.

After 30 minutes, the stacking gel, containing the ingredients of Table 1, was pipetted into the remaining slit to fill it completely and then the comb was shoved into it. After 30 more minutes the gel was solidified and could be loaded. If it was not used immediately, it was stored wet and cool (in wet paper towels at 4 °C)

3.5.2. Gel-electrophoresis

The gel was placed in the electrophoresis chamber (Biorad) and the chamber was filled with electrophoresis buffer (running buffer). 10 µL of each sample (pellet, supernatant, flow-through, wash and fractions 2 and 3) were mixed with 10 µL of 2x loading dye and heated to 95 °C for 10 min. Then, 5 µL each were pipetted into the slots of the gel. The gel was run at 120 V for 20 min and 180 V for 45 min.

3.6. Verification of the binding to *P. larvae*

To verify that the isolated protein is the binding agent of the phage, this property was tested empirically with the newly developed **Ni-NTA affinity chromatography test** and a variation of the RBP-assay of Simpson et al. [40]. The affinity chromatography test was developed relying on the strength of the binding of the His-Tag and the putative binding properties of the produced protein. The *P. larvae* used for the tests were provided by Nikola Vinko, grown according to Beims [13].

3.6.1. Affinity chromatography test

ONCs were prepared by inoculating 5 mL BHI-medium in a 50 mL falcon tube with a colony of *P. larvae* and *Bacillus subtilis* respectively and incubated at 37 °C (150 rpm) over night. *B. subtilis* has a chloramphenicol resistance. When the bacteria reached a sufficiently high OD_{600},

300 µL of the protein solution was mixed with the respective ONCs and filled up to 1 mL with BHI medium. The final theoretical OD_{600} of the ONC alone was 0.4 (e.g. 100µl with OD_{600} of 4). The solutions were shaken with at 37 °C (150 rpm) for 30 min. After that, 9 mL of the ultrasonication lysis buffer are added to each of the two solutions.

The instrumental setup for the affinity chromatography test is the same as the one described under 3.4.2. for the purification of the protein with a c-column, but the volume of the column was chosen with 1 mL instead of 5 mL. The tubes and columns were prepared accordingly and the full 10 mL were loaded onto the two columns with a peristaltic pump. The Proteins were eluted according to the protein purification protocol. The fraction sizes were chosen with 4 mL. 10 µL of the flow-through, the wash and all the fractions were tested with 200 µL Biuret reagent for protein to decide which fractions should be further analyzed. Fractions 1 and 2 were chosen.

100 µL of both original protein-bacteria-solutions, 100 µL of flow-through, wash and the fractions 1 and 2 respectively were plated out. 1 mL of each of the same samples was spun down in the centrifuge and the not visible pellet resuspended in a tenth of the supernatant. The same was done with dilutions of 1:10 of all of these samples, to cover a wider range of possible concentrations. The *P. larvae* samples were plated out on Columbia Sheep Blood Agar. The *B. subtilis* were plated out on LBC-Agar and incubated at 37 °C. The cfu were counted the following day.

3.6.2. Colony assay

A colony of each of the bacteria to be tested was suspended by pipetting in sterile water. 2 µL of the solution were pipetted onto LBK-agar or LB-agar depending on the resistance to create a bacteria spot, following the instructions in Table 14. The spots were labelled in order to recognize them later. For this test *E. coli*, Rosetta, BL21 and Top10 (Top10-pos), containing the confirmed construct, *E. coli* Top10 (Top10-neg) with pETM11 containing the peptide-binding protein of *Rhodobacter sphaeroides* (NCBI accession number CP015287) from three different colonies and not transformed *E. coli* Top10 were chosen.

Table 14: Spots of different *E. coli* strains per agar plate

Plate nr.	Rosetta	BL21	Top10-pos	Top10-neg	Top10
1	3	3	3	9	-
2	3	3	3	9	-
3	-	-	7	-	-
4	-	-	-	7	-
5	-	-	-	-	7

The plates were kept on 37 °C over night and the colonies were transferred onto nitrocellulose membranes, cut to plate-sized circles, by gently pressing the membranes on the agar plates for some seconds, until the membrane clearly touched all of the agar-surface. Membrane 1, 3 and 4 on LBK-IPTG (LB + 25 µg/mL Kanamycin + 0.4 mM IPTG), Membrane 2 on LBK and Membrane 5 on LB-IPTG. The membranes were incubated at 37 °C overnight colony side up on agar plates. The membranes were transferred onto Whatman paper, pre-saturated with a protein extraction solution. According to the protocol of Simpson et al. the solution should be the bacterial protein extraction reagent B-PER, containing 1x protease inhibitor cocktail, DNAse I (1 U/mL) and lysozyme (500 µg/mL). Instead Y-PER (yeast protein extraction reagent) with lysozyme was used, as the protease inhibitor (1 M PMSF in DMSO) precipitated. The cells were lysed at room temperature for 1 h and the membranes were transferred to Petri dishes and blocked with 5% skim milk (Roth) in PBST at room temperature for 1 h. The spots were gently cleaned using a Kimwipe and remaining colonies were exposed to UV-light for 15 minutes to kill any living cells. The membranes were incubated in 500 mM NaCl at 4 °C overnight and sterilized by UV-light. An ONC (37 °C) of *P. larvae* was prepared and the cell number was linked to the optical density by measuring the probes in the cell counter. The OD_{600} of the ONC was measured and the approximate cfu/mL was calculated . The membranes in the Petri dishes were sterilized by UV-light for 20 min and then incubated with the *P. larvae*-solution at room temperature and 50 rpm for 30 min. They were then washed by incubating them with PBST at room temperature and occasional manual shaking 3 times for 10 minutes. The membranes were placed on MYPGP-Agar to allow for the growth of *P. larvae* and kept on 37°C over night. The membranes were then matched to the spots on the initial agar plate.

4. Results

4.1 DNA-results

The amplification of g17 from the whole phage genome with forward and reverse primers containing NcoI and NotI restriction sites, respectively, worked fine and resulted in a clear band of around 1100 bp, which is the expected size of the gene. The subsequent ligation and transformation resulted in a dozen colonies, which was in the same range as the number of transformants obtained in the control experiment without the insert.

Miniprepping of three transformed colonies and digesting the isolated plasmids with NcoI and NotI resulted in a band of approximately 1200 bp for one of the three colonies and sequencing confirmed that the sequence was not the right insert g17.
A second PCR with the same primers resulted in bands of 1100 bp as well and again the ligation failed; none of three clones contained the right insert.

The blunt-end cloning of the PCR-product g17 with pJET1.2 was successful and the restriction digest of the cloned plasmids resulted in a backbone band around 3.2 kbp and a band at 1100 bp. The ligation of pETM11 and two of the inserts failed, even though the best ligation resulted in 50 transformants. Restriction digest of 9 clones with NcoI and NotI again showed again either no additional bands or bands at 1200 bp and sequencing confirmed that those sequences did not correspond to g17.

The blunt-end cloning of the synthetized tg17 with pJET1.2 was successful, too, and the restriction digest of the recombinant plasmids resulted in a backbone band at 3.2 kbp and a band at 1100 bp (Figure 2), but again the ligation did not lead to the right constructs (Figure 3).

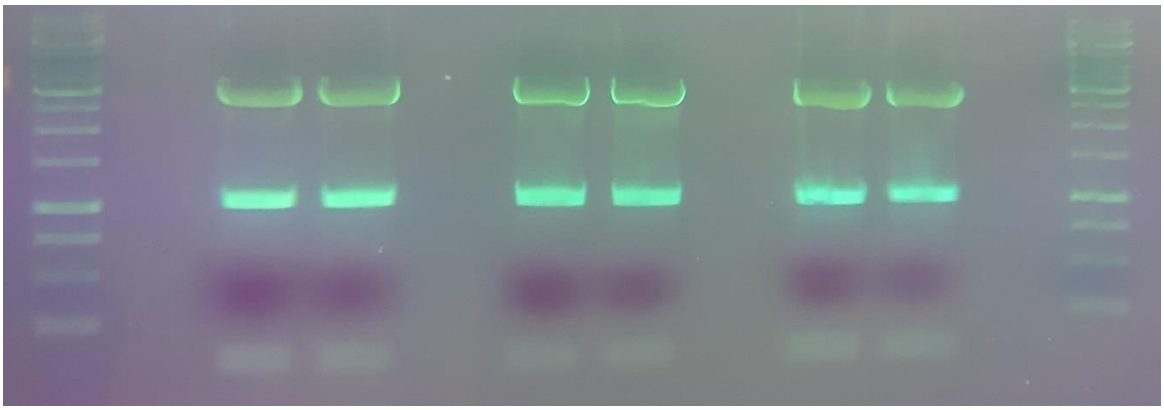

Figure 2: Preparative gel of the NcoI-NotI restriction digested blunt-end cloned construct pJET1.2-tg17. The insert band is at 1100 bp, the backbone-band at 3.2 kbp. The band below 250 bp are small fragments, as pJET1.2 itself contains both a NcoI and a NotI restriction site and the bands are DNA-sequences located between the sites of the tg17 and the sites of pJET1.2

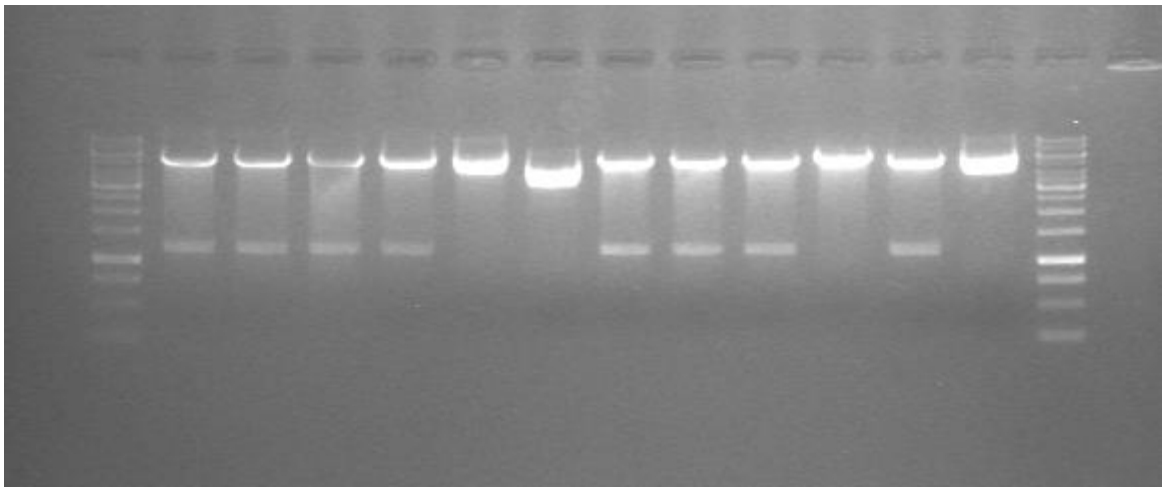

Figure 3: Control gel of the NcoI-NotI restriction digested putative pETM11-tg17 constructs (tg17 was cut from pJET1.2). Plasmids showing no insert probably are cut pETM11 vectors, which ligated with themselves. Plasmids containing an insert of 1200 bp most probably were neither cut nor ligated and still contained their original insert (peptide-binding protein of *R. sphaeroides*)

The ligation was repeated by Sara Hopf and Vera Wasserbacher, using the same vector and insert, but two different ligases, again with no success.

Marina Toplak repeated the experiment with the same protocols and with a different ligase and using a newly cut tg17 from pJET1.2. The transformation resulted in fewer colonies than the ligation control but nevertheless all of the colonies contained pETM11, correctly ligated with tg17 (Figure 4), confirmed by sequencing.

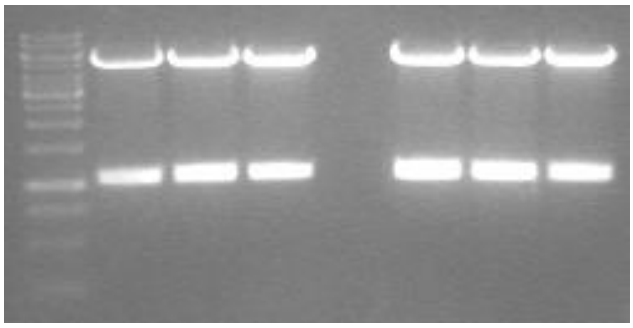

Figure 4: Control gel of the NcoI-NotI control restriction digested putative pETM11-tg17 clones, performed by Marina Toplak. All 6 tested clones have an insert with the size of about 1100 bp.

The PCR for obtaining amplified tg17 for the Gibson-cloning was successful. Upon transformation many colonies were obtained, but fewer than expected. 3 and later 9 more clones were isolated and digested with PstI instead of NcoI/NotI, in order to obtain a more clear band pattern. Three of the latter nine colonies showed the expected bands (Figure 5). Sequencing of one of the constructs confirmed the proper insertion of tg17.

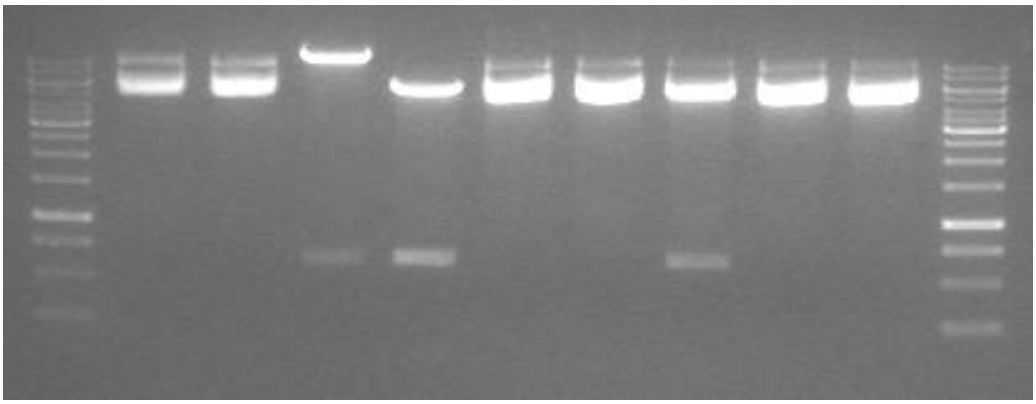

Figure 5: Control gel of the PstI restriction digest of putative pETM11-tg17 clones obtained by Gibson cloning (tg17 has been amplified with Gibson primers). Visible are bands at 600 bp and the backbone band at 5 kbp. The expected band at 40 bp can hardly be seen, as its brightness is less than a tenth of the brightness of the 600-bp-band. The seemingly very large backbone-band of slot 3 may correspond to a pETM11-pETM11-tg17 construct.

4.2 Protein production

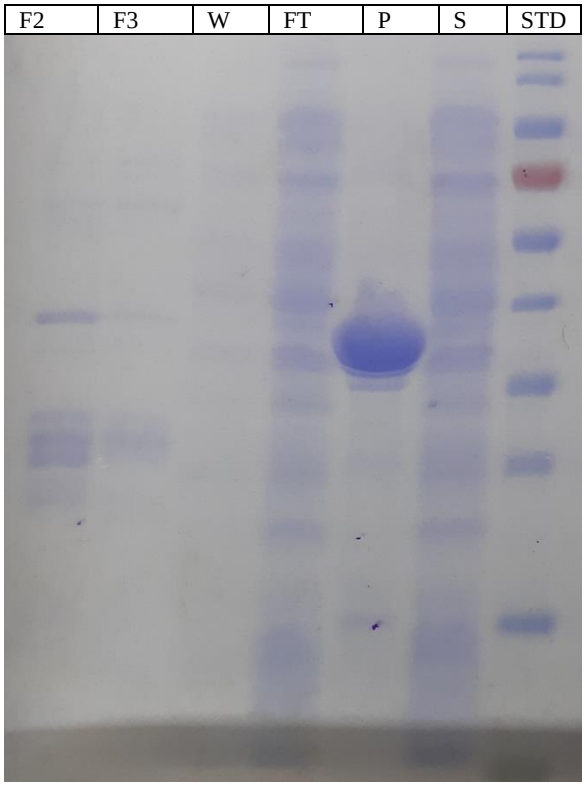

Figure 6: SDS-PAGE of samples of the p17-production with Rosetta: fraction 2 (F2), fraction 3 (F3), wash (W), flow-through (FT), pellet (P), supernatant (S) and protein standard (STD). 10 µL of each sample were mixed with 10 µL of 2x loading dye and heated to 95 °C for 10 min. Then, 5 µL each were pipetted into the slots of the gel. The gel was run at 120 V for 20 min and 180 V for 45 min.

The IPTG-induced overexpression produced a high amount of a protein with the approximate mass of 40 kDa (Figure 6), which is close to the expected mass of 43,4 kDa for p17 (calculated with Expasy pI/Mw). The majority of the overexpressed protein, however, clearly remained in the pellet and only a small fraction may have been in in the supernatant. In the fractions 2 and 3 a thin band shows the same motility as p17 in the pellet. This could be p17, but it remains unclear. The protein in F2 and F3 did neither precipitate during dialysis nor after concentrating using centricons.

The extinction E of the concentrated protein sample was 9.266. The concentration of the protein (calculated as stated in 3.5.3.) was about 146 µmol/L. If the uppermost bands of fraction 2 and fraction 3 are p17, the concentration of p17 in the protein sample could be estimated to be a fifth of this number (30µmol/L).

4.3 Verification tests

4.1.1. Affinity chromatography test

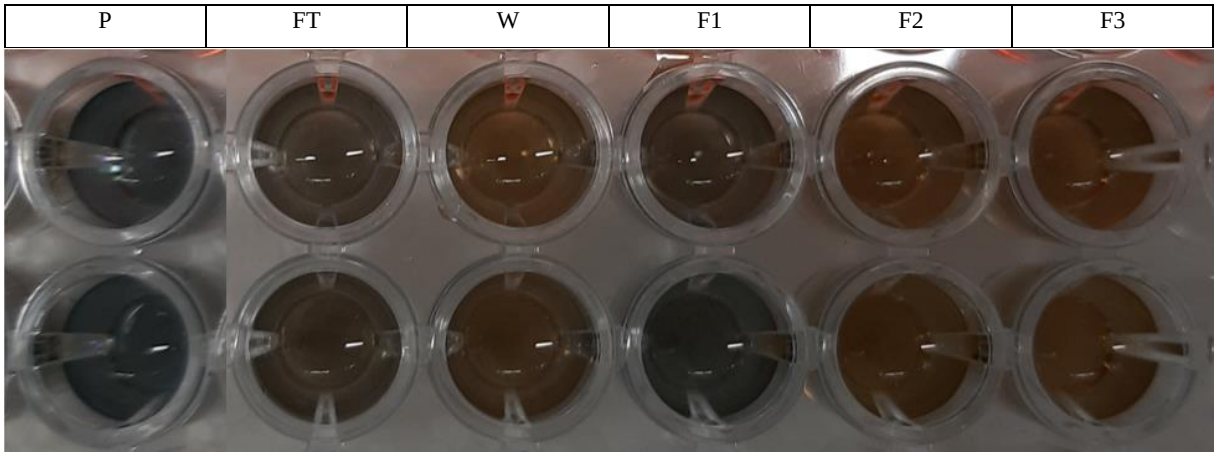

Figure 7: The Biuret test of protein-bacteria-solution (P), flow-through (FT), wash (W), fractions 1, 2, 3 (F1, F2, F3), upper line: *P. larvae*, lower line: *B. subtilis*. 10 µL of each sample were mixed with 200 µL of Biuret reagent, a blue color shows presence of proteins. The pure sample (P) and the flow-through contain protein as well as the first fraction.

The biuret-test confirmed the proper working of the column loading, washing and elution. The proteins in the sample have partially been loaded onto the column and eluted mainly in the first fraction.

Table 15: The approximate cell density in the different fractions of the affinity chromatography test. The fractions were collected and 100 µL, each, were plated out for single colonies on CSA. Samples are: protein-bacteria-solution (P), flow-through (FT), wash (W), fractions (F1, F2)

	P	FT	W	F1	F2
cfu/mL	44000	4300	700	500	60
Fractin volume (mL)	(10)	11	13,5	4	4
cfu	440000	47300	10850	2000	240
Total cfu	440000	60390			

The number of *P. larvae*-colonies decreases in the course of the elution process. Only a tenth (60,000 of 440,000) of the bacteria could be recovered viable in all of the liquid flown through. This might be due to the general conditions of the test or special properties of P. larvae.

The *B. subtilis* plates did not show countable single colonies, rather almost a bacterial lawn, but the number of colonies visibly decreased over the five consecutive fractions with the highest density on the plate with the protein-bacteria-solution and the lowest number on the plate with F2.

The cfu/mL has been calculated as the average cfu/mL of each fraction, which reduces the usability of the data for an exponential model. If plotted against the volume, which had passed through the column, the cfu/mL decreased clearly. An exponential model (Figure 8) describes the reduction properly. The cfu/mL of fraction 1 is slightly higher than the expected value, but the data are too doubtful to allow for any conclusion. The data indicate that solely the flow of the liquid was responsible for the "elution" of the bacteria.

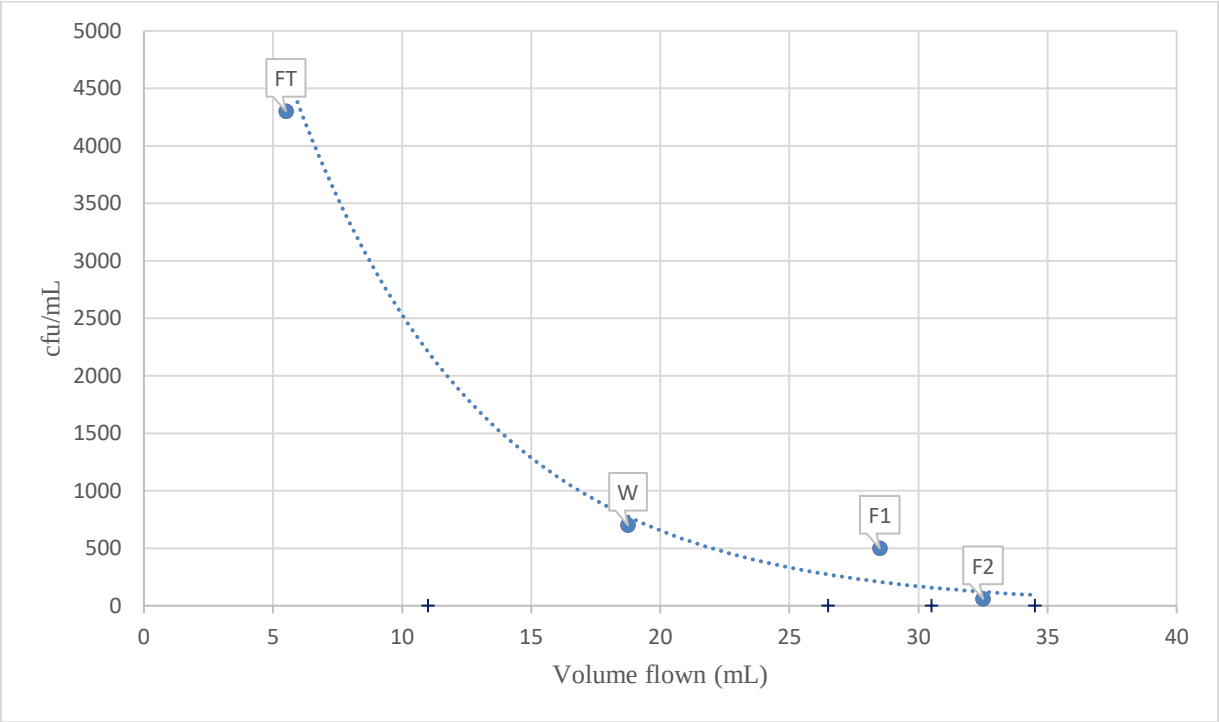

Figure 8: The cfu/mL of *P. larvae* in the course of the liquid flow in mL. The x-values have been chosen to represent the volume, of which half of the respective fraction volume had been collected. Full volumes of each fraction are marked with a plus on the axis. The dotted lines represent an exponential (blue) and a linear (grey) trendline, calculated by Excel 2016.

4.1.2. Colony assay

The incubation of the membranes with *P. larvae* resulted mainly in randomly growing colonies with little visible order. Close inspection of the plates allows for the to the following interpretations of the results.

Three of the membranes were covered with a bacterial lawn, two showed only few colonies. Clearly free areas of different sizes were visible on all three intensely covered membranes, however, they could in no case be matched to any original *E. coli*-colony. The second type of spots observed share at least some of the following properties:

- They have the size of the original *E. coli*-spots
- They show a ring of colony-free membrane
- They show a slightly stronger intensity of the bacterial lawn
- They have a small spot of colony-free membrane in the middle

All of these types of spots could be matched to original *E. coli*-spots, although not all of the original spots could be matched to these *P. larvae*-spots. IPTG seems not to have a significant influence on the appearance of the spots.

Detailed information about Plate 1 and 2 can be found as picture caption.

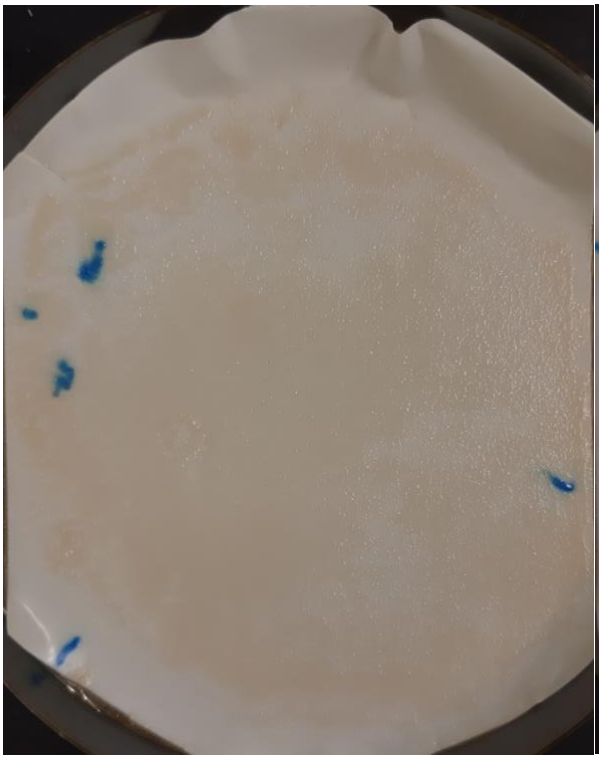

Figure 9: Plate 1, 3 sections *E. coli* with and 3 sections without RBP, IPTG. The membrane is completely covered with bacteria, but a few spots look different. One spot is almost free of bacteria, but could not be matched to an *E. coli* colony on the original plate. 4 out of 4 Rosetta colonies, 3 out of 3 BL21 colonies and 1 out of 3 Top10 colonies – each containing tg17 – could be matched to noticeable formations of the bacterial lawn on the membrane but to none of the negative controls (nHis(1)8, nHis(1)9 and pETM11).

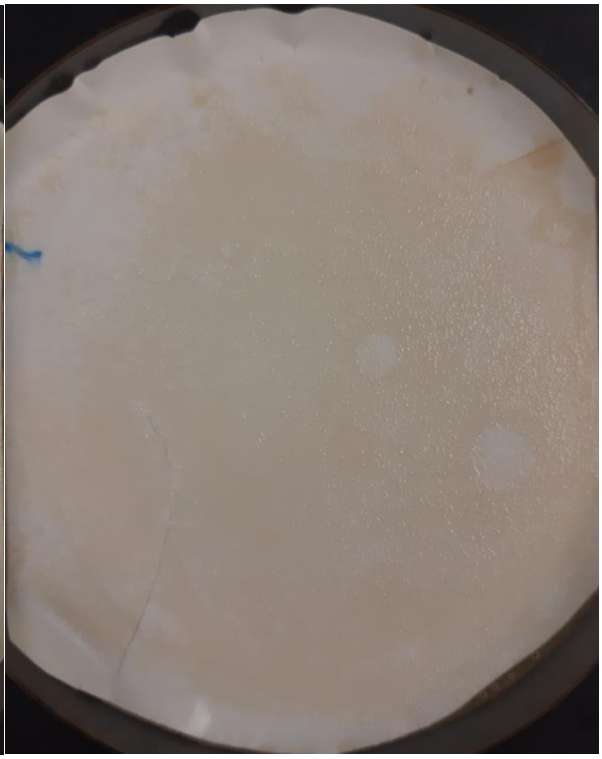

Figure 10: Plate 2, 3 sections *E. coli* with and 3 sections without RBP, no IPTG. The membrane is completely covered with bacteria, but a few spots look differently. Two spots are almost free of bacteria, but could not be matched to an *E. coli* colony on the original plate and one of them is bigger as the original spots. 1 out of 3 Rosetta colonies, 3 out of 3 BL21 colonies and 1 out of 3 Top10 colonies – each containing tg17 – can be matched to the described spots of P. larvae on the plate but to none of the negative controls (nHis(1)8, nHis(1)9 and pETM11) can.

Plate 3: Top10 with pETM11-p17, IPTG

Many single colonies were spread over the plate, but they did not form a bacterial lawn. The spots with the lysed original colonies were visible as yellow staining of the membrane and the colonies accumulated around these spots. When transferring the colonies onto the membrane,

the membrane was dragged over the agar plate, which could explain some additional colonies grown between the actual spots.

Plate 4: Top10 with pETM11, without p17, IPTG
The membrane was completely covered with bacteria. Two spots were almost free of bacteria, but they could not be matched to any of the colonies on the original plate.

Plate 5: Top10 without pETM11, IPTG
Few colonies were randomly spread over the membrane. The spots with the lysed original colonies were visible as yellow staining of the membrane, but colonies would hardly accumulated around these spots.

5. Discussion

5.1 SDS-PAGE analysis

p17 was produced but insoluble, which can be due to many reasons. It might be traced back to general difficulties of protein production or to specific reasons for RBPs. Specific reasons might be the following.

The assembly of the tail of *Myoviridae* and *Siphoviridae* is assisted by phage-encoded chaperons and sometimes also the assembly of the baseplate [41–43], [Uniprot: P17173]. The formation of the baseplate does not rely on the presence of an assembled or assembling tail and therefore the lack of tail tube proteins should not be the reason for any misfolding of the overexpressed RBP [44]. Even though the baseplate of HB10c2 is rather simple in comparison to the 150 proteins of the baseplate of phage T4 [45], it might still rely on a protein to start the assembly process, as it is the case for other phages [43]. Potentially, the RBP has problems folding correctly without the rest of the baseplate and thus aggregates.

Simpson et al. suggest that some RBPs might need their own chaperones to assemble correctly but do not provide any evidence for that [40]. Considering the tight management of information on the phage genome, this, however seems not very likely. However, if this really is the case for HB_00017 it will be difficult to produce significant amounts of correctly folded proteins when only expressing tg17 in *E.coli* BL21 or Rosetta.

The folding process of RBPs might indeed be difficult, as models show that the three protomers are tightly intertwined and not just accumulated [30,32]. This can can be a reason why the protein produced in *E. coli* Rosetta and BL21 was insoluble. On the other hand, if that really was the case, all or other RBPs, which are trimeric too, probably should have the same assembling issues. However, that has not been observed.

The solubility might therefore depend on factors difficult to determine. The natural property of structural phage proteins to self-assemble leads to higher proportions of insoluble proteins. Even when that is the case, small fractions of correctly folded RBP might be found in the cytoplasm [46].

However, it can be assumed that the difficulties producing soluble RBP should be minor, as all phage proteins are naturally produced in the cytoplasm of the host and self-assemble to the native structure. Some researchers experienced no problem producing the RBP of other phages [28]. Usual methods like producing the protein in a special expression strain or fusing it with

other proteins might therefore prove to be sufficient and will be the next approach to hopefully obtain native protein.

Small amounts of protein in the fractions F2 and F3 of the Ni-NTA affinity chromatography, visible as thin bands after the SDS-PAGE (Figure 6), could indeed be p17 in a soluble form. They seem to have a similar motility as the proteins in the large band containing p17. Several other researchers too were only able to produce little amounts of soluble RBPs [46]. However, it is not possible to tell solely from the SDS-PAGE, whether the thin band stems from correctly folded, soluble p17. The concerning fraction needs to be examined by mass spectrometry or by other means to clarify if the thin band contains p17 and to estimate the effort it will take to produce a higher amount of soluble protein.

5.2 Affinity chromatography test

The affinity chromatography test aims to evaluate the binding of the his-tagged RBP to the receptors on the bacterial surface. During the incubation and the subsequent retention of the marked bacteria on the Ni-NTA-column a protein-receptor interaction should become visible. Whether the bacteria were retarded can be checked by plating out the fractions and counting the CFU. Ideally, two negative controls should be conducted, one with the target bacterium, but no RBP, and one with RBP, but a different bacterium. Only the latter control was done due to resource and time limitations.

The protein solution used was gained from the production of p17 in *E. coli*, where most or all of the p17 was insoluble (see SDS-PAGE analysis). It was assumed that the small band in the fractions F2 and F3 showing a similar motility as the p17 in the pellet could be the right protein. However, this is not sure and needs to be confirmed by mass spectrometry.

The biuret test of the eluates confirmed that the elution of the protein worked properly, with unspecific proteins being eluted in the flow-through and the his-tagged proteins being eluted in the three fractions (Figure 7). The cfu/mL however decreased in all consecutive eluates for both *P. larvae* and *B. subtilis*, although the trend for *B. subtilis* could only be estimated visually, as the cell number on every plate was too high for counting. The decrease of the cfu/mL follows an exponential trend (Figure 8). The elution of *P. larvae* seems to be only dependent on the flow and not on the composition of the eluent and thus is independent from the elution of the protein.

The two different bacteria seemed to have had a different cfu/mL, which is possible as the concentration of both was only estimated via photospectrometry at 600 nm (OD_{600}) by assuming

that the bacteria have similar optical properties. The different number of colonies could also be the result of a lower viability of *P. larvae* or different growth characteristics. Control bacteria would need to be related much closer to the target bacteria, ideally stemming from the same genus.

P. larvae as a gram positive bacterium was plated out on a complex medium containing nalidixic acid, which is an antibiotic targeting gram negative bacteria and was not contaminated. Antibiotics might be necessary as the setup is not sterile, but by cleaning the tubes and columns with NaOH or other agents might be sufficient to keep the contamination to a minimum, which would only slightly affect the outcome of the experiment.

The tests shows severe weaknesses and the data gained does neither lead to the conclusion that this is a viable test for the presence of RBPs nor that p17 is the desired RBP of HB10c2. This could be due to the flaws of the concept or the protein solution lacking significant amounts of correctly folded RBP.

5.3 Colony assay

According to Simpson et al. the RBP-producing colonies should result in equally sized spots of the target bacterium and the rest of the membrane should be colony-free [40]. This is not what was observed, rather plates 1, 2 and, 4 were fully covered with bacteria and plates 3 and 5 contained a few seemingly randomly distributed colonies. Only a closer look revealed, that some hardly visible circles matched the original p17-producing colonies perfectly. The expected patterns were there, but could visually hardly be distinguished from the background noise and seemed to be random errors at first.

The interpretation of the circles in the bacterial lawn as indicators for the presence of the RBP is possible, but must be treated with caution., even more so, as the spots should attract more bacteria rather than less. However, the circles match the original p17-producing colonies perfectly and must be linked to them in some way. *Only* colonies producing p17 – although not all of them – produced the circles, while none of the control colonies producing the peptide-binding protein of *R. sphaeroides* left any mark in the bacterial lawn.

Surprisingly, the presence of IPGT seemed to have no distinguishable effect, as the seemingly RBP-indicating circles appeared on both plates, the one with and the one without IPTG added. (Figure 9, Figure 10 respectively). Conclusions from the exact number of the circles on both plates cannot be drawn, as the results are too weak for any quantitative interpretation. Either

the amount of protein produced by leaky transcription is sufficient or the circles do not result from p17 production, which, however, would contradict the rest of the findings.

Three of the membranes were covered with a bacterial lawn, while only few colonies grew on plates 3 and 5. A possible explanation for this observation is, that those membranes were inoculated last and the *P. larvae* solution was perhaps not properly mixed when the ONC was added to the skim milk. These plates did not show any clear spots, but the few colonies formed were distributed almost randomly, with only very slight accumulations of the colonies around the original RBP-producing colonies.

As the blocking of the membranes with skim milk and/or the washing did not work properly, the washing step should be intensified and lower cell numbers could be tested. Apart from that, the number of initial spots should be reduced.

5.4 Outlook

Several adaptations and improvements of the conducted experiments are possible.

p17 can be produced in different strains of *E. coli*, designed for the production of barely soluble proteins. p17 can be fused with another protein to improve the solubility. To remove potentially hydrophobic parts of the protein the receptor binding domain could be produced without the rest of the protein, as it was done previously for other RBPs [30]. This could even improve the stability of the protein [30].

The band of putative soluble p17 in the fractions F2 and F3 can be excised from the SDS-polyacrylamide gel and tested by mass spectrometry to determine whether the band really contains p17 or not.

The affinity chromatography test should be developed with an easier to handle bacterium, as *P. larvae* is growing slowly and is pathogenic for bees, which is why it should not be transformed with a plasmid granting antibiotic resistance without need.

The colony assay should be set up in other ways and more fitting to the actual problem, as the setup of the experiment by Simpson et al. deliberately aims a the selection of colonies from a genetic library and is more difficult than necessary.

6. Conclusion

Leaving the difficulties of interpreting the visual results aside, the colony experiment hints at the binding characteristics of the protein p17, which reinforces the assumption that HB_00017 codes for the receptor binding protein of the *Paenibacillus* phage HB10c2. However, further experiments will be needed to clearly confirm this hypothesis.

References

[1] D. P. Barrett, ‘History of American foulbrood in Michigan: period 1927–1954.’, *Glean. Bee Cult.*, no. 83, pp. 460–461, 1955.

[2] E. Genersch, ‘American Foulbrood in honeybees and its causative agent, Paenibacillus larvae’, *J. Invertebr. Pathol.*, vol. 103, pp. S10–S19, Jan. 2010, doi: 10.1016/j.jip.2009.06.015.

[3] Y.-W. Chen, C.-H. Wang, J. An, and H. Kai-Kuang, ‘Susceptibility of the Asian honey bee, *Apis cerana* , to American foulbrood, *Paenibacillus larvae larvae*’, *J. Apic. Res.*, vol. 39, no. 3–4, pp. 169–175, Jan. 2000, doi: 10.1080/00218839.2000.11101038.

[4] E. Genersch, ‘Reclassification of Paenibacillus larvae subsp. pulvifaciens and Paenibacillus larvae subsp. larvae as Paenibacillus larvae without subspecies differentiation’, *Int. J. Syst. Evol. Microbiol.*, vol. 56, no. 3, pp. 501–511, Mar. 2006, doi: 10.1099/ijs.0.63928-0.

[5] L. Haseman, ‘How long can spores of American foulbrood live.’, *Am Bee J*, no. 101, pp. 298–299, 1961.

[6] J. F. Bamrick and W. C. Rothenbuhler, ‘RESISTANCE TO AMERICAN FOULBROOD IN HONEY BEES. 4. RELATIONSHIP BETWEEN LARVAL AGE AT INOCULATION AND MORTALITY IN A RESISTANT AND IN A SUSCEPTIBLE LINE.’, *J. Insect Pathol.*, no. 3.4, p. 381, 1961.

[7] C. J. Brødsgaard, W. Ritter, and H. Hansen, ‘Response of in vitro reared honey bee larvae to various doses of Paenibacillus larvae larvae spores’, *Apidologie*, vol. 29, no. 6, pp. 569–578, 1998, doi: 10.1051/apido:19980609.

[8] A. Lindström, S. Korpela, and I. Fries, ‘The distribution of Paenibacillus larvae spores in adult bees and honey and larval mortality, following the addition of American foulbrood diseased brood or spore-contaminated honey in honey bee (Apis mellifera) colonies’, *J. Invertebr. Pathol.*, vol. 99, no. 1, pp. 82–86, Sep. 2008, doi: 10.1016/j.jip.2008.06.010.

[9] ‘Faulbrutquelle Glascontainer?’, *Allg. Dtsch. Imkerztg.*, no. Heft 11, p. 13, 2011.

[10] J. Kochansky, D. A. Knox, M. Feldlaufer, and J. S. Pettis, ‘Screening alternative antibiotics against oxytetracycline-susceptibleand -resistant *Paenibacillus larvae*’, *Apidologie*, vol. 32, no. 3, pp. 215–222, May 2001, doi: 10.1051/apido:2001123.

[11] V. A. Govan, M. H. Allsopp, and S. Davison, ‘A PCR detection method for rapid identification of Paenibacillus larvae’, *Appl. Environ. Microbiol.*, vol. 65, no. 5, pp. 2243–2245, May 1999.

[12] S. Ghorbani-Nezami, L. LeBlanc, D. G. Yost, and P. S. Amy, ‘Phage Therapy is Effective in Protecting Honeybee Larvae from American Foulbrood Disease’, *J. Insect Sci.*, vol. 15, no. 1, p. 84, 2015, doi: 10.1093/jisesa/iev051.

[13] H. Beims *et al.*, ‘Paenibacillus larvae-Directed Bacteriophage HB10c2 and Its Application in American Foulbrood-Affected Honey Bee Larvae’, *Appl. Environ. Microbiol.*, vol. 81, no. 16, pp. 5411–5419, Aug. 2015, doi: 10.1128/AEM.00804-15.

[14] D. P. Stahly, A. M. Alippi, N. Bakhiet, C. F. Campana, C. C. Novak, and R. Cox, ‘PPL1c, a Virulent Mutant Bacteriophage Useful for Identification of Paenibacillus larvae subspecies larvae’, *J. Invertebr. Pathol.*, vol. 74, no. 3, pp. 295–296, Nov. 1999, doi: 10.1006/jipa.1999.4893.

[15] C. Stamereilers *et al.*, ‘Genomic Analysis of 48 Paenibacillus larvae Bacteriophages’, *Viruses*, vol. 10, no. 7, p. 377, Jul. 2018, doi: 10.3390/v10070377.

[16] A. Oliveira, L. D. R. Melo, A. M. Kropinski, and J. Azeredo, ‘Complete Genome Sequence of the Broad-Host-Range Paenibacillus larvae Phage phiIBB_Pl23’, *Genome Announc.*, vol. 1, no. 5, pp. e00438-13, 1/5/e00438-13, Sep. 2013, doi: 10.1128/genomeA.00438-13.

[17] A. Oliveira, M. Leite, L. D. Kluskens, S. B. Santos, L. D. R. Melo, and J. Azeredo, ‘The First Paenibacillus larvae Bacteriophage Endolysin (PlyPl23) with High Potential to Control American Foulbrood’, *PLOS ONE*, vol. 10, no. 7, p. e0132095, Jul. 2015, doi: 10.1371/journal.pone.0132095.

[18] L. LeBlanc, S. Nezami, D. Yost, P. Tsourkas, and P. S. Amy, ‘Isolation and characterization of a novel phage lysin active against *Paenibacillus larvae* , a honeybee pathogen’, *Bacteriophage*, vol. 5, no. 4, p. e1080787, Oct. 2015, doi: 10.1080/21597081.2015.1080787.

[19] T. S. Brady, B. D. Merrill, J. A. Hilton, A. M. Payne, M. B. Stephenson, and S. Hope, ‘Bacteriophages as an alternative to conventional antibiotic use for the prevention or treatment of Paenibacillus larvae in honeybee hives’, *J. Invertebr. Pathol.*, vol. 150, pp. 94–100, Nov. 2017, doi: 10.1016/j.jip.2017.09.010.

[20] P. Hyman and S. T. Abedon, ‘Bacteriophage Host Range and Bacterial Resistance’, in *Advances in Applied Microbiology*, vol. 70, Elsevier, 2010, pp. 217–248.

[21] N. Sternberg and D. Hamilton, ‘Bacteriophage P1 site-specific recombination’, *J. Mol. Biol.*, vol. 150, no. 4, pp. 467–486, Aug. 1981, doi: 10.1016/0022-2836(81)90375-2.

[22] N. A. Campbell and J. B. Reece, *Biology*, 6th ed. San Francisco: Benjamin Cummings, 2002.

[23] C. A. Suttle, ‘Viruses in the sea’, *Nature*, vol. 437, no. 7057, pp. 356–361, Sep. 2005, doi: 10.1038/nature04160.

[24] A. A. Cisek, I. Dąbrowska, K. P. Gregorczyk, and Z. Wyżewski, ‘Phage Therapy in Bacterial Infections Treatment: One Hundred Years After the Discovery of Bacteriophages’, *Curr. Microbiol.*, vol. 74, no. 2, pp. 277–283, Feb. 2017, doi: 10.1007/s00284-016-1166-x.

[25] I. Takeuchi, K. Osada, A. H. Azam, H. Asakawa, K. Miyanaga, and Y. Tanji, ‘The Presence of Two Receptor-Binding Proteins Contributes to the Wide Host Range of Staphylococcal Twort-Like Phages’, *Appl. Environ. Microbiol.*, vol. 82, no. 19, pp. 5763–5774, Oct. 2016, doi: 10.1128/AEM.01385-16.

[26] Z. J. Storms and D. Sauvageau, ‘Modeling tailed bacteriophage adsorption: Insight into mechanisms’, *Virology*, vol. 485, pp. 355–362, Nov. 2015, doi: 10.1016/j.virol.2015.08.007.

[27] A. Garen, ‘THE FIRST TWO STEPS OF THE INVASION OF HOST CELLS BY BACTERIAL VIRUSES. II’, *J. Exp. Med.*, vol. 94, no. 3, pp. 177–189, Sep. 1951, doi: 10.1084/jem.94.3.177.

[28] D. M. Tremblay *et al.*, ‘Receptor-Binding Protein of Lactococcus lactis Phages: Identification and Characterization of the Saccharide Receptor-Binding Site’, *J. Bacteriol.*, vol. 188, no. 7, pp. 2400–2410, Apr. 2006, doi: 10.1128/JB.188.7.2400-2410.2006.

[29] H. Ando, S. Lemire, D. P. Pires, and T. K. Lu, ‘Engineering Modular Viral Scaffolds for Targeted Bacterial Population Editing’, *Cell Syst.*, vol. 1, no. 3, pp. 187–196, Sep. 2015, doi: 10.1016/j.cels.2015.08.013.

[30] S. Ricagno *et al.*, ‘Crystal Structure of the Receptor-Binding Protein Head Domain from Lactococcus lactis Phage bIL170’, *J. Virol.*, vol. 80, no. 18, pp. 9331–9335, Sep. 2006, doi: 10.1128/JVI.01160-06.

[31] A. S. A. Dowah and M. R. J. Clokie, ‘Review of the nature, diversity and structure of bacteriophage receptor binding proteins that target Gram-positive bacteria’, *Biophys. Rev.*, vol. 10, no. 2, pp. 535–542, Apr. 2018, doi: 10.1007/s12551-017-0382-3.

[32] C. Cambillau, ‘Bacteriophage module reshuffling results in adaptive host range as exemplified by the baseplate model of listerial phage A118’, *Virology*, vol. 484, pp. 86–92, Oct. 2015, doi: 10.1016/j.virol.2015.05.015.

[33] C. Koç *et al.*, ‘Structure of the host-recognition device of Staphylococcus aureus phage ϕ11’, *Sci. Rep.*, vol. 6, no. 1, p. 27581, Jun. 2016, doi: 10.1038/srep27581.

[34] G. F. Hatfull, ‘Bacteriophage genomics’, *Curr. Opin. Microbiol.*, vol. 11, no. 5, pp. 447–453, Oct. 2008, doi: 10.1016/j.mib.2008.09.004.

[35] NCBI Resource Coordinators, ‘Database resources of the National Center for Biotechnology Information’, *Nucleic Acids Res.*, vol. 41, no. D1, pp. D8–D20, Nov. 2012, doi: 10.1093/nar/gks1189.

[36] R. Bielmann *et al.*, ‘Receptor binding proteins of Listeria monocytogenes bacteriophages A118 and P35 recognize serovar-specific teichoic acids’, *Virology*, vol. 477, pp. 110–118, Mar. 2015, doi: 10.1016/j.virol.2014.12.035.

[37] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, and M. J. E. Sternberg, ‘The Phyre2 web portal for protein modeling, prediction and analysis’, *Nat. Protoc.*, vol. 10, p. 845, May 2015.

[38] Center of phage technology (Texas A&M University), ‘Protocol for Phage DNA Extraction with Phenol:Chloroform: https://cpt.tamu.edu/wordpress/wp-content/uploads/2018/09/Phage-DNA-Extraction-by-PhenolChloroform-Protocol.pdf’.

[39] E. Gasteiger *et al.*, ‘Protein Identification and Analysis Tools on the ExPASy Server’, in *The Proteomics Protocols Handbook*, Humana Press, 2005, pp. 571–607.

[40] D. J. Simpson, J. C. Sacher, and C. M. Szymanski, ‘Development of an Assay for the Identification of Receptor Binding Proteins from Bacteriophages’, *Viruses*, vol. 8, no. 1, Jan. 2016, doi: 10.3390/v8010017.

[41] ‘Viral tail assembly ~ ViralZone page’. [Online]. Available: https://viralzone.expasy.org/3955?outline=complete_by_species. [Accessed: 03-Dec-2019].

[42] M. Siponen *et al.*, ‘Crystal Structure of ORF12 from *Lactococcus lactis* Phage p2 Identifies a Tape Measure Protein Chaperone’, *J. Bacteriol.*, vol. 191, no. 3, pp. 728–734, Feb. 2009, doi: 10.1128/JB.01363-08.

[43] L. M. Kozloff and M. Lute, ‘Identification of bacteriophage T4D gene products 26 and 51 as baseplate hub structural components.’, *J. Virol.*, vol. 52, no. 2, p. 344, Nov. 1984.

[44] A. A. Aksyuk and M. G. Rossmann, ‘Bacteriophage Assembly’, *Viruses*, vol. 3, no. 3, pp. 172–203, Feb. 2011, doi: 10.3390/v3030172.

[45] V. A. Kostyuchenko *et al.*, ‘Three-dimensional structure of bacteriophage T4 baseplate’, *Nat. Struct. Mol. Biol.*, vol. 10, no. 9, pp. 688–693, Sep. 2003, doi: 10.1038/nsb970.

[46] M. Mondigler, R. T. Vögele, and K. J. Heller, ‘Overproduced and purified receptor binding protein pb5 of bacteriophage T5 binds to the T5 receptor protein FhuA’, *FEMS Microbiol. Lett.*, vol. 130, no. 2–3, pp. 293–300, Aug. 1995, doi: 10.1111/j.1574-6968.1995.tb07734.x.

Registers

List of Tables

List of Figures

Sequences

Sequence 1

>YP_009195207.1 hypothetical protein HB_00017 [*Paenibacillus* phage HB10c2]

MAETYRFFDSTDTDERLYTADEFAEYFRQVLSDGIFNGGTNLKVESTGKNMETYIQPGYAWL
QGYLYAVKDTKLNLQHPYPHATLDRIDRVVVRLDKRLDHRYVRAFVKEGTPSTTPSPPALTR
NDNVFEISLAQVKIVKGKSYIEAYQITDERLNKTVCGIVNSLIQADTTTIFNQFQKWFESRT
ADFEKEWKEWLEKMKDQGGGKFGVTSVNGKTGDVILMAKHVGAPSINDLRAYALKGEPAGQY

TPTFLNGWYVQAGEVKGVCYYKDQFGYVHLYGTCSGTKTEFGTPLFNLPAGFRPSGVIRVGC
LMIDFADYSRSIQFLGVYPSGEVLIESYGLPGFVSFSIFPSSFYGQR

Sequence 2

>g17 from the gene HB_00017 (*Paenibacillus* phage HB10c2)

TTGGCAGAAACCTACAGATTTTTTGACTCGACGGATACGGACGAGCGACTCTACACGGCAGA
CGAATTTGCGGAATATTTCAGGCAAGTCCTTAGTGACGGCATATTTAACGGCGGAACAAATC
TTAAGGTAGAAAGTACGGGCAAAAACATGGAAACGTATATTCAACCGGGCTATGCCTGGTTA
CAGGGGTACTTGTATGCCGTCAAGGATACGAAGCTAAATCTACAGCATCCCTATCCACATGC
CACGCTAGACCGCATTGACCGAGTCGTAGTAAGGTTGGACAAACGTCTAGATCATCGATACG
TAAGGGCTTTTGTCAAAGAAGGGACACCCTCGACTACTCCCAGCCCTCCTGCATTAACACGT
AACGATAACGTGTTTGAAATCAGTCTAGCACAAGTAAAGATTGTAAAAGGTAAATCTTATAT
CGAAGCCTATCAAATCACAGACGAGAGGCTCAACAAAACCGTTTGTGGCATTGTGAACTCTC
TTATACAAGCCGACACGACTACGATTTTCAACCAATTCCAGAAATGGTTTGAAAGCCGTACA
GCAGACTTTGAAAAAGAGTGGAAAGAATGGCTGGAGAAGATGAAAGATCAAGGGGGAGGGAA
ATTTGGTGTAACATCCGTCAATGGTAAGACCGGGGATGTAATACTGATGGCCAAACATGTAG
GGGCCCCAAGTATAAATGATCTCAGAGCATACGCCCTGAAAGGCGAACCCGCGGGGCAGTAC
ACGCCTACGTTTTTGAATGGTTGGTACGTACAAGCAGGTGAAGTCAAAGGGGTTTGCTACTA
CAAGGATCAATTCGGCTACGTCCACCTTTATGGCACTTGCTCAGGGACTAAAACTGAGTTTG
GCACACCTTTATTCAACCTCCCGGCTGGGTTTCGTCCAAGTGGTGTAATCCGCGTTGGTTGT
CTGATGATCGACTTTGCAGACTACTCCAGATCTATTCAATTTCTAGGTGTATATCCCAGTGG
AGAGGTATTGATAGAAAGTTATGGGTTACCCGGTTTTGTCTCGTTTAGCATTTTCCCCTCCT
CATTTTACGGGCAAAGGTAG

Sequence 3

>tg17 codon optimized sequence of HB_00017 (capital letters – coding sequence, bold - restriction sites for NcoI and NotI)

atat**ccATGG**CCGAAACCTATCGCTTTTTTGATAGCACCGATACCGATGAACGTCTGTATAC
CGCAGATGAATTTGCCGAATATTTTCGTCAGGTTCTGAGTGATGGTATCTTTAATGGTGGCA
CCAATCTGAAAGTTGAAAGCACCGGCAAAAACATGGAAACCTATATTCAGCCTGGTTATGCA
TGGCTGCAGGGTTATCTGTATGCAGTGAAAGATACAAAACTGAATCTGCAGCATCCGTATCC
GCATGCAACCCTGGATCGTATTGATCGTGTTGTTGTTCGTCTGGATAAACGTCTGGATCATC
GTTATGTTCGTGCCTTTGTTAAAGAAGGCACCCCGAGCACCACACCGAGTCCGCCTGCACTG
ACCCGTAATGATAATGTTTTTGAAATTTCTCTGGCCCAGGTGAAAATCGTGAAAGGCAAAAG
CTATATTGAGGCCTATCAGATTACAGATGAACGCCTGAATAAAACCGTTTGCGGTATTGTTA
ATAGCCTGATTCAGGCAGATACCACCACCATTTTTAACCAGTTTCAGAAATGGTTTGAAAGC
CGCACCGCAGATTTTGAAAAAGAATGGAAAGAATGGCTGGAAAAAATGAAAGATCAAGGCGG
TGGTAAATTTGGTGTTACCAGCGTTAATGGTAAAACCGGTGATGTTATTCTGATGGCCAAAC
ATGTTGGTGCACCGAGCATTAATGATCTGCGTGCCTATGCACTGAAAGGTGAACCGGCAGGT
CAGTATACCCCGACCTTTCTGAATGGTTGGTATGTTCAGGCAGGCGAAGTTAAAGGTGTTTG
CTACTATAAAGATCAGTTCGGCTATGTGCATCTGTATGGCACCTGTAGCGGCACCAAAACCG
AATTTGGTACACCGCTGTTTAATCTGCCTGCAGGTTTTCGTCCGAGCGGTGTTATTCGTGTT
GGTTGTCTGATGATTGATTTCGCAGATTATAGCCGTAGCATTCAGTTTCTGGGTGTTTATCC
GAGTGGTGAAGTTCTGATTGAAAGCTATGGTCTGCCTGGTTTTGTTAGCTTTAGCATTTTTC
CGAGCAGCTTTTATGGTCAGCGTTAA**gcggccgc**atat