“The Possibilities of Silk”
A silk protein guide.

Written by :

University of Copenhagen : 2022 iGEM team

Ecole Polytechnique Fédérale de Lausanne : 2022 iGEM team

University of Life Sciences and Natural Resources (BOKU-Vienna): 2022 iGEM team

06/2022-10/2022

INTRODUCTION

As silk proteins are such novel tools in synthetic biology, working with it was quite difficult for our teams at the beginning of our projects. So the iGEM teams Copenhagen, EPFL and Vienna all working with silks, started having regular meetings, to share protocols, experiences, and struggles and to help each other out. Eventually, we realized how valuable the collected knowledge could be to future iGEM teams and other research groups, so we decided to write it down and create a digital repository with all silk-related sequences, protocols, and tutorials that we used throughout our project. Additionally, we made Instagram posts, informing other iGEM teams and the general public on possible applications of silk.

UCPH - NETLANTIS

Design & Characterization

Dragline silk produced by spiders is one of the strongest natural materials to exist and it is mainly made up of structural proteins called spidroins. These spidroins consist of N-terminal and C-terminal domains and a central part consisting of tandem repeats of a certain amino acid sequence. These sequences are rich in alanine to glycine to form the crystalline and amorphous parts of the fiber respectively.

We came across many research articles whose authors could successfully produce recombinant spider silk proteins and spin them into fibers by mimicking the conditions of the spider’s silk gland where the fibers are formed naturally. One such protein that caught our eye was the minispidroin, which is a recombinant protein formed by combining the N-terminal of *E.australis* Masp1 (Major ampullate spidroin 1), C-terminal of *A.ventricosus* Misp (Minor ampullate spidroin) and the repetitive part of *E.australis* Masp. This chimeric protein has been designed such that it has high solubility and high pH sensitivity which are 2 crucial factors when it comes to spinning.

We chose *E.coli* as our chassis since it would allow us to obtain high yields of the protein to enable spinning. It is important to note that spidroins are highly repetitive in nature which makes it difficult to synthesize the DNA sequences. So, in order for ease of synthesis, we split the N-terminal and C-terminal of the minispidroin in 1 plasmid and the middle repetitive part in another plasmid and used Golden Gate Cloning to assemble the plasmid with the full sequence of the minispidroin. This protein will be referred to as 2rep from here on.

It has been shown that the nature and number of repeats in the spidroin plays a vital role in determining the mechanical properties of the dragline silk. Essentially this means that the strength of the fiber increases with more repeats. So, in another similar cloning experiment, we introduced the repetitive region twice between the N-terminal and C-terminal domains to form another protein which will be referred to as 4rep.

With the goal of producing the first-ever biocomposite fiber for industrial fishing applications, we decided to split the constitutive elements of the fiber into 3 modules - core of the fiber, protective layer and functionalities. All of these proteins were produced in the BL21(DE3) strain of E.coli.

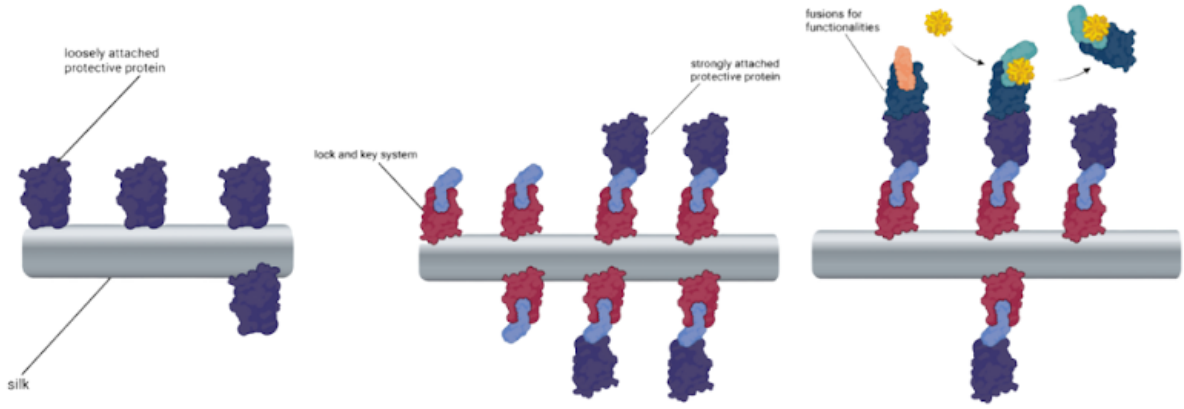


Figure 1: Conceptualization of our fiber, from left to right in order of complexity. Left) a silk fiber with a loosely bound protective protein. Middle) a silk fiber where the protective protein is covalently bound to the silk with a lock and key system (such as the Tag/Catcher system, more later). Right) a silk fiber with the addition of functionalization modules, these could be for example detection modules..

Module 1

This module represents the core fiber that we decided to base our nets on. We selected silk proteins since silk is a well known material with remarkable strength, light weight and softness which are all important characteristics for a material used to make fishing nets.

Silk proteins are usually made of non-repetitive N-terminal and C-terminal domains and a repetitive central part consisting of tandem repeats of a certain amino acid sequence. The center is characterized by their intrinsic disordered structure, since they contain very repetitive motifs (e.g. poly-A, poly-G). These repeats are generally in an alpha-helical structure before the fiber is formed but when the organism (e.g. spider or silk worms) spins the proteins into a fiber, the alpha-helices shift to beta-sheets leading to the formation of a network of H-bonds. Further, the structurally ordered N and C terminus enable macroscale protein interactions through disulphide and salt bridges.

We also plan to produce the core silk proteins with a SnoopTag so that these proteins can be spun into fibers and when these fibers are coated with another protein containing the SnoopCatcher, a spontaneous isopeptide bond is formed between the Tag and the Catcher and hence, the functional protein forms a protective layer on the fiber.

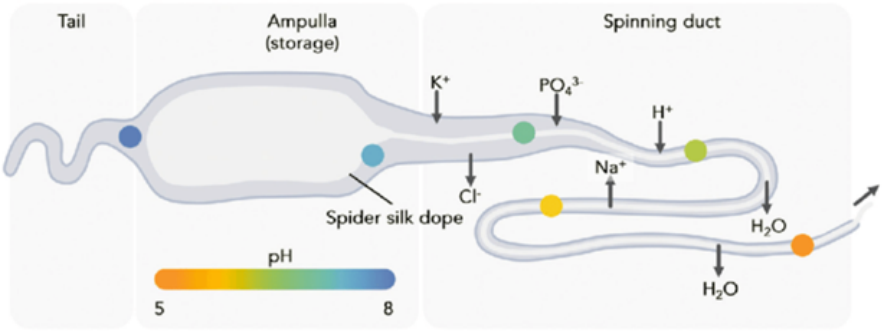


Figure 2: The spider organelle for spinning (“spinneret”) and physiological changes that occur across the spinning duct. These changes cause the structural changes in the proteins and make them interact to form a fiber. (Whittall et al., 2021)

The composite parts we produced for this module are as follows:

1. Minispidroin-2rep (BBa_K4247004, BBa_K4247007)
2. Minispidroin-4rep (BBa_K4247011, BBa_K4247012)
3. MAKI-marine minispidroin (BBa_K4247027)
4. Aneroin (BBa_K4247017)

Minispidroin-2rep and 4rep

It is well known that the solubility and pH sensitivity have a huge effect on the N- and C-terminus of the spidroin which thus affects the formation of fibers. It has been found that the N-terminus of MaSp1 (Major ampullate spidroin 1) from Euprosthenops australis, shows extremely high solubility and pH sensitivity whereas the C-terminus has low solubility and is inert to pH changes and vice versa for the MiSp (Minor ampullate spidroin) of Araneus ventricosus. So, minispidroin, a recombinant protein was produced by combining the N-terminal of E.australis Masp1, C-terminal of A.ventricosus Misp and the repetitive part of E.australis Masp. This chimeric protein was designed for optimal spinning since it has high solubility and high sensitivity to pH changes which are two factors that are crucial for fibre spinning. In our constructs, Minispidroin-2rep has one copy of the repetitive region whereas Minispidroin-4rep has two copies. These proteins were of particular interest to us because they are small and practical, but still display notable strength. Their solubility helps with achieving high concentrations of the proteins, while pH sensitivity facilitates conformational changes in the protein (Andersson et al., 2017).

MAKI-marine minispidroin

The core of our project is all about building a fiber that lasts in the harsh marine environment, but silks can have some drawbacks when it comes to resistance to salts and water.

It has been shown that spider webs from terrestrial spiders undergo structural changes with humidity wherein high humidity causes supercontraction. Supercontraction is a phenomenon where when spider silk is exposed to water, water infiltrates the fiber and causes it to reduce in length by nearly half. Major ampullate silk proteins (MaSp) contain a lot of GPGXX motifs where X can be any amino acid from a small set of amino acids. These motifs form the non-crystalline fractions of the spider silk and when the silk is in a dry state, hydrogen bonds keep these non-crystalline fractions parallel to the fiber -axis whereas when the silk is wet, these hydrogen bonds are disrupted which causes a loss of orientation and drives the shrinking and thickening of the fiber.

Hence, supercontraction hinders the use of spiders for underwater applications. However, there are certain spiders in nature that can produce silk in water such as Argyroneta aquatica (freshwater) and Desis marina (marine). Desis marina spiders construct retreats with their silk for protection from tides and pressure. Further, they can trap air in their retreat and remain submerged for up to 19 days.

Aneroin

In 2013, Yang and colleagues discovered a new silk-like protein in the sea anemone *Nematostella vectensis*. This sea-star anemone lives along the coasts of England and the USA, and has been used for a long time as a model organism for the phylum Cnidaria. The protein was an hypothetical protein found in the model genome when Yang et al. (2013) discovered it. It’s characterized by its localisation in the nematocysts of the tentacles (the harpoon-like weapons of Cnidarians) and by the repetitive motif GPGNTGYPGQ (Yang et al., 2013).

This protein has high levels of glycine and proline, the GXX repeat motif of collagens and the GPGXX repeat motif of spider silks, suggesting its involvement in prey capture and potential for fiber formation. Considering its properties, we decided to include it as a possible protein that can be used for the core fiber.

Module 2

This module represents the protective layer of the fiber. We selected the mussel foot protein MFP151. MFP151 is a hybrid protein between MFP5 and MFP1, and has been successfully produced and purified in the literature. Since collagen and silks have peculiar similarities, and since byssus lasts long periods of time in the marine environment, we decided to include them in our bio-composite material.

Mussels have the ability to attach themselves to various surfaces underwater by permanent adhesion. This adhesion is facilitated by their byssus, which is secreted from their foot. The byssus comprises a bundle of threads and at the end of each thread, there is an adhesion plaque containing a water-resistant adhesive that enables the mussel to anchor itself to surfaces. Several types of foot proteins have been characterized and each of them have a

different function as per their location in the byssus. Of these, MFP3 and MFP5 are found in the distal end of the byssus.

MFP3 and MFP5 have peculiar DOPA groups that are obtained by an hydroxylation reaction catalyzed by the enzyme tyrosinase (supported by a copper cofactor). DOPA groups are well known for their adhesive and cohesive properties thanks to covalent bonds, hydrogen bonds, metal-ligand complexes, Michael-type addition compounds, and quinhydrone charge-transfer complexes (Guvendiren et al., 2009). Furthermore, they have also been shown to possess antibiotic properties. Overall, these proteins could help make our fibers more strong and durable.

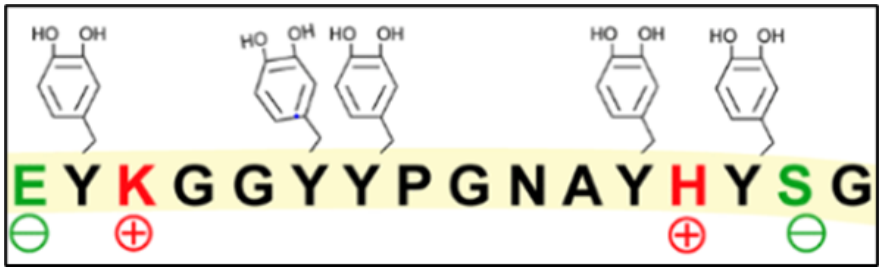


Figure 3: Representation of the DOPA groups in the MFPs, with the hydroxylations in position 3 and 4 of the tyrosine residues.

For this protein, we therefore also included a system for carrying out this post translational modification. While mfp151 can become sticky and thus be linked (non-covalently) to the fiber, when it oxidizes, it would lose most of the stickiness, thus potentially being lost and not be an effective fiber protection. So, the MFP151 proteins were also produced with a SnoopCatcher so that they can be used to coat the fibers.

Another protein that we considered for this module is an RFP called mCherry with a SnoopCatcher. We designed this protein with the aim of seeing for ourselves the efficacy of the SnoopTag-Catcher system for coating (e.g. grafting) the fibers with the proteins of Module 2. Since mCherry is a fluorescent protein, the efficacy of the system can be quantified by measuring the fluorescence intensity. Unfortunately, we did not get to test this out entirely, but we did produce the mCherry_SnoopCatcher proteins.

Module 3

This module represents the potential functionalizations that can be used to make our bio-composite fiber viable for various diverse applications. For example, how cool would it be if our net could be bioluminescent to keep marine mammals away from getting caught in the net? Or repel all the crabs that fishermen always have to remove after a fishing trip? In many years, possibly also include a way to tune the net for target-specific fishing so that they do not catch any non-target species. Unfortunately, the duration of the project was not sufficient for us to test this development. Click on functionality design to see our ideas!

Linking the modules

In order to enable the linkage of the modules, we have incorporated a SnoopTag-Catcher system in our proteins such that the silk proteins can be expressed with a SnoopTag and spun into fibres and when these fibres are coated with another protein containing the SnoopCatcher, a spontaneous isopeptide bond is formed between the Tag and the Catcher and hence, the functional protein forms a protective layer on the fiber.

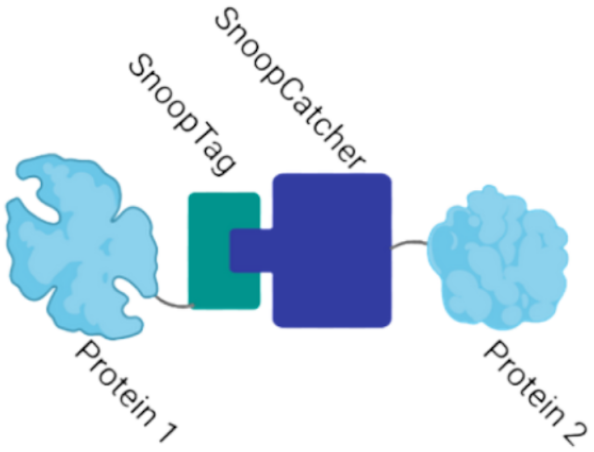


Figure 4: A schematic representing how the SnoopTag-Catcher system can be used to bind one protein to another.

3D Models

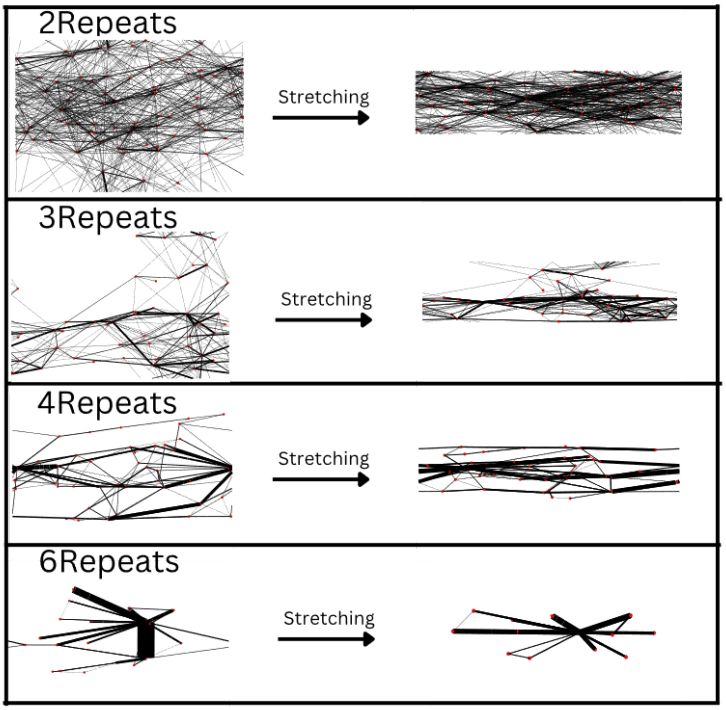


Figure 5: Node-bridge diagrams of silk fiber-networks for spider silk proteins with 2-, 3-, 4-, and 6-repeats of hydrophobic motif. Shown at a time frame after spinning (left) and after stretching (right).

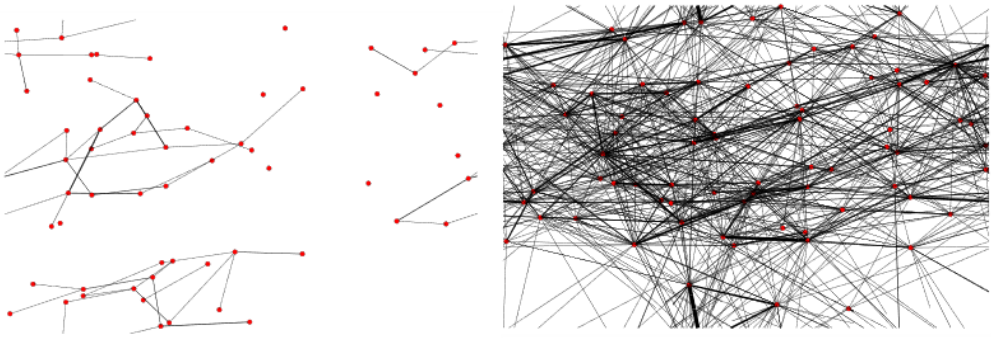


Figure 6: Comparison between node-bridge diagram of silk fiber-network for spider silk with 2 repetitions of the hydrophobic motif at a concentration of 15% (left) and 30% (right) after applying forces, which mimics spinning.

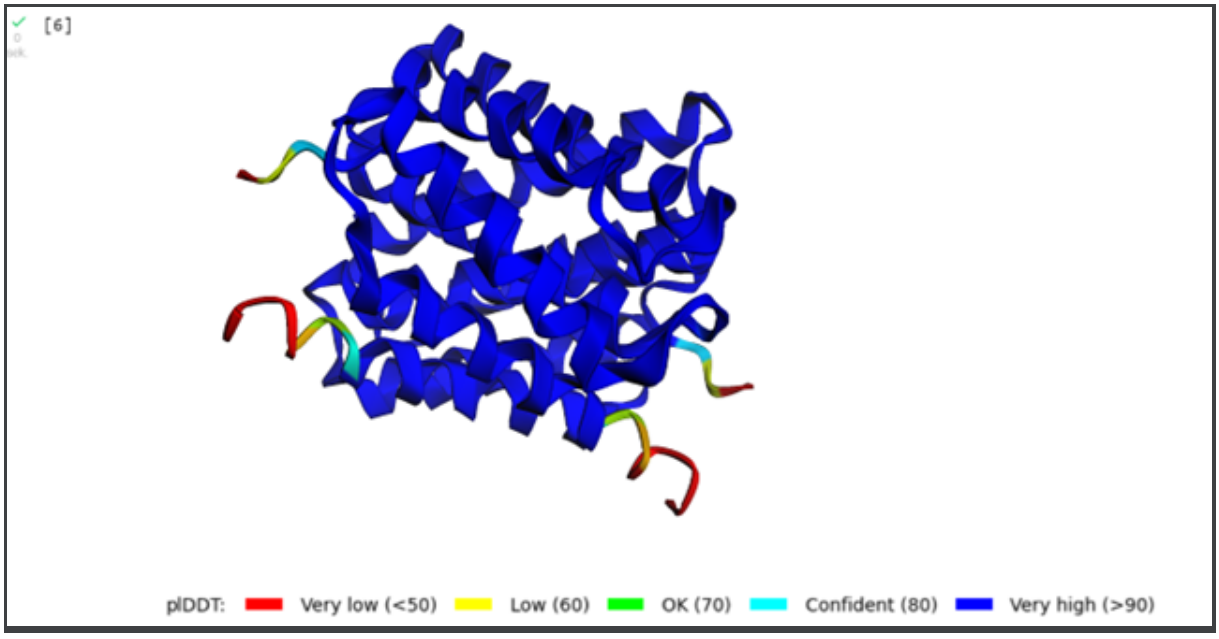


Figure 7: Modeling of the C-terminus domain of 2rep/4rep.

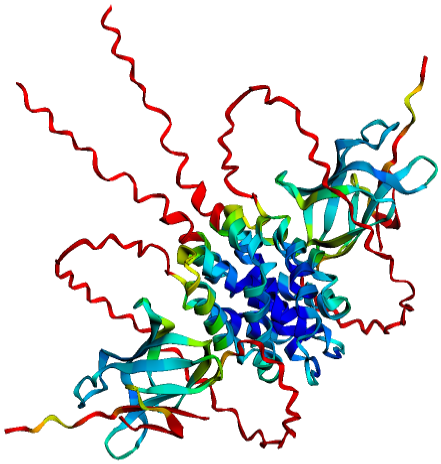


Figure 8: Modeling with the Tag and Catcher system

Final constructs

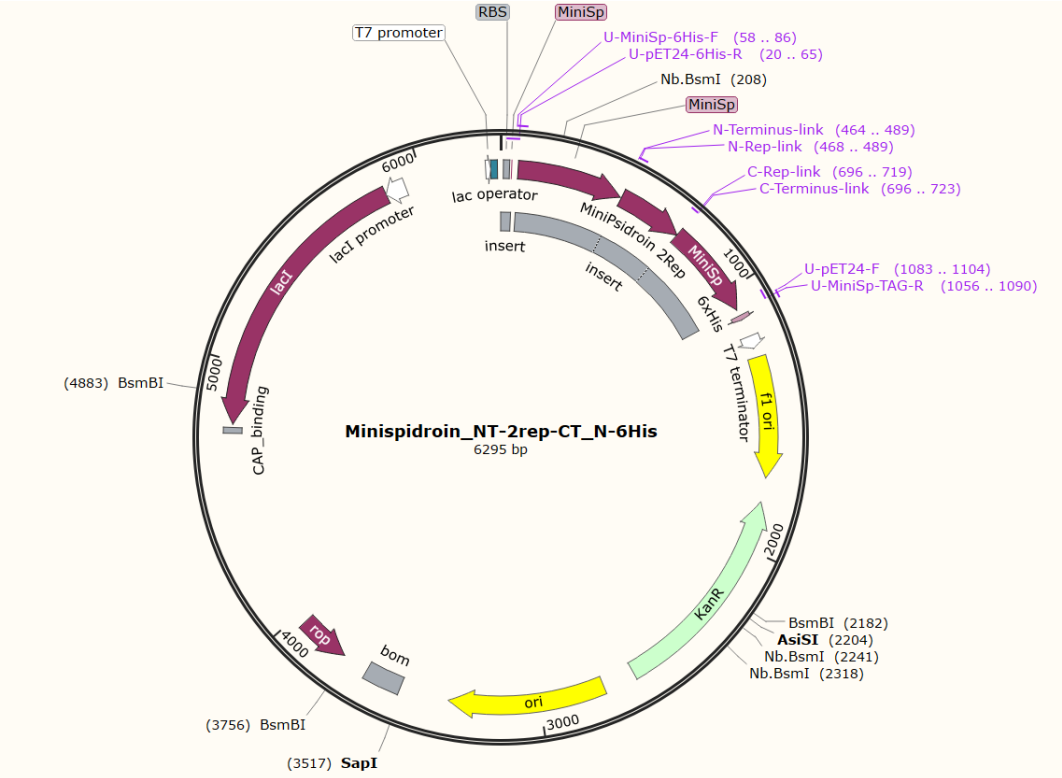


Figure 9: Minispidroin_NT-2rep-CT_N-6His plasmid map.

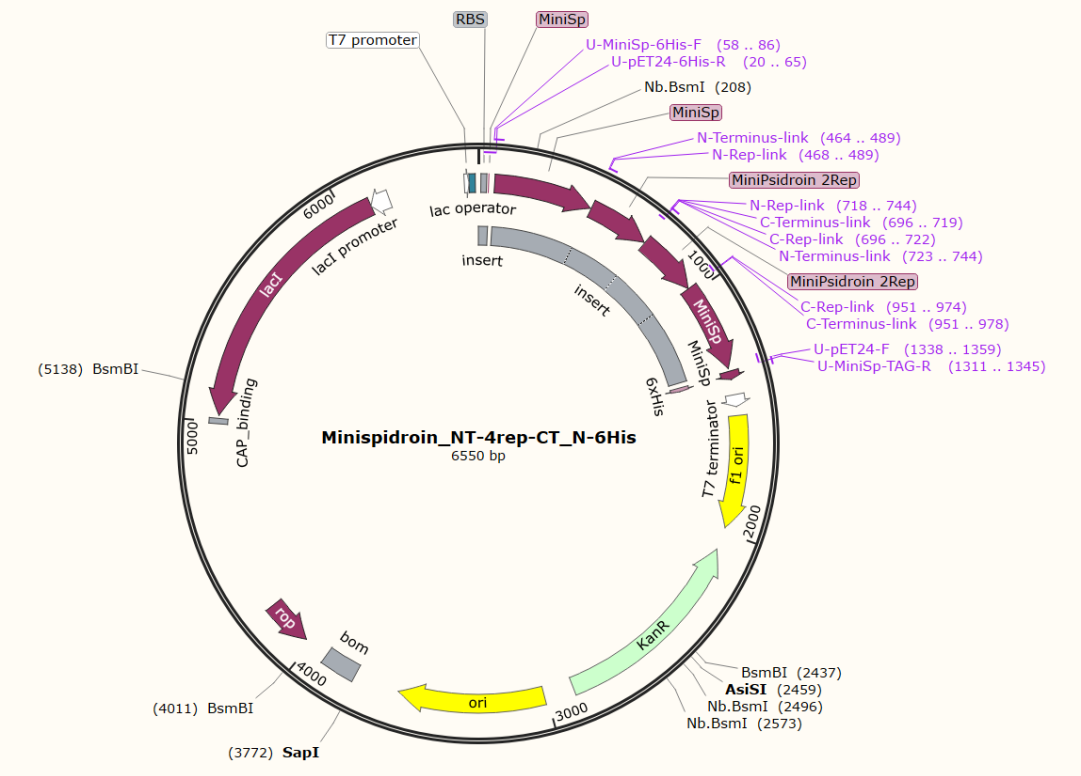


Figure 10: Minispidroin_NT-4rep-CT_N-6His plasmid map.

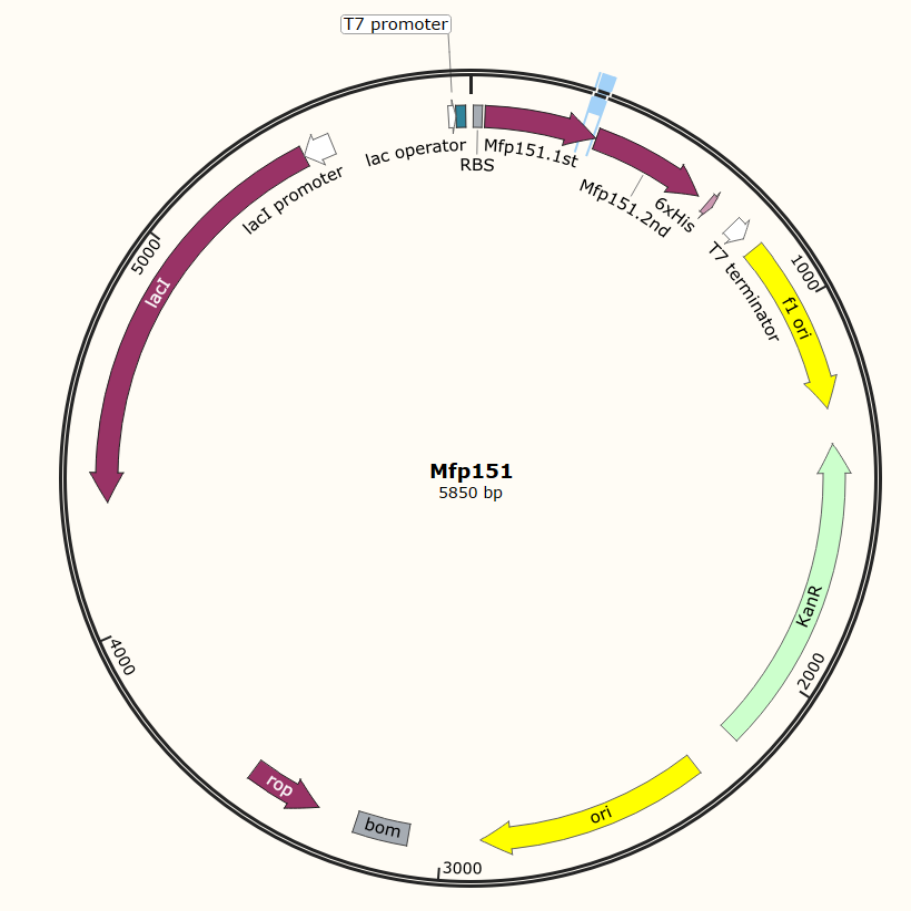


Figure 11: Mfp151 plasmid map.

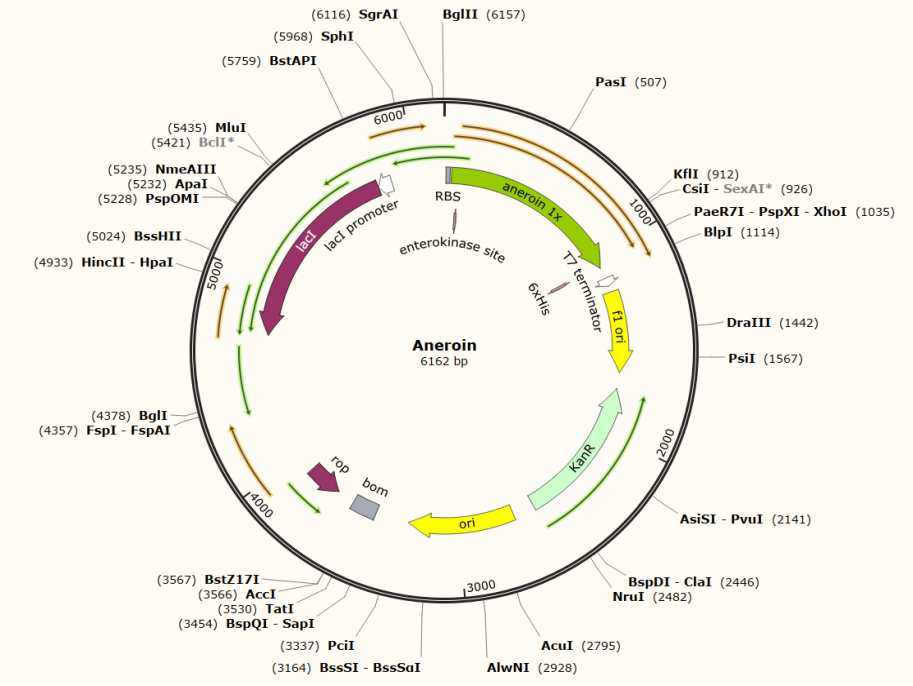


Figure 12: Aneroin plasmid map.

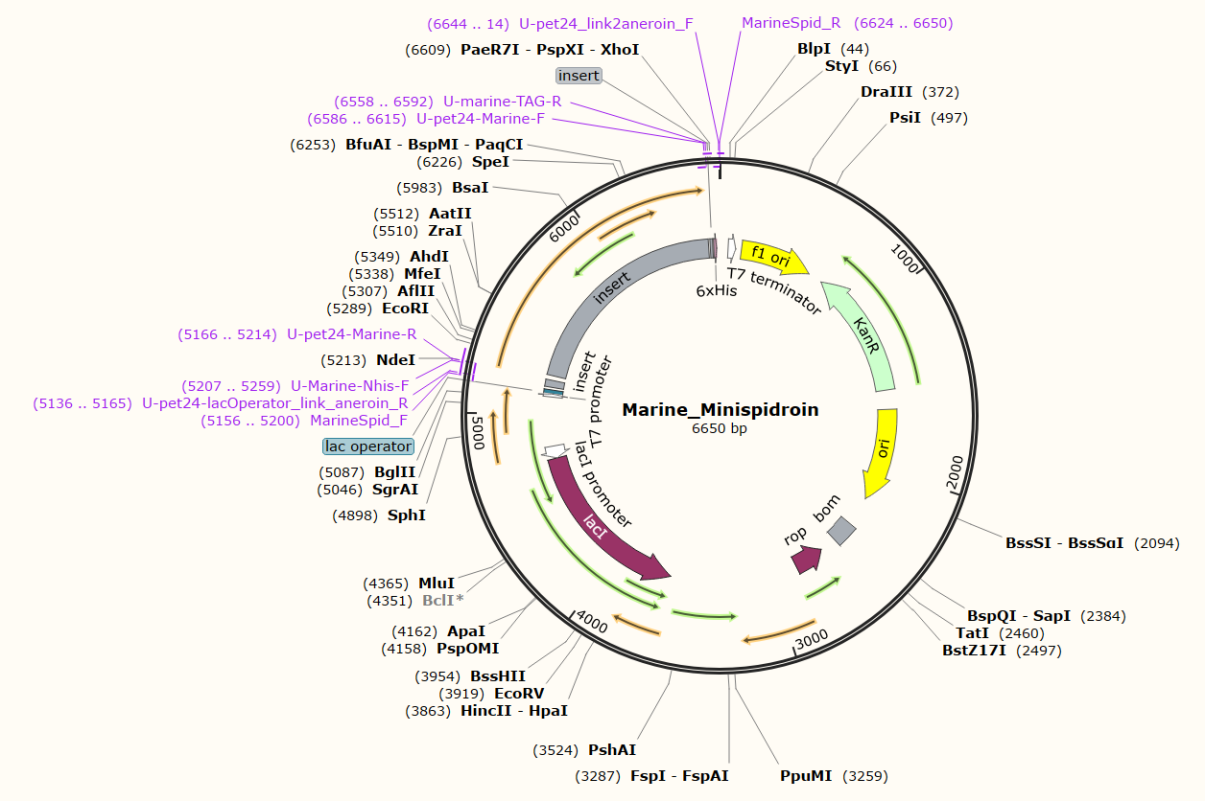


Figure 13: Marine Minispidroin plasmid map.

Protocols + Tips & Tricks

Protein Expression of Minispidroin_2rep and Minispidroin_4rep

- Take the appropriate amount of LB media and add kanamycin (concentration - 1mLantibiotic/1L media)
- Inoculate LB media containing the appropriate amount of kanamycin with the cells from a plate or from another starter culture
- Grow the cells in a shaking incubator at 37deg until they reach an OD600 of 0.8
- Induce the cells with 0.3mM IPTG
- Grow the induced cell culture at 28deg in a shaking incubator from 8h to overnight
- **Note:** It is best to optimize the inducer concentration and temperature for growing the cells post-induction since this varies for each protein expression system. So, different concentrations of the inducer and different incubation temperatures after induction can be tried and the one that has the highest protein expression can be selected

Expression of mfp

Take the appropriate amount of LB media and add kanamycin (concentration - 1mL/1L media)

- Inoculate LB media containing kanamycin (1:1000; 1uL for 1mL) with the cells from a plate or from another culture
- Grow the cells in a shaking incubator at 37deg until they reach OD600 of 0.8

- Induce the cells with 0.2mM IPTG
- Grow the induced cell culture at 18deg in a shaking incubator from 8h to overnight
- **Note:** It is best to optimize the inducer concentration and temperature for growing the cells post-induction since this varies for each protein expression system. So, different concentrations of the inducer and different incubation temperatures after induction can be tried and the one that has the highest protein expression can be selected

Protein Expression of Aneroin

- Take the appropriate amount of LB media and add kanamycin (concentration - 1mL/1L media)
- Inoculate LB media containing kanamycin (1:1000; 1uL for 1mL) with the cells from a plate or from another culture
- Grow the cells in a shaking incubator at 37deg until they reach OD600 of 0.8
- Induce the cells with 0.3mM IPTG
- Grow the induced cell culture at 37deg in a shaking incubator from 8h to overnight
- **Note:** It is best to optimize the inducer concentration and temperature for growing the cells post-induction since this varies for each protein expression system. So, different concentrations of the inducer and different incubation temperatures after induction can be tried and the one that has the highest protein expression can be selected

Agarose gel electrophoresis

- Fix the combs and the gel tray
- Pour 1% agarose until it reaches 3/4th of the comb’s depth
- Let it sit for at least 20mins
- **Note:** Be careful not to disturb the gel while it’s setting
- Once the gel is set, carefully remove the combs
- Remove the tray containing the gel and place it in the machine containing the buffer
- Mix 5ul of the sample with 1ul of the loading dye
- Load 6ul of the 1kB plus DNA ladder in the first well
- Then, load 6ul of the samples containing the dye in the other wells
- Run the gel at 100V for 30 mins or 135V for 20 mins
- Place the gel on the gel imaging system to visualize the gel

SDS-PAGE (whole cells)

Preparation of samples

- Take 1 mL of OD600 = 0.1 cells
- Centrifuge the cells at 10,000g for 10 mins
- Discard the supernatant
- Resuspend the cell pellet in 100uL of 1x SDS sample buffer
- **Note:** It is essential that the pellet is thoroughly resuspended in the SDS sample buffer
- Boil the samples at 95deg for 5-10 mins
- Centrifuge them at atleast 10,000g or more for 10 mins
- Take the required amount of sample from the top without disturbing the bottom

- **Note:** It is very important to take the sample only from the supernatant without disturbing the bottom since all the cell debris and insoluble proteins settle at the bottom. Loading these would lower the resolution of the SDS-gel

Running the gel

- Fill the chamber with SDS running buffer until the mark is reached
- Remove the tape at the bottom of the premade gel
- Insert the gel into the chamber
- Pour some more buffer into the gel until the wells are immersed
- Remove the comb gently
- **Note:** The comb bust is pulled upward gently to be removed. This must be done very carefully because if not, the wells get ruined and the gel needs to be replaced
- Flush the wells by pipetting up and down a few times in each well
- Load wells with the ladder samples
- Load each well with the required amount of protein sample
- Close the chamber while ensuring that the electrodes are connected right
- Set the voltage to 250V at constant voltage and run for 20-30 mins
- Stop the voltage and turn off the machine
- Carefully break the cassette to get the gel
- Wash the gel with distilled water
- Cut away the extra pieces of gel from the wells
- Place the gel on the gel imaging system to visualize the gel

Western blot

Blotting

- Place one set of filter papers on the cassette
- Place the membrane on top of the filter paper
- Place the gel on the membrane
- Place the other set of filter papers on the gel
- **Note:** To ensure an even transfer, remove the air bubbles between the layers with a Blot Roller. Don't put too much pressure on it, otherwise, the gel will be damaged
- Put the cassette in the "Trans-Blot turbo" machine and follow the instructions
- Place the cassette in the machine
- Use the right setting as per the membrane (mini or maxi membrane)
- After the blotting is done, take the blotted membrane
- Incubate the membrane in blocking solution for 1h at room temperature or overnight in the cold with shaking conditions
- Then, incubate the membrane with primary antibody solution for 1h or overnight in the cold with shaking conditions
- Wash the membrane thrice, 5 mins each with PBS-T
- Incubate the membrane with secondary antibody solution for 1h with shaking conditions
- Wash the membrane thrice for 10 mins each with PBS-T

- Add the colour developing solution to the blot
- **Note:** The blot should be incubated with the colour developing solution in the dark for no more than 5 mins since a longer incubation ruins the imaging of the blot
- Place the blot on the gel imaging system to visualize the blot

Recipe for solutions

- Blocking solution - 5% skim nonfat dry milk powder in PBS-T
- Primary antibody solution - 1:1000 dilution of the primary antibody (mouse anti-hexa his antibody) in the blocking solution
- Secondary antibody solution - 1:5,000 dilution of the secondary antibody (goat anti-mouse HRP conjugated secondary antibody) in the blocking solution
- **Note:** The antibody solutions can be reused for 3-5 times or until the milk in the solution goes bad

Making LB plates

- Melt the LB agar in the microwave until all of the agar is melted
- Check periodically that it's not boiling. If it is, take the bottle out with the rubber gloves and swirl it a bit to cool it slightly, then put it back in
- Take out the bottle and let the agar cool down to 60 deg
- It is important to make sure it's 60 deg since if it's too hot the antibiotic will be degraded and if it's not hot enough the agar solidifies
- Add appropriate antibiotic to the LB agar at a concentration of 1mL/ 1L LB agar
- Pour about ~20mL per plate till it covers/fills the plate's entire surface
- Let the plates rest until the agar is solid, then put them in the fridge with the resistance labeled on the lid
- It is important to store plates with Kanamycin in dark since kanamycin gets degraded with light and hence, the plates get contaminated easily

Transformation

- Take 50mL of competent cells in a microcentrifuge tube from the freezer and let the tube thaw on ice for 20-30 mins
- Let agar plates (containing the appropriate antibiotic) warm up to room temperature
- Add 1 - 5 µl of DNA (usually 10 pg - 100 ng) to the competent E.coli cells in a microcentrifuge or falcon tube
- Mix gently by flicking the bottom of the tube a few times
- Incubate the mixture on ice for 20-30 mins
- Heat shock the tube by placing the bottom 1/2 to 2/3 of the tube into a 42°C water bath for 45 sec
- Put the tubes back on ice for 2 min
- Add 700uL of LB media (without antibiotic) to the bacteria
- Grow in 37°C shaking incubator for 45 min
- Plate 100uL of the transformation on an LB plate with the appropriate antibiotic
- Incubate plates at 37°C overnight

Cell lysis (French press homogenizer)

- Centrifuge the cell culture and discard the supernatant to obtain the cell pellets
- Add lysis buffer and resuspend the cell pellet in the lysis buffer
- **Note:** Amount of lysis buffer to be added should be 10-20% of the original volume of cell culture that was centrifuged
- Filter the solution before lysis
- Set the pressure to 10kPsi in the French press
- Perform 2 washes: 30ml water followed by 30ml lysis buffer
- Then set the required pressure
- Add the cell suspension and lyse the cells
- Add 30ml lysis buffer to the French press and collect this too
- Wash the centrifuge thrice with 100ml water and once with 20% ethanol
- Turn off the machine

Obtaining the soluble lysate

- Centrifuge the lysate at maximum speed
- Collect the supernatant as the soluble fraction of the lysate
- **Note:** Do not disturb the pellet since it contains all the cell debris and insoluble proteins

Protein purification using Ni-NTA agarose

- Take the cell lysate in a tube and add enough Ni-NTA agarose according to its binding capacity. For instance, if the binding capacity of Ni-NTA agarose is 50mg/mL and the lysate is expected to have 100 mg of protein then add 2mL of resin
- Keep the tubes shaking in a cold room for at least 45 mins to allow the Ni-NTA resin to bind the proteins
- Centrifuge the tubes at maximum speed for 5 mins at 4 deg to let the resin settle at the bottom
- Discard the flowthrough without disturbing the resin at the bottom while saving some flowthrough (~2mL) in another tube labeled “Flowthrough”
- Add 10 bed volumes (amount of Ni-NTA resin) of the wash buffer
- Keep the tubes shaking in a cold room for 4 mins
- Centrifuge the tubes at maximum speed for 5 mins at 4 deg to let the resin settle at the bottom
- Discard the wash without disturbing the resin at the bottom while saving some wash (~2mL) in another tube labeled “Wash 1”
- Repeat the wash step one more time and save some wash (~2mL) in another tube labeled “Wash 2”
- Finally, add 2 bed volumes (amount of Ni-NTA resin) of elution buffer
- Keep the tubes shaking in a cold room for 10 mins
- Centrifuge the tubes at maximum speed for 5 mins at 4 deg to let the resin settle at the bottom
- Collect all the eluate carefully without disturbing the resin to obtain the pure proteins

- Run an SDS-PAGE gel on the different fractions - Flowthrough, wash 1, wash 2 and eluate - to evaluate whether all the protein was obtained in the eluate or if some of it was lost in any of the flowthrough or wash fractions
- **Note:** To clean the resin, wash the resin with 15 bed volumes of 0.5M NaOH. The resin and NaOH should be kept shaking in the cold for at least 30 `mins.
- **Note:** Add some 30% ethanol to the Ni-NTA left in the tube and store in the fridge. This protects the resin from any microbial growth and hence, allows the reuse of the resin for the same protein.
- **Note:** If the resin changes colour from blue to brown, it shows that the resin is contaminated. In that case, there is an elaborate protocol to regenerate the resin. The protocol can be found in the QIAexpressionist under “Reuse of Ni-NTA MAtrices” in page 109. Link for the book: https://www.qiagen.com/us/resources/resourcedetail?id=79ca2f7d-42fe-4d62-8676-4cfa948c9435&lang=en

Dialysis

- Use a 10kDa molecular weight cutoff dialysis membrane
- Make a secure knot at the bottom of the tubing
- Add the purified protein sample obtained after elution to the tubing
- Make another secure knot at the top while leave at least a 1cm headspace between the protein sample and the top knot
- Keep the dialysis membranes containing the proteins in a bucket containing dialysis buffer
- **Note:** The volume of the dialysis buffer in the bucket should be at least 10 times the volume of the protein sample in the dialysis membrane

Recipe for buffers:

1. 2rep/4rep:
    - Lysis buffer - 20mM Tris-Cl, pH 8.0
    - Wash buffer - 20mM Tris-Cl, 5mM Imidazole, pH 8.0
    - Elution buffer - 20mM Tris-Cl, 200mM Imidazole, pH 8.0
    - Dialysis buffer - 20mM Tris-Cl, pH 8.0
2. Mussel foot protein:
    - Lysis buffer - 10mM Tris-Cl, 100mM NaH_2PO_4, 8M Urea, pH 8
    - Wash buffer - 10mM Tris-Cl, 100mM NaH_2PO_4, 8M Urea, pH 6
    - Elution buffer 1 - 10mM Tris-Cl, 100mM NaH_2PO_4, 8M Urea, pH 4.5
    - Elution buffer 2 - 0.5M HCl
    - Dialysis buffer - distilled water
3. Aneroin:
    - Lysis buffer - 50mM NaH_2PO_4, 300 mM NaCl, 10mM imidazole, pH 8.0
    - Wash buffer - 50mM NaH_2PO_4, 300 mM NaCl, 30 mM imidazole, pH 8.0
    - Elution buffer - 50mM NaH_2PO_4, 300 mM NaCl, 300 mM imidazole, pH 8.0
    - Dialysis buffer - Distilled water

Concentration of proteins

- After purification and dialysis you should have a sample of the pure protein whose concentration is determined through BCA assay
- Take a centrifugal filter unit with a filter size fitting the size of your protein
- First, wash the unit with water for 15 mins
- Then, wash with the dialysis buffer for 15 mins
- Then, transfer your protein sample to the centrifugal filter unit
- Centrifuge at 4000g, 4deg for at least 1h
- If the initial volume of your sample exceeds the maximum volume for the centrifugal filter unit, add the remaining protein sample to the filter unit and centrifuge again
- **Note:** Remember to remove the liquid that flows through the filter
- Repeat the steps until all the protein sample is concentrated
- When the volume has been reduced to match the required concentration of protein, remove the sample in the filter with a pipette
- **Note:** The membrane of the filter unit can be fragile so gently insert the pipette tip to remove the concentrate protein without damaging the filter unit
- **Note:** Wash the sides of the filter with the sample still in the filter to ensure that all your proteins are in the bottom of the filter
- **Note:** This can also be used as a coarse way to remove salts and small molecules that aren't bound to your protein, if you wash the sample a couple of times with MilliQ in the filter. This is because the small molecules can pass through the filter, whereas the proteins cannot

IPTG/Temp optimization

- We used an experiment approach for optimizing our yields. See Link_To_model

Fibre spinning

- **Preparing the physical set-up**
    - Set-up a syringe pump with a speed of 17 uL/min
    - Use a bunsen burner to stretch a capillary tube, and break it in two. Test that milliQ water can go through the tip. This acts as the syringe tip that goes in the spinning bath.
    - Use a syringe tip to connect the syringe to a silicon tupe, which is then connected to the capillary syringe tip.
    - **Note:** Don’t insert the capillary syringe tip in the spinning bath before starting the pump, as this will suck some of the spinning bath into the

- Making of spinning dope (solute of protein for spinning)
    - Approach 1
        - Freeze dry the protein which will give you a powder (see Freeze Drying protocol)
        - Redissolve the freeze dried protein in 20 mM Tris in amount that would resolve in a 30% w/w

    - Approach 2
        - Use a filtered falcon tube to increase the concentration of the protein until you have the desired concentration (30 % w/w)
- Making the bath to spin into
    - Make a bath with 500 mM sodium acetate buffer and 200 mM NaCl at pH 5
    - Note: We also experimented with adding trivalent anions to the bath. In our case: Phosphate (X M)
- Set-up for stretching the fibre
- Insert your concentrated protein solution in a syringe attached to an automated syringe pump
- Start the syringe pump, and lower the capillary syringe pump into the spinning bath.

BOKU-Vienna - Pichitecture

Engineering

Design

Major strain spider silk consists of four major parts, the first T-sequence, two S-sequences and another T-sequence at the end [fig. 2]. Due to a high number of repeats in the sequence, synthesizing the whole sequence is not possible (alternative: not done by the company IDT) . The engineering design was therefore split into two steps, allowing us to order the insert in smaller parts, which needed to be combined to gain the whole sequence for the major strain spider silk.

Step 1:

In a first step, only the T-sequence in the beginning and the first of the two S-sequences were created using polymerase chain reaction (PCR)[fig. 3]. The S-sequence could be ordered at the company IDT, when split into three different parts (S#1, S#2 and S#3). The T-sequence on the other hand could not be ordered at the company IDT. Thus, various primers, resulting in the T-sequence when combined, were ordered at the company IDT.
In a first trial run a PCR was done containing all the components for part 1 of the engineering design. However, no band in the expected size could be seen on the electrophoresis performed after the PCR.
After discussing these results and brainstorming solutions, sequence-T, Part S#1, Part S#2 and Part S#3 were amplified by PCR separately. The reaction resulting in the T-sequence consisted of the primers “FW1 A1T”, “BW1 A1T”, “FW2 A1T” and “BW2 A1T”. For the parts S#1, S#2 and S#3 the template sequences were mixed with the according primers resulting in the three following reactions. Template “A1S#1” and primers “FW2 A1T” and “A1Spart1and2BW” for the part S#1, template “A1S#2” and primers “A1Spart1and2FW” and “A1Spart2and3BW” for the part S#2 as well as template “A1S#3” and primers “A1Spart2and3FW” and “A1Spart3BW”. To combine these products in a second reaction, the primers “FW1 A1T”, “BW2 A1T”, “A1Spart1and2FW”, “A1Spart1and2BW”, “A1Spart2and3FW”, “A1Spart2and3BW”, and “A1Spart3BW” were used. The success of this experiment was proven by sequencing the cleaned PCR product at the company “Microsynth”.

Step 2:

The PCR product of step one was now amplified with three different sets of primers to gain the parts A, B and C together creating the insert used in the following Golden Gate cloning to create the designed plasmids. These primers also contained recognition sites for the restriction enzyme BsaI, resulting in sticky ends with the ability to combine the PCR products to the correct sequence, as well as the possibility to add the secretion factor and clone the whole insert into a backbone 1 with the fusion sites 2 and 3.
Part A was constructed with the use of the primers “A1T FW1” on the on hand and

“PCR-2-BW” on the other hand [fig. 4], resulting in a N-terminal sticky end with the overhang sequence “GAGA”, fitting the C-terminal sticky end of the secretion factor for Pichia pastoris, and a C-terminal sticky end with the overhang sequence “CGCT”, fitting the N-terminal sticky end of part B. This part consists of the N-terminal T-sequence as well as the first of the two S-sequences.
For part B the primers “PCR-3-FW” and “PCR-3-BW” were used for the creation of a PCR-product [fig. 5] containing only one S-sequence including the N-terminal sticky end with the overhang sequence “CGCT”, fitting the C-terminal sticky end of part A and the C-terminal sticky end with the overhang sequence “GGCG”, fitting the N-terminal sticky end of part C. Part B is used as the second of the two S-sequences in the complete insert.

The primers “PCR-4-FW” and “PCR-4-BW” were necessary to produce the small PCR product C, only including the C-terminal T-sequence [fig.6]. This PCR product contains the C-terminal fusion site 3, needed for the cloning in backbone 1, as well as the N-terminal sticky end with the overhang sequence “GGCG”, fitting the C-terminal sticky end of part B.
The resulting parts A, B and C were determined to contain the expected sequences by sending them to the company “Mikrosynth” for sequencing, after cleaning up the PCR-products.

Golden Gate cloning:

Backbone 1
The cloning in backbone 1 combines the three parts of the main strain spider silk sequence with the secretion signal for Pichia pastoris. This secretion signal contains the N-terminal fusion site 2, used for cloning into backbone 1 and the C-terminal sticky end with the overhang sequence “GAGA”, fitting the N-terminal sticky end of part A, when cut with the restriction enzyme BsaI. To later gain inside on the success of the Golden Gate reaction, the backbone 1 contains a kanamycine resistance gene. After the Golden Gate reaction, the resulting plasmid was transformed into chemically competent E.coli of the strain DH10b by heatshock treatment. For the determination of a successful reaction a negative control was added to the Golden Gate reaction. After plating the transformed E.coli on LB-media plates containing the antibiotic kanamycine and incubating over night at 37 degrees Celsius, cell growth could only be witnessed on the plates containing the E.coli with the correct plasmid. However, after extracting the plasmid, using the “MiniPrep” method, choosing the samples to send for sequencing by doing a restriction digest to get slight knowledge about the composition of the plasmid and sending the chosen samples for sequencing, slight mutations in the coding sequence could be noticed in all the samples.
Since only a few mutations occurred and the sequencing results of the PCR products clearly showed the expected sequences, the decision to just repeat the Golden Gate experiment was made. The second run proved to be very successful as every sample

sent to sequencing matched the expected sequence.

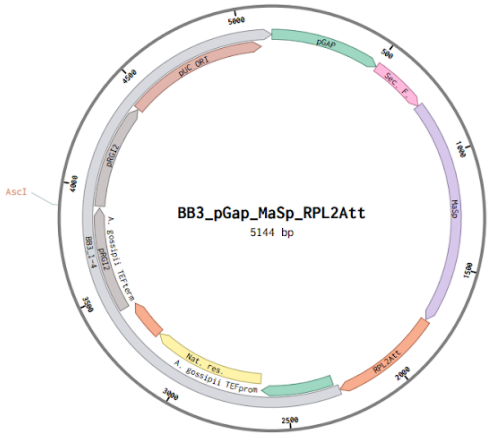
Backbone 2 [fig. 8-11]
Backbone 2 was only needed for the plasmid containing the four coding sequences, since it can be skipped by using by using an empty backbone 3 with the fusion sites 1 and 4, created by digest with BbsI.
For the plasmid version including the coding sequence for main strain spider silk four times, four different backbone 2 plasmids were made. To avoid problems caused by inadequate binding of DNA in repeating regions, two promoters and three terminators were used in the four backbone 2 plasmids.

Every backbone 2, used for the main strain spider silk contains the fusion sites 1 and 4, created by digest with BbsI, as well as an ampicillin resistance used for the determination of a successful cloning. The used promoters contain the N-terminal fusion site 1 and the C-terminal fusion site 2, fitting the N-terminal fusion site 2 of the secretion factor whereas the used terminators contain the N-terminal fusion site 3, fitting the C-terminal fusion site 3 of the main strain spider silk and the C-terminal fusion site 4, used for cloning in backbone 2, when digested with the restriction enzyme BbsI:
The difference between the used backbones lies in the fusion sites, created by a digest with BsaI. This digest is used for the cloning in backbone 3.
To assemble the backbone 2-versions, a Golden Gate reaction with the restriction enzyme BbsI, was performed. The reactions entailed the different promoters and terminators as well as the previously created backbone 1, containing the major strain spider silk insert including the secretion signal for Pichia pastoris.
The first created backbone 2 consists of the promoter “pGAP”, the terminator “RPL2Att”, the MaSp insert including the secretion signal and a backbone 2 version building the fusion sites A and B after a digest with the restriction enzyme BsaI. The N-terminal fusion site A is used for the cloning in backbone 3 while the C-terminal fusion site B fits the N-terminal fusion site B of the second created backbone 2.
For the second backbone 2-version the promoter “pTEF”, the terminator “RPP1Btt” and the MaSp insert including the secretion signal were cloned in an empty backbone 2, building the N-terminal fusion site B, fitting the C-terminal fusion site B of the first backbone 2 and the C-terminal fusion site C, fitting the N-terminal fusion site C of the third backbone 2 version, when digested with the restriction enzyme BsaI.
The third created backbone 2 consists of the promoter “pGAP”, the terminator “RPS25Att”, the MaSp insert including the secretion signal and a backbone 2 version building the fusion sites C and E after a digest with the restriction enzyme BsaI. The N-terminal fusion site C fits the C-terminal fusion site C of the second backbone 2 while the C-terminal fusion site E fits the N-terminal fusion site B of the fourth created backbone 2.
For the fourth backbone 2-version the promoter “pTEF”, the terminator “RPP1Btt” and the MaSp insert including the secretion signal were cloned in an empty backbone 2, building the N-terminal fusion site C, fitting the C-terminal fusion site C of the third

backbone 2 and the C-terminal fusion site E, used for cloning in a backbone 3 with the fusion sites A and E, when digested with the restriction enzyme BsaI.
The Golden Gate products of the four backbone 2 versions were transformed into chemically competent E.coli DH10b and plated on LB-Ampicillin Agar plates. After an incubation period of 18 hours at 37 degrees Celsius, colony growth could only be witnessed on the plates containing the complete backbone 2 versions, but not on the control plates only containing the empty and, after the Golden Gate, linearized backbone 2 versions. As further proof for the successful cloning a control digest was performed using the restriction enzyme BsaI. After loading the digest on a 1% agarose gel, the bands for all backbone versions could be found at the expected sizes.

Backbone 3:
To obtain a comparison regarding expression efficiency, three different backbone 3 were created. The first and second version each contain one expression cassette that varys only in terms of the used promoter. The third expression cassette entails four different expression cassettes.
The backbone 3 were obtained by Golden Gate cloning. All backbone three versions contain a nourseothricin resistance to enable a selection of the E.coli clones including the plasmid after the Golden Gate reaction was transformed into E.coli.

BB3-pGAP-MaSp-RPL2Att
The first created backbone 3 is the plasmid BB3-pGAP-MaSp-RPL2Att. It contains the promoter “pGAP”, the terminator “RPL2Att” and the MaSp insert including the secretion signal. This plasmid was synthesized by doing a Golden Gate reaction including the Backbone 1 with the MaSp coding sequence and the secretion signal for Pichia Pastoris, the “pGAP” promoter and the “RPL2Att terminator. The restriction enzyme BbsI was used to create the desired sticky ends. For the empty backbone the backbone 3 version, building the fusion sites 1 and 4, when digested with BbsI, was chosen.
After an incubation period of 18 hours at 37 degrees Celsius, colony growth could only be witnessed on the nourseothricin-plates containing the complete BB3-pGAP-MaSp-RPL2Att plasmid but not on the control plates only containing the empty and, after the Golden Gate, linearized backbone 3. As further proof for the successful cloning a control digest was performed using the restriction enzyme BglI. After loading the digest on a 1% agarose gel, the bands could be found at the expected sizes.



1. Insert:

- Major Spider Silk (MaSp)
- Secretion Factor (Sec. F.) for Pichia Pastoris

1.1. Amino Acid Sequence:

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPFSNSTN
NGLLFINTTIASIAAKEEGVSLEKRGAPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
PPGPPGPPGPPGPAGASGPGQQGRGGYGPGQQGPSGPGSGAAAAAAGGAGQGGYGGLGGQ
GAGRGGLGGQAAAAAASGPGGYGPGQQGPGGYGPGQQGPSGPGSMAAAAAAAMGPGGYGP
GQQGPGGYGSPAGASGPGQQGRGGYGPGQQGPSGPGSGAAAAAAGGAGQGGYGGLGGQGA
GRGGLGGQAAAAAASGPGGYGPGQQGPGGYGPGQQGPSGPGSMAAAAAAAMGPGGYGPGQ
QGPGGYGSPAGAPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPAG
G**

1.2. Nucleic Acid Sequence:

ATGAGATTCCCATCTATTTTCACCGCTGTCTTGTTCGCTGCCTCCTCTGCATTGGCTGCC
CCTGTTAACACTACCACTGAAGATGAGACTGCTCAAATTCCAGCTGAAGCAGTTATCGGT
TACTCTGACCTTGAGGGTGATTTCGACGTCGCTGTTTTGCCTTTCTCTAACTCCACTAAC
AACGGTTTGTTGTTCATTAACACCACTATCGCTTCCATTGCTGCTAAGGAAGAGGGTGTC
TCTCTCGAGAAGAGAGGCGCTCCCCCAGGCCCACCCGGACCCCCCGGACCACCTGGCCCC
CCTGGTCCACCTGGCCCACCTGGCCCACCTGGACCACCAGGACCACCAGGTCCCCCCGGC
CCTCCCGGTCCTCCCGGACCCCCTGGACCCCCTGGACCCGCCGGCGCAAGTGGTCCAGGA
CAGCAGGGTAGGGGAGGTTACGGTCCCGGTCAACAAGGTCCATCTGGTCCTGGTTCCGGT
GCTGCAGCAGCTGCCGCTGGAGGTGCTGGTCAGGGCGGATATGGTGGTCTTGGCGGTCAA
GGTGCTGGAAGAGGTGGATTAGGAGGACAGGCCGCAGCCGCTGCTGCCTCTGGTCCCGGA
GGATATGGACCTGGACAGCAAGGTCCAGGCGGATATGGTCCCGGTCAGCAAGGCCCCTCA
GGACCCGGTTCTATGGCTGCAGCCGCTGCTGCAGCTATGGGACCCGGTGGTTACGGCCCT
GGACAGCAAGGACCTGGCGGATACGGTTCTCCCGCTGGCGCAAGTGGTCCAGGACAGCAG
GGTAGGGGAGGTTACGGTCCCGGTCAACAAGGTCCATCTGGTCCTGGTTCCGGTGCTGCA
GCAGCTGCCGCTGGAGGTGCTGGTCAGGGCGGATATGGTGGTCTTGGCGGTCAAGGTGCT
GGAAGAGGTGGATTAGGAGGACAGGCCGCAGCCGCTGCTGCCTCTGGTCCCGGAGGATAT
GGACCTGGACAGCAAGGTCCAGGCGGATATGGTCCCGGTCAGCAAGGCCCCTCAGGACCC
GGTTCTATGGCTGCAGCCGCTGCTGCAGCTATGGGACCCGGTGGTTACGGCCCTGGACAG

CAAGGACCTGGCGGATACGGTTCTCCCGCTGGCGCTCCCCCAGGCCCACCCGGACCCCCC
GGACCACCTGGCCCCCCTGGTCCACCTGGCCCACCTGGCCCACCTGGACCACCAGGACCA
CCAGGTCCCCCCGGCCCTCCCGGTCCTCCCGGACCCCCTGGACCCCCTGGACCCGCCGGT
GGATGATAA

2. Promoter:

- pGap-Promoter

2.1. Nucleic Acid Sequence:

AGGATCCTTTTTTGTAGAAATGTCTTGGTGTCCTCGTCCAATCAGGTAGCCATCTCTGAA
ATATCTGGCTCCGTTGCAACTCCGAACGACCTGCTGGCAACGTAAAATTCTCCGGGGTAA
AACTTAAATGTGGAGTAATGGAACCAGAAACGTCTCTTCCCTTCTCTCTCCTTCCACCGC
CCGTTACCGTCCCTAGGAAATTTTACTCTGCTGGAGAGCTTCTTCTACGGCCCCCTTGCA
GCAATGCTCTTCCCAGCATTACGTTGCGGGTAAAACGGAGGTCGTGTACCCGACCTAGCA
GCCCAGGGATGGAAAAGTCCCGGCCGTCGCTGGCAATAATAGCGGGCGGACGCATGTCAT
GAGATTATTGGAAACCACCAGAATCGAATATAAAAGGCGAACACCTTTCCCAATTTTGGT
TTCTCCTGACCCAAAGACTTTAAATTTAATTTATTTGTCCCTATTTCAATCAATTGAACA
ACTATCAAAACAC

3. Terminator:

- RPL2Att-Terminator

3.1. Nucleic Acid Sequence:

CTATGTAACTAACGAAACAGCATGTACTAATAGAACCGTATCGAGAATATTTATTTAGGT
GAGTAGTAGGAGTGAACCAGACAGTCAATTTAGTGAGCTGTCCCAGCTTTTGTGCATTCC
AGAATTGCCGGTCAAATTGGTTATGGGTTATGGGGCTTTTCCGATTGAGGTTCAGTTTCT
GCGGTTATCTCTTTCTTGACCTGGTCTTTTACAGGCTGTTCTTTCTCCCCATGATTATTC
TTTAGCTGAAGATACCGCTTAGCCTGATAATGTCGTCGTTTTGTAATCAAAATCTTTAGT
TGGGCATCGTCTGAGGTTTCCTTTGGCTTCTGGGGTTGTTAGTAGGAACGTAGGAACCAT
AGTAACTTTTACACATACATTCTTATGATTGCGAAGTAAGCTGAGTCTGCTGCTTGGCTC
CCGAAGTACTTTCTCTTTCTCTACCGGTTGATTCTCCTTCTGGTGCTCCTAAACGATTGT
GTTAGAAGGGATTGAC

4. Backbone:

- Backbone 3 (BB3)
- Fusion-Site 1-4
- Nourseothrycin resistence (Nat. res.)

4.1. Nucleic Acid Sequence:

CGCTCCGGCATGCCGGCATGGACATGGAGGCCCAGAATACCCTCCTTGACAGTCTTGACG
TGCGCAGCTCAGGGGCATGATGTGACTGTCGCCCGTACATTTAGCCCATACATCCCCATG
TATAATCATTTGCATCCATACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTACGGCT
CCTCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCTCACAGACGCGTTGAATTGTCCCC
ACGCCGCGCCCCTGTAGAGAAATATAAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTTC
ATATACTTCCTTTTAAAATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAAC
AACCATGGGTACCACTCTTGACGACACGGCTTACCGGTACCGCACCAGTGTCCCGGGGGA
CGCCGAGGCCATCGAGGCACTGGATGGGTCCTTCACCACCGACACCGTTTTTCGCGTCAC
CGCCACCGGGGACGGCTTCACCCTGCGGGAGGTGCCGGTGGACCCGCCCCTGACCAAGGT
GTTCCCCGACGACGAATCGGACGACGAATCGGACGACGGGGAGGACGGCGACCCGGACTC

CCGGACGTTCGTCGCGTACGGGGACGACGGCGACCTGGCGGGCTTCGTGGTCGTCTCGTA
CTCCGGCTGGAACCGCCGGCTGACCGTCGAGGACATCGAGGTCGCCCCGGAGCACCGGGG
GCACGGGGTCGGGCGCGCGTTGATGGGGCTCGCGACGGAGTTCGCCCGCGAGCGGGGCGC
CGGGCACCTCTGGCTGGAGGTCACCAACGTCAACGCACCGGCGATCCACGCGTACCGGCG
GATGGGGTTCACCCTCTGCGGCCTGGACACCGCCCTGTACGACGGCACCGCCTCGGACGG
CGAGCAGGCGCTCTACATGAGCATGCCCTGCCCCTAATCAGTACTGACAATAAAAAGATT
CTTGTTTTCAAGAACTTGTCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAA
TCAAATGTTAGCGTGATTTATATTTTTTTTCGCCTCGACATCATCTGCCCAGATGCGAAG
TTAAGTGCGCAGAAAGTAATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTG
GCTTCATACCCAGCATTGACCTTTGGTATGAGCATCTGAAAAACAACCAGGTGTTGCAAA
GTTAAACATCCTTCTTTGTTCATATAGAACCCACTATTCATGGTACTCCCCAATCGAATT
TCACATTCTGGTTTTGAAATTACACACCACGTTAGCTTATAAGATTTCATATAACTTATT
GATATACGGTTTCCATTGTTCGAATAGTTGAGGTTGTATGTAATTCGATTGAAGGGGCCA
TTTTTGTTTCCTACTTTTCCTGGGAGCTTATCCGATGCGCTTCAAAGCTGGAATTGTAAA
TATAGAGAAAAAGAAGGATGTTGTTTTATTCTTGAAAGAGTATAATTTTACTTCTAGCAA
CTCTCCCACTTCGCTTGACTTCATTTATTTCTTGGGCACATAGGCGTAGTAATCTAGACC
AACAGATAATTTGCCGGAATGATATAGCGATTGGAAAATGAACTGAAATTTTTTGCTGTC
TTTCAATTTGACGGGCAGTTCATCAGGCGCGCCGTGACCGACCATATAAATACGTTGAGA
ATGTTATTCTTCCTCGTAGTTGAAGTGGCTTCATAATTTCAGAACTCAATAGATAAACTA
GGATGTTTTAAAGCAATTAATGCTCACAAGTAAGGAGCGACTCTCTTGCTTTTCGAATAC
TAAAAGTATCGTCCCAACCCAGAAAAAAAGACCTCTTAACTGCAAAATAAACTCTATATA
TTTCTTCTAAAACAGTTTCAGGTTGGATAGTATCGCATTCTCATCACTTCTAACTAGTAG
GCCATGAGATATATTAACGTTTACTTGAGTTCTAAGTTCTCCGAATTAGATGCACAGCAC
AAACAAGATTAGGTTTCACTTGGTACAAAATACGAACAGAGTTTAAGGTCGTAATTTCAT
TTCGTTATTGATCCCCACAATCTATTCTTATCACAGTCATCAGATAGTCGCGAAAAAGCA
TGCAGAAAAGGGGGTCGTCCCTATCTAAGTTGTAGCATTACAACAAATATGACCGCTGTG
AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA
TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA
CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC
TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC
GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT
GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG
TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG
GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA
CGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG
AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT
TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT
TTCTACGGGGTGGAGCGCAAAAAACCCCGCCCCTGACAGGGCGGGGTTTTTTCGCGATCG
G

5. Complete Plasmid:

5.1. Amino Acid Sequence:

AGGATCCTTTTTTGTAGAAATGTCTTGGTGTCCTCGTCCAATCAGGTAGCCATCTCTGAA
ATATCTGGCTCCGTTGCAACTCCGAACGACCTGCTGGCAACGTAAAATTCTCCGGGGTAA
AACTTAAATGTGGAGTAATGGAACCAGAAACGTCTCTTCCCTTCTCTCTCCTTCCACCGC
CCGTTACCGTCCCTAGGAAATTTTACTCTGCTGGAGAGCTTCTTCTACGGCCCCCTTGCA
GCAATGCTCTTCCCAGCATTACGTTGCGGGTAAAACGGAGGTCGTGTACCCGACCTAGCA

GCCCAGGGATGGAAAAGTCCCGGCCGTCGCTGGCAATAATAGCGGGCGGACGCATGTCAT
GAGATTATTGGAAACCACCAGAATCGAATATAAAAGGCGAACACCTTTCCCAATTTTGGT
TTCTCCTGACCCAAAGACTTTAAATTTAATTTATTTGTCCCTATTTCAATCAATTGAACA
ACTATCAAAACACCATGAGATTCCCATCTATTTTCACCGCTGTCTTGTTCGCTGCCTCCT
CTGCATTGGCTGCCCCTGTTAACACTACCACTGAAGATGAGACTGCTCAAATTCCAGCTG
AAGCAGTTATCGGTTACTCTGACCTTGAGGGTGATTTCGACGTCGCTGTTTTGCCTTTCT
CTAACTCCACTAACAACGGTTTGTTGTTCATTAACACCACTATCGCTTCCATTGCTGCTA
AGGAAGAGGGTGTCTCTCTCGAGAAGAGAGGCGCTCCCCCAGGCCCACCCGGACCCCCCG
GACCACCTGGCCCCCCTGGTCCACCTGGCCCACCTGGCCCACCTGGACCACCAGGACCAC
CAGGTCCCCCCGGCCCTCCCGGTCCTCCCGGACCCCCTGGACCCCCTGGACCCGCCGGCG
CAAGTGGTCCAGGACAGCAGGGTAGGGGAGGTTACGGTCCCGGTCAACAAGGTCCATCTG
GTCCTGGTTCCGGTGCTGCAGCAGCTGCCGCTGGAGGTGCTGGTCAGGGCGGATATGGTG
GTCTTGGCGGTCAAGGTGCTGGAAGAGGTGGATTAGGAGGACAGGCCGCAGCCGCTGCTG
CCTCTGGTCCCGGAGGATATGGACCTGGACAGCAAGGTCCAGGCGGATATGGTCCCGGTC
AGCAAGGCCCCTCAGGACCCGGTTCTATGGCTGCAGCCGCTGCTGCAGCTATGGGACCCG
GTGGTTACGGCCCTGGACAGCAAGGACCTGGCGGATACGGTTCTCCCGCTGGCGCAAGTG
GTCCAGGACAGCAGGGTAGGGGAGGTTACGGTCCCGGTCAACAAGGTCCATCTGGTCCTG
GTTCCGGTGCTGCAGCAGCTGCCGCTGGAGGTGCTGGTCAGGGCGGATATGGTGGTCTTG
GCGGTCAAGGTGCTGGAAGAGGTGGATTAGGAGGACAGGCCGCAGCCGCTGCTGCCTCTG
GTCCCGGAGGATATGGACCTGGACAGCAAGGTCCAGGCGGATATGGTCCCGGTCAGCAAG
GCCCCTCAGGACCCGGTTCTATGGCTGCAGCCGCTGCTGCAGCTATGGGACCCGGTGGTT
ACGGCCCTGGACAGCAAGGACCTGGCGGATACGGTTCTCCCGCTGGCGCTCCCCCAGGCC
CACCCGGACCCCCCGGACCACCTGGCCCCCCTGGTCCACCTGGCCCACCTGGCCCACCTG
GACCACCAGGACCACCAGGTCCCCCCGGCCCTCCCGGTCCTCCCGGACCCCCTGGACCCC
CTGGACCCGCCGGTGGATGATAAGCTTCTATGTAACTAACGAAACAGCATGTACTAATAG
AACCGTATCGAGAATATTTATTTAGGTGAGTAGTAGGAGTGAACCAGACAGTCAATTTAG
TGAGCTGTCCCAGCTTTTGTGCATTCCAGAATTGCCGGTCAAATTGGTTATGGGTTATGG
GGCTTTTCCGATTGAGGTTCAGTTTCTGCGGTTATCTCTTTCTTGACCTGGTCTTTTACA
GGCTGTTCTTTCTCCCCATGATTATTCTTTAGCTGAAGATACCGCTTAGCCTGATAATGT
CGTCGTTTTGTAATCAAAATCTTTAGTTGGGCATCGTCTGAGGTTTCCTTTGGCTTCTGG
GGTTGTTAGTAGGAACGTAGGAACCATAGTAACTTTTACACATACATTCTTATGATTGCG
AAGTAAGCTGAGTCTGCTGCTTGGCTCCCGAAGTACTTTCTCTTTCTCTACCGGTTGATT
CTCCTTCTGGTGCTCCTAAACGATTGTGTTAGAAGGGATTGACCGCTCCGGCATGCCGGC
ATGGACATGGAGGCCCAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGGGCA
TGATGTGACTGTCGCCCGTACATTTAGCCCATACATCCCCATGTATAATCATTTGCATCC
ATACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTACGGCTCCTCGCTGCAGACCTGC
GAGCAGGGAAACGCTCCCCTCACAGACGCGTTGAATTGTCCCCACGCCGCGCCCCTGTAG
AGAAATATAAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTTCATATACTTCCTTTTAAA
ATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAACAACCATGGGTACCACTC
TTGACGACACGGCTTACCGGTACCGCACCAGTGTCCCGGGGGACGCCGAGGCCATCGAGG
CACTGGATGGGTCCTTCACCACCGACACCGTTTTTCGCGTCACCGCCACCGGGGACGGCT
TCACCCTGCGGGAGGTGCCGGTGGACCCGCCCCTGACCAAGGTGTTCCCCGACGACGAAT
CGGACGACGAATCGGACGACGGGGAGGACGGCGACCCGGACTCCCGGACGTTCGTCGCGT
ACGGGGACGACGGCGACCTGGCGGGCTTCGTGGTCGTCTCGTACTCCGGCTGGAACCGCC
GGCTGACCGTCGAGGACATCGAGGTCGCCCCGGAGCACCGGGGGCACGGGGTCGGGCGCG
CGTTGATGGGGCTCGCGACGGAGTTCGCCCGCGAGCGGGGCGCCGGGCACCTCTGGCTGG
AGGTCACCAACGTCAACGCACCGGCGATCCACGCGTACCGGCGGATGGGGTTCACCCTCT

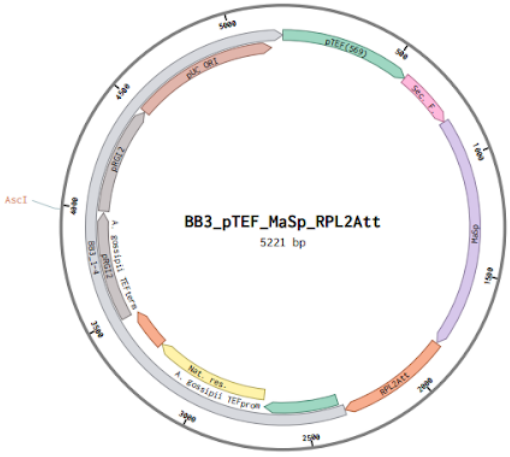
GCGGCCTGGACACCGCCCTGTACGACGGCACCGCCTCGGACGGCGAGCAGGCGCTCTACA
TGAGCATGCCCTGCCCCTAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAACTT
GTCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCAAATGTTAGCGTGAT
TTATATTTTTTTTCGCCTCGACATCATCTGCCCAGATGCGAAGTTAAGTGCGCAGAAAGT
AATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTGGCTTCATACCCAGCATT
GACCTTTGGTATGAGCATCTGAAAAACAACCAGGTGTTGCAAAGTTAAACATCCTTCTTT
GTTCATATAGAACCCACTATTCATGGTACTCCCCAATCGAATTTCACATTCTGGTTTTGA
AATTACACACCACGTTAGCTTATAAGATTTCATATAACTTATTGATATACGGTTTCCATT
GTTCGAATAGTTGAGGTTGTATGTAATTCGATTGAAGGGGCCATTTTTGTTTCCTACTTT
TCCTGGGAGCTTATCCGATGCGCTTCAAAGCTGGAATTGTAAATATAGAGAAAAAGAAGG
ATGTTGTTTTATTCTTGAAAGAGTATAATTTTACTTCTAGCAACTCTCCCACTTCGCTTG
ACTTCATTTATTTCTTGGGCACATAGGCGTAGTAATCTAGACCAACAGATAATTTGCCGG
AATGATATAGCGATTGGAAAATGAACTGAAATTTTTTGCTGTCTTTCAATTTGACGGGCA
GTTCATCAGGCGCGCCGTGACCGACCATATAAATACGTTGAGAATGTTATTCTTCCTCGT
AGTTGAAGTGGCTTCATAATTTCAGAACTCAATAGATAAACTAGGATGTTTTAAAGCAAT
TAATGCTCACAAGTAAGGAGCGACTCTCTTGCTTTTCGAATACTAAAAGTATCGTCCCAA
CCCAGAAAAAAAGACCTCTTAACTGCAAAATAAACTCTATATATTTCTTCTAAAACAGTT
TCAGGTTGGATAGTATCGCATTCTCATCACTTCTAACTAGTAGGCCATGAGATATATTAA
CGTTTACTTGAGTTCTAAGTTCTCCGAATTAGATGCACAGCACAAACAAGATTAGGTTTC
ACTTGGTACAAAATACGAACAGAGTTTAAGGTCGTAATTTCATTTCGTTATTGATCCCCA
CAATCTATTCTTATCACAGTCATCAGATAGTCGCGAAAAAGCATGCAGAAAAGGGGGTCG
TCCCTATCTAAGTTGTAGCATTACAACAAATATGACCGCTGTGAGCAAAAGGCCAGCAAA
AGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTG
ACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA
GATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGC
TTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAC
GCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAAC
CCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGG
TAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGT
ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAA
CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCT
CTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGA
TTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTGGAGCG
CAAAAAACCCCGCCCCTGACAGGGCGGGGTTTTTTCGCGATCGG

BB3-pTEF-MaSp-RPL2Att

The first created backbone 3 is the plasmid BB3-pTEF-MaSp-RPL2Att. It contains the promoter “pTEF”, the terminator “RPL2Att” and the MaSp insert including the secretion signal. This plasmid was synthesized by doing a Golden Gate reaction including the Backbone 1 with the MaSp coding sequence and the secretion signal for Pichia Pastoris, the “pTEF” promoter and the “RPL2Att terminator. The restriction enzyme BbsI was used to create the desired sticky ends. For the empty backbone the backbone 3 version, building the fusion sites 1 and 4, when digested with BbsI, was chosen.

After an incubation period of 18 hours at 37 degrees Celsius, colony growth could only be witnessed on the nourseothricin-plates containing the complete BB3-pTEF-

MaSp-RPL2Att plasmid but not on the control plates only containing the empty and, after the Golden Gate, linearized backbone 3. As further proof for the successful cloning a control digest was performed using the restriction enzyme BglI. After loading the digest on a 1% agarose gel, the bands could be found at the expected sizes.



1. Insert:

- Major Spider Silk (MaSp)
- Secretion Factor (Sec. F.) for Pichia Pastoris

2.1. Amino Acid Sequence:

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPFSNSTN
NGLLFINTTIASIAAKEEGVSLEKRGAPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
PPGPPGPPGPPGPAGASGPGQQGRGGYGPGQQGPSGPGSGAAAAAAGGAGQGGYGGLGGQ
GAGRGGLGGQAAAAAASGPGGYGPGQQGPGGYGPGQQGPSGPGSMAAAAAAAMGPGGYGP
GQQGPGGYGSPAGASGPGQQGRGGYGPGQQGPSGPGSGAAAAAAGGAGQGGYGGLGGQGA
GRGGLGGQAAAAAASGPGGYGPGQQGPGGYGPGQQGPSGPGSMAAAAAAAMGPGGYGPGQ
QGPGGYGSPAGAPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPAG
G**

2.2. Nucleic Acid Sequence:

ATGAGATTCCCATCTATTTTCACCGCTGTCTTGTTCGCTGCCTCCTCTGCATTGGCTGCC
CCTGTTAACACTACCACTGAAGATGAGACTGCTCAAATTCCAGCTGAAGCAGTTATCGGT
TACTCTGACCTTGAGGGTGATTTCGACGTCGCTGTTTTGCCTTTCTCTAACTCCACTAAC
AACGGTTTGTTGTTCATTAACACCACTATCGCTTCCATTGCTGCTAAGGAAGAGGGTGTC
TCTCTCGAGAAGAGAGGCGCTCCCCCAGGCCCACCCGGACCCCCCGGACCACCTGGCCCC
CCTGGTCCACCTGGCCCACCTGGCCCACCTGGACCACCAGGACCACCAGGTCCCCCCGGC
CCTCCCGGTCCTCCCGGACCCCCTGGACCCCCTGGACCCGCCGGCGCAAGTGGTCCAGGA
CAGCAGGGTAGGGGAGGTTACGGTCCCGGTCAACAAGGTCCATCTGGTCCTGGTTCCGGT
GCTGCAGCAGCTGCCGCTGGAGGTGCTGGTCAGGGCGGATATGGTGGTCTTGGCGGTCAA

GGTGCTGGAAGAGGTGGATTAGGAGGACAGGCCGCAGCCGCTGCTGCCTCTGGTCCCGGA
GGATATGGACCTGGACAGCAAGGTCCAGGCGGATATGGTCCCGGTCAGCAAGGCCCCTCA
GGACCCGGTTCTATGGCTGCAGCCGCTGCTGCAGCTATGGGACCCGGTGGTTACGGCCCT
GGACAGCAAGGACCTGGCGGATACGGTTCTCCCGCTGGCGCAAGTGGTCCAGGACAGCAG
GGTAGGGGAGGTTACGGTCCCGGTCAACAAGGTCCATCTGGTCCTGGTTCCGGTGCTGCA
GCAGCTGCCGCTGGAGGTGCTGGTCAGGGCGGATATGGTGGTCTTGGCGGTCAAGGTGCT
GGAAGAGGTGGATTAGGAGGACAGGCCGCAGCCGCTGCTGCCTCTGGTCCCGGAGGATAT
GGACCTGGACAGCAAGGTCCAGGCGGATATGGTCCCGGTCAGCAAGGCCCCTCAGGACCC
GGTTCTATGGCTGCAGCCGCTGCTGCAGCTATGGGACCCGGTGGTTACGGCCCTGGACAG
CAAGGACCTGGCGGATACGGTTCTCCCGCTGGCGCTCCCCCAGGCCCACCCGGACCCCCC
GGACCACCTGGCCCCCCTGGTCCACCTGGCCCACCTGGCCCACCTGGACCACCAGGACCA
CCAGGTCCCCCCGGCCCTCCCGGTCCTCCCGGACCCCCTGGACCCCCTGGACCCGCCGGT
GGATGATAA

2. Promoter:

- pTEF-Promoter

3.1.Nucleic Acid Sequence:

CGGTACTACCATTGACACTACACATCCTTAATTCCAATCCTGTCTGGCCTCCTTCACCTT
TTAACCATCTTGCCCATTCCAACTCGTGTCAGATTGCGTATCAAGTGAAAAAAAAAAAAT
TTTAAATCTTTAACCCAATCAGGTAATAACTGTCGCCTCTTTTATCTGCCGCACTGCATG
AGGTGTCCCCTTAGTGGGAAAGAGTACTGAGCCAACCCTGGAGGACAGCAAGGGAAAAAT
ACCTACAACTTGCTTCATAATGGTCGTAAAAACAATCCTTGTCGGATATAAGTGTTGTAG
ACTGTCCCTTATCCTCTGCGATGTTCTTCCTCTCAAAGTTTGCGATTTCTCTCTATCAGA
ATTGCCATCAAGAGACTCAGGACTAATTTCGCAGTCCCACACGCACTCGTACATGATTGG
CTGAAATTTCCCTAAAGAATTTCTTTTTCACGAAAATTTTTTTTTTACACAAGATTTTCA
GCAGATATAAAATGGAGAGCAGGACCTCCGCTGTGACTCTTCTTTTTTTTCTTTTATTCT
CACTACATACATTTTAGTTATTCGCCAAC

3. Terminator:

- RPL2Att-Terminator

4.1. Nucleic Acid Sequence:

CTATGTAACTAACGAAACAGCATGTACTAATAGAACCGTATCGAGAATATTTATTTAGGT
GAGTAGTAGGAGTGAACCAGACAGTCAATTTAGTGAGCTGTCCCAGCTTTTGTGCATTCC
AGAATTGCCGGTCAAATTGGTTATGGGTTATGGGGCTTTTCCGATTGAGGTTCAGTTTCT
GCGGTTATCTCTTTCTTGACCTGGTCTTTTACAGGCTGTTCTTTCTCCCCATGATTATTC
TTTAGCTGAAGATACCGCTTAGCCTGATAATGTCGTCGTTTTGTAATCAAAATCTTTAGT
TGGGCATCGTCTGAGGTTTCCTTTGGCTTCTGGGGTTGTTAGTAGGAACGTAGGAACCAT
AGTAACTTTTACACATACATTCTTATGATTGCGAAGTAAGCTGAGTCTGCTGCTTGGCTC
CCGAAGTACTTTCTCTTTCTCTACCGGTTGATTCTCCTTCTGGTGCTCCTAAACGATTGT
GTTAGAAGGGATTGAC

4. Backbone:

- Backbone 3 (BB3)
- Fusion-Site 1-4
- Nourseothrycin resistence (Nat. res.)

5.1. Nucleic Acid Sequence:

CGCTCCGGCATGCCGGCATGGACATGGAGGCCCAGAATACCCTCCTTGACAGTCTTGACG
TGCGCAGCTCAGGGGCATGATGTGACTGTCGCCCGTACATTTAGCCCATACATCCCCATG
TATAATCATTTGCATCCATACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTACGGCT
CCTCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCTCACAGACGCGTTGAATTGTCCCC
ACGCCGCGCCCCTGTAGAGAAATATAAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTTC
ATATACTTCCTTTTAAAATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAAC
AACCATGGGTACCACTCTTGACGACACGGCTTACCGGTACCGCACCAGTGTCCCGGGGGA
CGCCGAGGCCATCGAGGCACTGGATGGGTCCTTCACCACCGACACCGTTTTTCGCGTCAC
CGCCACCGGGGACGGCTTCACCCTGCGGGAGGTGCCGGTGGACCCGCCCCTGACCAAGGT
GTTCCCCGACGACGAATCGGACGACGAATCGGACGACGGGGAGGACGGCGACCCGGACTC
CCGGACGTTCGTCGCGTACGGGGACGACGGCGACCTGGCGGGCTTCGTGGTCGTCTCGTA
CTCCGGCTGGAACCGCCGGCTGACCGTCGAGGACATCGAGGTCGCCCCGGAGCACCGGGG
GCACGGGGTCGGGCGCGCGTTGATGGGGCTCGCGACGGAGTTCGCCCGCGAGCGGGGCGC
CGGGCACCTCTGGCTGGAGGTCACCAACGTCAACGCACCGGCGATCCACGCGTACCGGCG
GATGGGGTTCACCCTCTGCGGCCTGGACACCGCCCTGTACGACGGCACCGCCTCGGACGG
CGAGCAGGCGCTCTACATGAGCATGCCCTGCCCCTAATCAGTACTGACAATAAAAAGATT
CTTGTTTTCAAGAACTTGTCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAA
TCAAATGTTAGCGTGATTTATATTTTTTTTCGCCTCGACATCATCTGCCCAGATGCGAAG
TTAAGTGCGCAGAAAGTAATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTG
GCTTCATACCCAGCATTGACCTTTGGTATGAGCATCTGAAAAACAACCAGGTGTTGCAAA
GTTAAACATCCTTCTTTGTTCATATAGAACCCACTATTCATGGTACTCCCCAATCGAATT
TCACATTCTGGTTTTGAAATTACACACCACGTTAGCTTATAAGATTTCATATAACTTATT
GATATACGGTTTCCATTGTTCGAATAGTTGAGGTTGTATGTAATTCGATTGAAGGGGCCA
TTTTTGTTTCCTACTTTTCCTGGGAGCTTATCCGATGCGCTTCAAAGCTGGAATTGTAAA
TATAGAGAAAAAGAAGGATGTTGTTTTATTCTTGAAAGAGTATAATTTTACTTCTAGCAA
CTCTCCCACTTCGCTTGACTTCATTTATTTCTTGGGCACATAGGCGTAGTAATCTAGACC
AACAGATAATTTGCCGGAATGATATAGCGATTGGAAAATGAACTGAAATTTTTTGCTGTC
TTTCAATTTGACGGGCAGTTCATCAGGCGCGCCGTGACCGACCATATAAATACGTTGAGA
ATGTTATTCTTCCTCGTAGTTGAAGTGGCTTCATAATTTCAGAACTCAATAGATAAACTA
GGATGTTTTAAAGCAATTAATGCTCACAAGTAAGGAGCGACTCTCTTGCTTTTCGAATAC
TAAAAGTATCGTCCCAACCCAGAAAAAAAGACCTCTTAACTGCAAAATAAACTCTATATA
TTTCTTCTAAAACAGTTTCAGGTTGGATAGTATCGCATTCTCATCACTTCTAACTAGTAG
GCCATGAGATATATTAACGTTTACTTGAGTTCTAAGTTCTCCGAATTAGATGCACAGCAC
AAACAAGATTAGGTTTCACTTGGTACAAAATACGAACAGAGTTTAAGGTCGTAATTTCAT
TTCGTTATTGATCCCCACAATCTATTCTTATCACAGTCATCAGATAGTCGCGAAAAAGCA
TGCAGAAAAGGGGGTCGTCCCTATCTAAGTTGTAGCATTACAACAAATATGACCGCTGTG
AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA
TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA
CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC
TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC
GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT
GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG
TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG
GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA
CGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG
AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT
TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT

TTCTACGGGGTGGAGCGCAAAAAACCCCGCCCCTGACAGGGCGGGGTTTTTTCGCGATCG
G

5. Complete Plasmid:

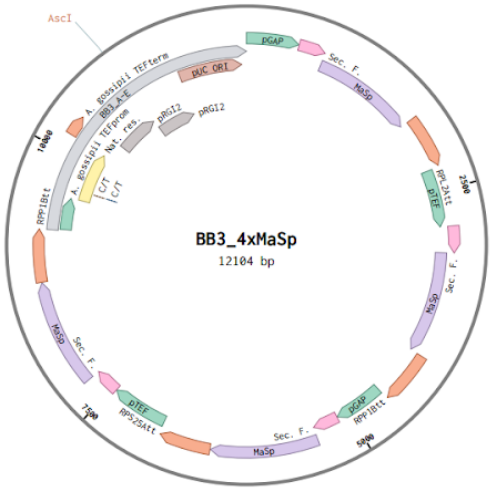
6.1. Amino Acid Sequence:

CGGTACTACCATTGACACTACACATCCTTAATTCCAATCCTGTCTGGCCTCCTTCACCTT
TTAACCATCTTGCCCATTCCAACTCGTGTCAGATTGCGTATCAAGTGAAAAAAAAAAAAT
TTTAAATCTTTAACCCAATCAGGTAATAACTGTCGCCTCTTTTATCTGCCGCACTGCATG
AGGTGTCCCCTTAGTGGGAAAGAGTACTGAGCCAACCCTGGAGGACAGCAAGGGAAAAAT
ACCTACAACTTGCTTCATAATGGTCGTAAAAACAATCCTTGTCGGATATAAGTGTTGTAG
ACTGTCCCTTATCCTCTGCGATGTTCTTCCTCTCAAAGTTTGCGATTTCTCTCTATCAGA
ATTGCCATCAAGAGACTCAGGACTAATTTCGCAGTCCCACACGCACTCGTACATGATTGG
CTGAAATTTCCCTAAAGAATTTCTTTTTCACGAAAATTTTTTTTTTACACAAGATTTTCA
GCAGATATAAAATGGAGAGCAGGACCTCCGCTGTGACTCTTCTTTTTTTTCTTTTATTCT
CACTACATACATTTTAGTTATTCGCCAACATGAGATTCCCATCTATTTTCACCGCTGTCT
TGTTCGCTGCCTCCTCTGCATTGGCTGCCCCTGTTAACACTACCACTGAAGATGAGACTG
CTCAAATTCCAGCTGAAGCAGTTATCGGTTACTCTGACCTTGAGGGTGATTTCGACGTCG
CTGTTTTGCCTTTCTCTAACTCCACTAACAACGGTTTGTTGTTCATTAACACCACTATCG
CTTCCATTGCTGCTAAGGAAGAGGGTGTCTCTCTCGAGAAGAGAGGCGCTCCCCCAGGCC
CACCCGGACCCCCCGGACCACCTGGCCCCCCTGGTCCACCTGGCCCACCTGGCCCACCTG
GACCACCAGGACCACCAGGTCCCCCCGGCCCTCCCGGTCCTCCCGGACCCCCTGGACCCC
CTGGACCCGCCGGCGCAAGTGGTCCAGGACAGCAGGGTAGGGGAGGTTACGGTCCCGGTC
AACAAGGTCCATCTGGTCCTGGTTCCGGTGCTGCAGCAGCTGCCGCTGGAGGTGCTGGTC
AGGGCGGATATGGTGGTCTTGGCGGTCAAGGTGCTGGAAGAGGTGGATTAGGAGGACAGG
CCGCAGCCGCTGCTGCCTCTGGTCCCGGAGGATATGGACCTGGACAGCAAGGTCCAGGCG
GATATGGTCCCGGTCAGCAAGGCCCCTCAGGACCCGGTTCTATGGCTGCAGCCGCTGCTG
CAGCTATGGGACCCGGTGGTTACGGCCCTGGACAGCAAGGACCTGGCGGATACGGTTCTC
CCGCTGGCGCAAGTGGTCCAGGACAGCAGGGTAGGGGAGGTTACGGTCCCGGTCAACAAG
GTCCATCTGGTCCTGGTTCCGGTGCTGCAGCAGCTGCCGCTGGAGGTGCTGGTCAGGGCG
GATATGGTGGTCTTGGCGGTCAAGGTGCTGGAAGAGGTGGATTAGGAGGACAGGCCGCAG
CCGCTGCTGCCTCTGGTCCCGGAGGATATGGACCTGGACAGCAAGGTCCAGGCGGATATG
GTCCCGGTCAGCAAGGCCCCTCAGGACCCGGTTCTATGGCTGCAGCCGCTGCTGCAGCTA
TGGGACCCGGTGGTTACGGCCCTGGACAGCAAGGACCTGGCGGATACGGTTCTCCCGCTG
GCGCTCCCCCAGGCCCACCCGGACCCCCCGGACCACCTGGCCCCCCTGGTCCACCTGGCC
CACCTGGCCCACCTGGACCACCAGGACCACCAGGTCCCCCCGGCCCTCCCGGTCCTCCCG
GACCCCCTGGACCCCCTGGACCCGCCGGTGGATGATAAGCTTCTATGTAACTAACGAAAC
AGCATGTACTAATAGAACCGTATCGAGAATATTTATTTAGGTGAGTAGTAGGAGTGAACC
AGACAGTCAATTTAGTGAGCTGTCCCAGCTTTTGTGCATTCCAGAATTGCCGGTCAAATT
GGTTATGGGTTATGGGGCTTTTCCGATTGAGGTTCAGTTTCTGCGGTTATCTCTTTCTTG
ACCTGGTCTTTTACAGGCTGTTCTTTCTCCCCATGATTATTCTTTAGCTGAAGATACCGC
TTAGCCTGATAATGTCGTCGTTTTGTAATCAAAATCTTTAGTTGGGCATCGTCTGAGGTT
TCCTTTGGCTTCTGGGGTTGTTAGTAGGAACGTAGGAACCATAGTAACTTTTACACATAC
ATTCTTATGATTGCGAAGTAAGCTGAGTCTGCTGCTTGGCTCCCGAAGTACTTTCTCTTT
CTCTACCGGTTGATTCTCCTTCTGGTGCTCCTAAACGATTGTGTTAGAAGGGATTGACCG
CTCCGGCATGCCGGCATGGACATGGAGGCCCAGAATACCCTCCTTGACAGTCTTGACGTG
CGCAGCTCAGGGGCATGATGTGACTGTCGCCCGTACATTTAGCCCATACATCCCCATGTA
TAATCATTTGCATCCATACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTACGGCTCC

TCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCTCACAGACGCGTTGAATTGTCCCCAC
GCCGCGCCCCTGTAGAGAAATATAAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTTCAT
ATACTTCCTTTTAAAATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAACAA
CCATGGGTACCACTCTTGACGACACGGCTTACCGGTACCGCACCAGTGTCCCGGGGGACG
CCGAGGCCATCGAGGCACTGGATGGGTCCTTCACCACCGACACCGTTTTTCGCGTCACCG
CCACCGGGGACGGCTTCACCCTGCGGGAGGTGCCGGTGGACCCGCCCCTGACCAAGGTGT
TCCCCGACGACGAATCGGACGACGAATCGGACGACGGGGAGGACGGCGACCCGGACTCCC
GGACGTTCGTCGCGTACGGGGACGACGGCGACCTGGCGGGCTTCGTGGTCGTCTCGTACT
CCGGCTGGAACCGCCGGCTGACCGTCGAGGACATCGAGGTCGCCCCGGAGCACCGGGGGC
ACGGGGTCGGGCGCGCGTTGATGGGGCTCGCGACGGAGTTCGCCCGCGAGCGGGGCGCCG
GGCACCTCTGGCTGGAGGTCACCAACGTCAACGCACCGGCGATCCACGCGTACCGGCGGA
TGGGGTTCACCCTCTGCGGCCTGGACACCGCCCTGTACGACGGCACCGCCTCGGACGGCG
AGCAGGCGCTCTACATGAGCATGCCCTGCCCCTAATCAGTACTGACAATAAAAAGATTCT
TGTTTTCAAGAACTTGTCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATC
AAATGTTAGCGTGATTTATATTTTTTTTCGCCTCGACATCATCTGCCCAGATGCGAAGTT
AAGTGCGCAGAAAGTAATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTGGC
TTCATACCCAGCATTGACCTTTGGTATGAGCATCTGAAAAACAACCAGGTGTTGCAAAGT
TAAACATCCTTCTTTGTTCATATAGAACCCACTATTCATGGTACTCCCCAATCGAATTTC
ACATTCTGGTTTTGAAATTACACACCACGTTAGCTTATAAGATTTCATATAACTTATTGA
TATACGGTTTCCATTGTTCGAATAGTTGAGGTTGTATGTAATTCGATTGAAGGGGCCATT
TTTGTTTCCTACTTTTCCTGGGAGCTTATCCGATGCGCTTCAAAGCTGGAATTGTAAATA
TAGAGAAAAAGAAGGATGTTGTTTTATTCTTGAAAGAGTATAATTTTACTTCTAGCAACT
CTCCCACTTCGCTTGACTTCATTTATTTCTTGGGCACATAGGCGTAGTAATCTAGACCAA
CAGATAATTTGCCGGAATGATATAGCGATTGGAAAATGAACTGAAATTTTTTGCTGTCTT
TCAATTTGACGGGCAGTTCATCAGGCGCGCCGTGACCGACCATATAAATACGTTGAGAAT
GTTATTCTTCCTCGTAGTTGAAGTGGCTTCATAATTTCAGAACTCAATAGATAAACTAGG
ATGTTTTAAAGCAATTAATGCTCACAAGTAAGGAGCGACTCTCTTGCTTTTCGAATACTA
AAAGTATCGTCCCAACCCAGAAAAAAAGACCTCTTAACTGCAAAATAAACTCTATATATT
TCTTCTAAAACAGTTTCAGGTTGGATAGTATCGCATTCTCATCACTTCTAACTAGTAGGC
CATGAGATATATTAACGTTTACTTGAGTTCTAAGTTCTCCGAATTAGATGCACAGCACAA
ACAAGATTAGGTTTCACTTGGTACAAAATACGAACAGAGTTTAAGGTCGTAATTTCATTT
CGTTATTGATCCCCACAATCTATTCTTATCACAGTCATCAGATAGTCGCGAAAAAGCATG
CAGAAAAGGGGGTCGTCCCTATCTAAGTTGTAGCATTACAACAAATATGACCGCTGTGAG
CAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATA
GGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC
CGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTG
TTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGC
TTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGG
GCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTC
TTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA
TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACG
GCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA
AAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTG
TTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTT
CTACGGGGTGGAGCGCAAAAAACCCCGCCCCTGACAGGGCGGGGTTTTTTCGCGATCGGA
G

BB3-4xMaSp

The third plasmid version was created out of the four previously described backbone 2 versions. These backbone 2 versions were used in a Golden Gate reaction with BsaI as restriction enzyme, also including an empty backbone 3 containing the fusion sites A and E after digest with BsaI.
After an incubation period of 18 hours at 37 degrees Celsius, extensive colony growth could be witnessed on the nourseothricin-plates, whereas only few colonies appeared on the plates only containing the empty and, after the Golden Gate, linearized backbone 3. A first control digest with the enzyme BglI was not entirely rewarding since it led to several not very distinctive bands. After some consideration a different restriction enzyme, XhoI, was chosen for a second digest. The second digest was more successful, resulting in bands at the expected sizes on a 1% agarose gel.



1. Expression cassettes

1.1. Cassette 1

- Promoter: pGAP
- Terminator: RPL2Att
- Insert: MaSp coding sequence + secretion factor for Pichia Pastoris

1.2. Cassette 2

- Promoter: pTEF
- Terminator: RPP1Btt
- Insert: MaSp coding sequence + secretion factor for Pichia Pastoris

1.3.Cassette 3

- Promoter: pGAP
- Terminator: RPS25Att
- Insert: MaSp coding sequence + secretion factor for Pichia Pastoris

1.4.Cassette 4

- Promoter: pTEF
- Terminator: RPL2Att
- Insert: MaSp coding sequence + secretion factor for Pichia Pastoris

2. Insert:

- Major Spider Silk (MaSp)
- Secretion Factor (Sec. F.) for Pichia Pastoris

2.1. Amino Acid Sequence:

```
MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPFSNSTN
NGLLFINTTIASIAAKEEGVSLEKRGAPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
PPGPPGPPGPPGPAGASGPGQQGRGGYGPGQQGPSGPGSGAAAAAAGGAGQGGYGGLGGQ
GAGRGGLGGQAAAAAASGPGGYGPGQQGPGGYGPGQQGPSGPGSMAAAAAAAMGPGGYGP
GQQGPGGYGSPAGASGPGQQGRGGYGPGQQGPSGPGSGAAAAAAGGAGQGGYGGLGGQGA
GRGGLGGQAAAAAASGPGGYGPGQQGPGGYGPGQQGPSGPGSMAAAAAAAMGPGGYGPGQ
QGPGGYGSPAGAPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPAG
G**
```

2.2. Nucleic Acid Sequence:

```
ATGAGATTCCCATCTATTTTCACCGCTGTCTTGTTCGCTGCCTCCTCTGCATTGGCTGCC
CCTGTTAACACTACCACTGAAGATGAGACTGCTCAAATTCCAGCTGAAGCAGTTATCGGT
TACTCTGACCTTGAGGGTGATTTCGACGTCGCTGTTTTGCCTTTCTCTAACTCCACTAAC
AACGGTTTGTTGTTCATTAACACCACTATCGCTTCCATTGCTGCTAAGGAAGAGGGTGTC
TCTCTCGAGAAGAGAGGCGCTCCCCCAGGCCCACCCGGACCCCCCGGACCACCTGGCCCC
CCTGGTCCACCTGGCCCACCTGGCCCACCTGGACCACCAGGACCACCAGGTCCCCCCGGC
CCTCCCGGTCCTCCCGGACCCCCTGGACCCCCTGGACCCGCCGGCGCAAGTGGTCCAGGA
CAGCAGGGTAGGGGAGGTTACGGTCCCGGTCAACAAGGTCCATCTGGTCCTGGTTCCGGT
GCTGCAGCAGCTGCCGCTGGAGGTGCTGGTCAGGGCGGATATGGTGGTCTTGGCGGTCAA
GGTGCTGGAAGAGGTGGATTAGGAGGACAGGCCGCAGCCGCTGCTGCCTCTGGTCCCGGA
GGATATGGACCTGGACAGCAAGGTCCAGGCGGATATGGTCCCGGTCAGCAAGGCCCCTCA
GGACCCGGTTCTATGGCTGCAGCCGCTGCTGCAGCTATGGGACCCGGTGGTTACGGCCCT
GGACAGCAAGGACCTGGCGGATACGGTTCTCCCGCTGGCGCAAGTGGTCCAGGACAGCAG
GGTAGGGGAGGTTACGGTCCCGGTCAACAAGGTCCATCTGGTCCTGGTTCCGGTGCTGCA
GCAGCTGCCGCTGGAGGTGCTGGTCAGGGCGGATATGGTGGTCTTGGCGGTCAAGGTGCT
GGAAGAGGTGGATTAGGAGGACAGGCCGCAGCCGCTGCTGCCTCTGGTCCCGGAGGATAT
GGACCTGGACAGCAAGGTCCAGGCGGATATGGTCCCGGTCAGCAAGGCCCCTCAGGACCC
GGTTCTATGGCTGCAGCCGCTGCTGCAGCTATGGGACCCGGTGGTTACGGCCCTGGACAG
CAAGGACCTGGCGGATACGGTTCTCCCGCTGGCGCTCCCCCAGGCCCACCCGGACCCCCC
GGACCACCTGGCCCCCCTGGTCCACCTGGCCCACCTGGCCCACCTGGACCACCAGGACCA
CCAGGTCCCCCCGGCCCTCCCGGTCCTCCCGGACCCCCTGGACCCCCTGGACCCGCCGGT
GGATGATAA
```

3. Promoters

3.1.pGAP-Promoter

3.1.1. Nucleic Acid Sequence

AGGATCCTTTTTTGTAGAAATGTCTTGGTGTCCTCGTCCAATCAGGTAGCCATCTCTGAA
ATATCTGGCTCCGTTGCAACTCCGAACGACCTGCTGGCAACGTAAAATTCTCCGGGGTAA
AACTTAAATGTGGAGTAATGGAACCAGAAACGTCTCTTCCCTTCTCTCTCCTTCCACCGC
CCGTTACCGTCCCTAGGAAATTTTACTCTGCTGGAGAGCTTCTTCTACGGCCCCCTTGCA
GCAATGCTCTTCCCAGCATTACGTTGCGGGTAAAACGGAGGTCGTGTACCCGACCTAGCA
GCCCAGGGATGGAAAAGTCCCGGCCGTCGCTGGCAATAATAGCGGGCGGACGCATGTCAT
GAGATTATTGGAAACCACCAGAATCGAATATAAAAGGCGAACACCTTTCCCAATTTTGGT
TTCTCCTGACCCAAAGACTTTAAATTTAATTTATTTGTCCCTATTTCAATCAATTGAACA
ACTATCAAAACAC

3.2.pTEF-Promoter

3.2.1. Nucleic Acid Sequence

CGGTACTACCATTGACACTACACATCCTTAATTCCAATCCTGTCTGGCCTCCTTCACCTT
TTAACCATCTTGCCCATTCCAACTCGTGTCAGATTGCGTATCAAGTGAAAAAAAAAAAAT
TTTAAATCTTTAACCCAATCAGGTAATAACTGTCGCCTCTTTTATCTGCCGCACTGCATG
AGGTGTCCCCTTAGTGGGAAAGAGTACTGAGCCAACCCTGGAGGACAGCAAGGGAAAAAT
ACCTACAACTTGCTTCATAATGGTCGTAAAAACAATCCTTGTCGGATATAAGTGTTGTAG
ACTGTCCCTTATCCTCTGCGATGTTCTTCCTCTCAAAGTTTGCGATTTCTCTCTATCAGA
ATTGCCATCAAGAGACTCAGGACTAATTTCGCAGTCCCACACGCACTCGTACATGATTGG
CTGAAATTTCCCTAAAGAATTTCTTTTTCACGAAAATTTTTTTTTTACACAAGATTTTCA
GCAGATATAAAATGGAGAGCAGGACCTCCGCTGTGACTCTTCTTTTTTTTCTTTTATTCT
CACTACATACATTTTAGTTATTCGCCAAC

4. Terminators

4.1.RPL2Att

4.1.1. Nucleic Acid Sequence

CTATGTAACTAACGAAACAGCATGTACTAATAGAACCGTATCGAGAATATTTATTTAGGT
GAGTAGTAGGAGTGAACCAGACAGTCAATTTAGTGAGCTGTCCCAGCTTTTGTGCATTCC
AGAATTGCCGGTCAAATTGGTTATGGGTTATGGGGCTTTTCCGATTGAGGTTCAGTTTCT
GCGGTTATCTCTTTCTTGACCTGGTCTTTTACAGGCTGTTCTTTCTCCCCATGATTATTC
TTTAGCTGAAGATACCGCTTAGCCTGATAATGTCGTCGTTTTGTAATCAAAATCTTTAGT
TGGGCATCGTCTGAGGTTTCCTTTGGCTTCTGGGGTTGTTAGTAGGAACGTAGGAACCAT
AGTAACTTTTACACATACATTCTTATGATTGCGAAGTAAGCTGAGTCTGCTGCTTGGCTC
CCGAAGTACTTTCTCTTTCTCTACCGGTTGATTCTCCTTCTGGTGCTCCTAAACGATTGT
GTTAGAAGGGATTGAC

4.2.RPP1Btt

4.2.1. Nucleic Acid Sequence

GGATGATACTTTAATTGATGCATGTGAAGAATTGAAGAAACAGTAATCTAGTAAAATATA
AGAAAGTCCTGAAAGTAGTTCATTTAGTAAGCTATGAACCCAAGTTGGGGTAACCGATTT
TTCAGGTAAACAGCTGTAGTGAGGAGACATCTCAAATCGCCTGAGAAGAAAAATATCTAT
GAGACAACTACAATGCGAAGAATGTGTTCAAATGCTCTATTATTCACTAGTTTACCTATT
CAATGACCAACTCCTGGGAGGCTTCGTTACCGATAATTGGTAACTTAACTAGTACGTCCT
TCATCAACTCGTAGACAGTGAAAGTGACGGCTTGTCCCGGAGCAACACGCATGATACGCG

GCGTGATACCTTTGTACAGAGCACGCATACCTTCCTCTTGGATCAATCTCTTGGAAATAC
GAAGAATACGGGCCACACCAGACTCATTGGTGACACCAGTCTCTCTCTGCATACGAGTCT
TGATGGTATCTAGAGGAGC

4.3.RPS25Att

4.3.1. Nucleic Acid Sequence

TTAGTGTACATCTGATAATATAGTACTACCACGTATGATAATGTAGAGAATAGGCTTCCT
TGTCGAGTGTGTTTGCAGTTTTCTTGAGTTTCAAGGTTTAAATGCTGGTATATTAGTTCA
TCGAAGGTTTCAGCCAATAGCACCTTAAATCAATCAAACTAATTCGACTCTTACGAAAGA
GCCTACTGTGTTTAGTATCGAAGTCGTTTACCTTTCATGTTGAATAGCTTCCTCTCTGAC
CCTAACATTTCAAGATCCTCCTAAAGTTACCCGGATTGTGAAATTCTAATGATCCACCTG
CCCAATGCATTTTTTCTTTATTCAGTTTACCTTTTTTACCTAATATACGAGCTTGTTAAA
GTAAGTGGCACTGCAATACTAGGCTTATTGTTGATATTATGATGAATCGTTTTCACAAAC
TTGATTTCCTGTGAACTCACCATGTACTAAGGAAAAAAACATGCATCACCATCTGAATAT
TTGAC

5. Backbone:

- Backbone 3 (BB3)
- Fusion-Site A-E
- Nourseothrycin resistence (Nat. res.)

5.1. Nucleic Acid Sequence:

CGCTGGAGCATGGACATGGAGGCCCAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGC
TCAGGGGCATGATGTGACTGTCGCCCGTACATTTAGCCCATACATCCCCATGTATAATCA
TTTGCATCCATACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTACGGCTCCTCGCTG
CAGACCTGCGAGCAGGGAAACGCTCCCCTCACAGACGCGTTGAATTGTCCCCACGCCGCG
CCCCTGTAGAGAAATATAAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTTCATATACTT
CCTTTTAAAATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAACAACCATGG
GTACCACTCTTGACGACACGGCTTACCGGTACCGCACCAGTGTCCCGGGGGACGCCGAGG
CCATCGAGGCACTGGATGGGTCCTTCACCACCGACACCGTTTTTCGCGTCACCGCCACCG
GGGACGGCTTCACCCTGCGGGAGGTGCCGGTGGACCCGCCCCTGACCAAGGTGTTCCCCG
ACGACGAATCGGACGACGAATCGGACGACGGGGAGGACGGCGACCCGGACTCCCGGACGT
TCGTCGCGTACGGGGACGACGGCGACCTGGCGGGCTTCGTGGTCGTCTCGTACTCCGGCT
GGAACCGCCGGCTGACCGTCGAGGACATCGAGGTCGCCCCGGAGCACCGGGGGCACGGGG
TCGGGCGCGCGTTGATGGGGCTCGCGACGGAGTTCGCCCGCGAGCGGGGCGCCGGGCACC
TCTGGCTGGAGGTCACCAACGTCAACGCACCGGCGATCCACGCGTACCGGCGGATGGGGT
TCACCCTCTGCGGCCTGGACACCGCCCTGTACGACGGCACCGCCTCGGACGGCGAGCAGG
CGCTCTACATGAGCATGCCCTGCCCCTAATCAGTACTGACAATAAAAAGATTCTTGTTTT
CAAGAACTTGTCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCAAATGT
TAGCGTGATTTATATTTTTTTTCGCCTCGACATCATCTGCCCAGATGCGAAGTTAAGTGC
GCAGAAAGTAATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTGGCTTCATA
CCCAGCATTGACCTTTGGTATGAGCATCTGAAAAACAACCAGGTGTTGCAAAGTTAAACA
TCCTTCTTTGTTCATATAGAACCCACTATTCATGGTACTCCCCAATCGAATTTCACATTC
TGGTTTTGAAATTACACACCACGTTAGCTTATAAGATTTCATATAACTTATTGATATACG
GTTTCCATTGTTCGAATAGTTGAGGTTGTATGTAATTCGATTGAAGGGGCCATTTTTGTT
TCCTACTTTTCCTGGGAGCTTATCCGATGCGCTTCAAAGCTGGAATTGTAAATATAGAGA

AAAAGAAGGATGTTGTTTTATTCTTGAAAGAGTATAATTTTACTTCTAGCAACTCTCCCA
CTTCGCTTGACTTCATTTATTTCTTGGGCACATAGGCGTAGTAATCTAGACCAACAGATA
ATTTGCCGGAATGATATAGCGATTGGAAAATGAACTGAAATTTTTTGCTGTCTTTCAATT
TGACGGGCAGTTCATCAGGCGCGCCGTGACCGACCATATAAATACGTTGAGAATGTTATT
CTTCCTCGTAGTTGAAGTGGCTTCATAATTTCAGAACTCAATAGATAAACTAGGATGTTT
TAAAGCAATTAATGCTCACAAGTAAGGAGCGACTCTCTTGCTTTTCGAATACTAAAAGTA
TCGTCCCAACCCAGAAAAAAAGACCTCTTAACTGCAAAATAAACTCTATATATTTCTTCT
AAAACAGTTTCAGGTTGGATAGTATCGCATTCTCATCACTTCTAACTAGTAGGCCATGAG
ATATATTAACGTTTACTTGAGTTCTAAGTTCTCCGAATTAGATGCACAGCACAAACAAGA
TTAGGTTTCACTTGGTACAAAATACGAACAGAGTTTAAGGTCGTAATTTCATTTCGTTAT
TGATCCCCACAATCTATTCTTATCACAGTCATCAGATAGTCGCGAAAAAGCATGCAGAAA
AGGGGGTCGTCCCTATCTAAGTTGTAGCATTACAACAAATATGACCGCTGTGAGCAAAAG
GCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCC
GCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAG
GACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGA
CCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTC
ATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG
TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGT
CCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA
GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACA
CTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAG
TTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCA
AGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGG
GGTGGAGCGCAAAAAACCCCGCCCCTGACAGGGCGGGGTTTTTTCGCGATCGG

6. Complete Sequence

6.1. Amino Acid Sequence

AGGATCCTTTTTTGTAGAAATGTCTTGGTGTCCTCGTCCAATCAGGTAGCCATCTCTGAA
ATATCTGGCTCCGTTGCAACTCCGAACGACCTGCTGGCAACGTAAAATTCTCCGGGGTAA
AACTTAAATGTGGAGTAATGGAACCAGAAACGTCTCTTCCCTTCTCTCTCCTTCCACCGC
CCGTTACCGTCCCTAGGAAATTTTACTCTGCTGGAGAGCTTCTTCTACGGCCCCCTTGCA
GCAATGCTCTTCCCAGCATTACGTTGCGGGTAAAACGGAGGTCGTGTACCCGACCTAGCA
GCCCAGGGATGGAAAAGTCCCGGCCGTCGCTGGCAATAATAGCGGGCGGACGCATGTCAT
GAGATTATTGGAAACCACCAGAATCGAATATAAAAGGCGAACACCTTTCCCAATTTTGGT
TTCTCCTGACCCAAAGACTTTAAATTTAATTTATTTGTCCCTATTTCAATCAATTGAACA
ACTATCAAAACACCATGAGATTCCCATCTATTTTCACCGCTGTCTTGTTCGCTGCCTCCT
CTGCATTGGCTGCCCCTGTTAACACTACCACTGAAGATGAGACTGCTCAAATTCCAGCTG
AAGCAGTTATCGGTTACTCTGACCTTGAGGGTGATTTCGACGTCGCTGTTTTGCCTTTCT
CTAACTCCACTAACAACGGTTTGTTGTTCATTAACACCACTATCGCTTCCATTGCTGCTA
AGGAAGAGGGTGTCTCTCTCGAGAAGAGAGGCGCTCCCCCAGGCCCACCCGGACCCCCCG
GACCACCTGGCCCCCCTGGTCCACCTGGCCCACCTGGCCCACCTGGACCACCAGGACCAC
CAGGTCCCCCCGGCCCTCCCGGTCCTCCCGGACCCCCTGGACCCCCTGGACCCGCCGGCG
CAAGTGGTCCAGGACAGCAGGGTAGGGGAGGTTACGGTCCCGGTCAACAAGGTCCATCTG
GTCCTGGTTCCGGTGCTGCAGCAGCTGCCGCTGGAGGTGCTGGTCAGGGCGGATATGGTG
GTCTTGGCGGTCAAGGTGCTGGAAGAGGTGGATTAGGAGGACAGGCCGCAGCCGCTGCTG

CCTCTGGTCCCGGAGGATATGGACCTGGACAGCAAGGTCCAGGCGGATATGGTCCCGGTC
AGCAAGGCCCCTCAGGACCCGGTTCTATGGCTGCAGCCGCTGCTGCAGCTATGGGACCCG
GTGGTTACGGCCCTGGACAGCAAGGACCTGGCGGATACGGTTCTCCCGCTGGCGCAAGTG
GTCCAGGACAGCAGGGTAGGGGAGGTTACGGTCCCGGTCAACAAGGTCCATCTGGTCCTG
GTTCCGGTGCTGCAGCAGCTGCCGCTGGAGGTGCTGGTCAGGGCGGATATGGTGGTCTTG
GCGGTCAAGGTGCTGGAAGAGGTGGATTAGGAGGACAGGCCGCAGCCGCTGCTGCCTCTG
GTCCCGGAGGATATGGACCTGGACAGCAAGGTCCAGGCGGATATGGTCCCGGTCAGCAAG
GCCCCTCAGGACCCGGTTCTATGGCTGCAGCCGCTGCTGCAGCTATGGGACCCGGTGGTT
ACGGCCCTGGACAGCAAGGACCTGGCGGATACGGTTCTCCCGCTGGCGCTCCCCCAGGCC
CACCCGGACCCCCCGGACCACCTGGCCCCCCTGGTCCACCTGGCCCACCTGGCCCACCTG
GACCACCAGGACCACCAGGTCCCCCCGGCCCTCCCGGTCCTCCCGGACCCCCTGGACCCC
CTGGACCCGCCGGTGGATGATAAGCTTCTATGTAACTAACGAAACAGCATGTACTAATAG
AACCGTATCGAGAATATTTATTTAGGTGAGTAGTAGGAGTGAACCAGACAGTCAATTTAG
TGAGCTGTCCCAGCTTTTGTGCATTCCAGAATTGCCGGTCAAATTGGTTATGGGTTATGG
GGCTTTTCCGATTGAGGTTCAGTTTCTGCGGTTATCTCTTTCTTGACCTGGTCTTTTACA
GGCTGTTCTTTCTCCCCATGATTATTCTTTAGCTGAAGATACCGCTTAGCCTGATAATGT
CGTCGTTTTGTAATCAAAATCTTTAGTTGGGCATCGTCTGAGGTTTCCTTTGGCTTCTGG
GGTTGTTAGTAGGAACGTAGGAACCATAGTAACTTTTACACATACATTCTTATGATTGCG
AAGTAAGCTGAGTCTGCTGCTTGGCTCCCGAAGTACTTTCTCTTTCTCTACCGGTTGATT
CTCCTTCTGGTGCTCCTAAACGATTGTGTTAGAAGGGATTGACCGCTCCGGGGAGCGGTA
CTACCATTGACACTACACATCCTTAATTCCAATCCTGTCTGGCCTCCTTCACCTTTTAAC
CATCTTGCCCATTCCAACTCGTGTCAGATTGCGTATCAAGTGAAAAAAAAAAAATTTTAA
ATCTTTAACCCAATCAGGTAATAACTGTCGCCTCTTTTATCTGCCGCACTGCATGAGGTG
TCCCCTTAGTGGGAAAGAGTACTGAGCCAACCCTGGAGGACAGCAAGGGAAAAATACCTA
CAACTTGCTTCATAATGGTCGTAAAAACAATCCTTGTCGGATATAAGTGTTGTAGACTGT
CCCTTATCCTCTGCGATGTTCTTCCTCTCAAAGTTTGCGATTTCTCTCTATCAGAATTGC
CATCAAGAGACTCAGGACTAATTTCGCAGTCCCACACGCACTCGTACATGATTGGCTGAA
ATTTCCCTAAAGAATTTCTTTTTCACGAAAATTTTTTTTTTACACAAGATTTTCAGCAGA
TATAAAATGGAGAGCAGGACCTCCGCTGTGACTCTTCTTTTTTTTCTTTTATTCTCACTA
CATACATTTTAGTTATTCGCCAACATGAGATTCCCATCTATTTTCACCGCTGTCTTGTTC
GCTGCCTCCTCTGCATTGGCTGCCCCTGTTAACACTACCACTGAAGATGAGACTGCTCAA
ATTCCAGCTGAAGCAGTTATCGGTTACTCTGACCTTGAGGGTGATTTCGACGTCGCTGTT
TTGCCTTTCTCTAACTCCACTAACAACGGTTTGTTGTTCATTAACACCACTATCGCTTCC
ATTGCTGCTAAGGAAGAGGGTGTCTCTCTCGAGAAGAGAGGCGCTCCCCCAGGCCCACCC
GGACCCCCCGGACCACCTGGCCCCCCTGGTCCACCTGGCCCACCTGGCCCACCTGGACCA
CCAGGACCACCAGGTCCCCCCGGCCCTCCCGGTCCTCCCGGACCCCCTGGACCCCCTGGA
CCCGCCGGCGCAAGTGGTCCAGGACAGCAGGGTAGGGGAGGTTACGGTCCCGGTCAACAA
GGTCCATCTGGTCCTGGTTCCGGTGCTGCAGCAGCTGCCGCTGGAGGTGCTGGTCAGGGC
GGATATGGTGGTCTTGGCGGTCAAGGTGCTGGAAGAGGTGGATTAGGAGGACAGGCCGCA
GCCGCTGCTGCCTCTGGTCCCGGAGGATATGGACCTGGACAGCAAGGTCCAGGCGGATAT
GGTCCCGGTCAGCAAGGCCCCTCAGGACCCGGTTCTATGGCTGCAGCCGCTGCTGCAGCT
ATGGGACCCGGTGGTTACGGCCCTGGACAGCAAGGACCTGGCGGATACGGTTCTCCCGCT
GGCGCAAGTGGTCCAGGACAGCAGGGTAGGGGAGGTTACGGTCCCGGTCAACAAGGTCCA
TCTGGTCCTGGTTCCGGTGCTGCAGCAGCTGCCGCTGGAGGTGCTGGTCAGGGCGGATAT
GGTGGTCTTGGCGGTCAAGGTGCTGGAAGAGGTGGATTAGGAGGACAGGCCGCAGCCGCT
GCTGCCTCTGGTCCCGGAGGATATGGACCTGGACAGCAAGGTCCAGGCGGATATGGTCCC
GGTCAGCAAGGCCCCTCAGGACCCGGTTCTATGGCTGCAGCCGCTGCTGCAGCTATGGGA

CCCGGTGGTTACGGCCCTGGACAGCAAGGACCTGGCGGATACGGTTCTCCCGCTGGCGCT
CCCCCAGGCCCACCCGGACCCCCCGGACCACCTGGCCCCCCTGGTCCACCTGGCCCACCT
GGCCCACCTGGACCACCAGGACCACCAGGTCCCCCCGGCCCTCCCGGTCCTCCCGGACCC
CCTGGACCCCCTGGACCCGCCGGTGGATGATAAGCTTGGATGATACTTTAATTGATGCAT
GTGAAGAATTGAAGAAACAGTAATCTAGTAAAATATAAGAAAGTCCTGAAAGTAGTTCAT
TTAGTAAGCTATGAACCCAAGTTGGGGTAACCGATTTTTCAGGTAAACAGCTGTAGTGAG
GAGACATCTCAAATCGCCTGAGAAGAAAAATATCTATGAGACAACTACAATGCGAAGAAT
GTGTTCAAATGCTCTATTATTCACTAGTTTACCTATTCAATGACCAACTCCTGGGAGGCT
TCGTTACCGATAATTGGTAACTTAACTAGTACGTCCTTCATCAACTCGTAGACAGTGAAA
GTGACGGCTTGTCCCGGAGCAACACGCATGATACGCGGCGTGATACCTTTGTACAGAGCA
CGCATACCTTCCTCTTGGATCAATCTCTTGGAAATACGAAGAATACGGGCCACACCAGAC
TCATTGGTGACACCAGTCTCTCTCTGCATACGAGTCTTGATGGTATCTAGAGGAGCCGCT
AATTGGAGGATCCTTTTTTGTAGAAATGTCTTGGTGTCCTCGTCCAATCAGGTAGCCATC
TCTGAAATATCTGGCTCCGTTGCAACTCCGAACGACCTGCTGGCAACGTAAAATTCTCCG
GGGTAAAACTTAAATGTGGAGTAATGGAACCAGAAACGTCTCTTCCCTTCTCTCTCCTTC
CACCGCCCGTTACCGTCCCTAGGAAATTTTACTCTGCTGGAGAGCTTCTTCTACGGCCCC
CTTGCAGCAATGCTCTTCCCAGCATTACGTTGCGGGTAAAACGGAGGTCGTGTACCCGAC
CTAGCAGCCCAGGGATGGAAAAGTCCCGGCCGTCGCTGGCAATAATAGCGGGCGGACGCA
TGTCATGAGATTATTGGAAACCACCAGAATCGAATATAAAAGGCGAACACCTTTCCCAAT
TTTGGTTTCTCCTGACCCAAAGACTTTAAATTTAATTTATTTGTCCCTATTTCAATCAAT
TGAACAACTATCAAAACACCATGAGATTCCCATCTATTTTCACCGCTGTCTTGTTCGCTG
CCTCCTCTGCATTGGCTGCCCCTGTTAACACTACCACTGAAGATGAGACTGCTCAAATTC
CAGCTGAAGCAGTTATCGGTTACTCTGACCTTGAGGGTGATTTCGACGTCGCTGTTTTGC
CTTTCTCTAACTCCACTAACAACGGTTTGTTGTTCATTAACACCACTATCGCTTCCATTG
CTGCTAAGGAAGAGGGTGTCTCTCTCGAGAAGAGAGGCGCTCCCCCAGGCCCACCCGGAC
CCCCCGGACCACCTGGCCCCCCTGGTCCACCTGGCCCACCTGGCCCACCTGGACCACCAG
GACCACCAGGTCCCCCCGGCCCTCCCGGTCCTCCCGGACCCCCTGGACCCCCTGGACCCG
CCGGCGCAAGTGGTCCAGGACAGCAGGGTAGGGGAGGTTACGGTCCCGGTCAACAAGGTC
CATCTGGTCCTGGTTCCGGTGCTGCAGCAGCTGCCGCTGGAGGTGCTGGTCAGGGCGGAT
ATGGTGGTCTTGGCGGTCAAGGTGCTGGAAGAGGTGGATTAGGAGGACAGGCCGCAGCCG
CTGCTGCCTCTGGTCCCGGAGGATATGGACCTGGACAGCAAGGTCCAGGCGGATATGGTC
CCGGTCAGCAAGGCCCCTCAGGACCCGGTTCTATGGCTGCAGCCGCTGCTGCAGCTATGG
GACCCGGTGGTTACGGCCCTGGACAGCAAGGACCTGGCGGATACGGTTCTCCCGCTGGCG
CAAGTGGTCCAGGACAGCAGGGTAGGGGAGGTTACGGTCCCGGTCAACAAGGTCCATCTG
GTCCTGGTTCCGGTGCTGCAGCAGCTGCCGCTGGAGGTGCTGGTCAGGGCGGATATGGTG
GTCTTGGCGGTCAAGGTGCTGGAAGAGGTGGATTAGGAGGACAGGCCGCAGCCGCTGCTG
CCTCTGGTCCCGGAGGATATGGACCTGGACAGCAAGGTCCAGGCGGATATGGTCCCGGTC
AGCAAGGCCCCTCAGGACCCGGTTCTATGGCTGCAGCCGCTGCTGCAGCTATGGGACCCG
GTGGTTACGGCCCTGGACAGCAAGGACCTGGCGGATACGGTTCTCCCGCTGGCGCTCCCC
CAGGCCCACCCGGACCCCCCGGACCACCTGGCCCCCCTGGTCCACCTGGCCCACCTGGCC
CACCTGGACCACCAGGACCACCAGGTCCCCCCGGCCCTCCCGGTCCTCCCGGACCCCCTG
GACCCCCTGGACCCGCCGGTGGATGATAAGCTTTTAGTGTACATCTGATAATATAGTACT
ACCACGTATGATAATGTAGAGAATAGGCTTCCTTGTCGAGTGTGTTTGCAGTTTTCTTGA
GTTTCAAGGTTTAAATGCTGGTATATTAGTTCATCGAAGGTTTCAGCCAATAGCACCTTA
AATCAATCAAACTAATTCGACTCTTACGAAAGAGCCTACTGTGTTTAGTATCGAAGTCGT
TTACCTTTCATGTTGAATAGCTTCCTCTCTGACCCTAACATTTCAAGATCCTCCTAAAGT
TACCCGGATTGTGAAATTCTAATGATCCACCTGCCCAATGCATTTTTTCTTTATTCAGTT

TACCTTTTTTACCTAATATACGAGCTTGTTAAAGTAAGTGGCACTGCAATACTAGGCTTA
TTGTTGATATTATGATGAATCGTTTTCACAAACTTGATTTCCTGTGAACTCACCATGTAC
TAAGGAAAAAAACATGCATCACCATCTGAATATTTGACCGCTAGCTGGAGCGGTACTACC
ATTGACACTACACATCCTTAATTCCAATCCTGTCTGGCCTCCTTCACCTTTTAACCATCT
TGCCCATTCCAACTCGTGTCAGATTGCGTATCAAGTGAAAAAAAAAAAATTTTAAATCTT
TAACCCAATCAGGTAATAACTGTCGCCTCTTTTATCTGCCGCACTGCATGAGGTGTCCCC
TTAGTGGGAAAGAGTACTGAGCCAACCCTGGAGGACAGCAAGGGAAAAATACCTACAACT
TGCTTCATAATGGTCGTAAAAACAATCCTTGTCGGATATAAGTGTTGTAGACTGTCCCTT
ATCCTCTGCGATGTTCTTCCTCTCAAAGTTTGCGATTTCTCTCTATCAGAATTGCCATCA
AGAGACTCAGGACTAATTTCGCAGTCCCACACGCACTCGTACATGATTGGCTGAAATTTC
CCTAAAGAATTTCTTTTTCACGAAAATTTTTTTTTTACACAAGATTTTCAGCAGATATAA
AATGGAGAGCAGGACCTCCGCTGTGACTCTTCTTTTTTTTCTTTTATTCTCACTACATAC
ATTTTAGTTATTCGCCAACATGAGATTCCCATCTATTTTCACCGCTGTCTTGTTCGCTGC
CTCCTCTGCATTGGCTGCCCCTGTTAACACTACCACTGAAGATGAGACTGCTCAAATTCC
AGCTGAAGCAGTTATCGGTTACTCTGACCTTGAGGGTGATTTCGACGTCGCTGTTTTGCC
TTTCTCTAACTCCACTAACAACGGTTTGTTGTTCATTAACACCACTATCGCTTCCATTGC
TGCTAAGGAAGAGGGTGTCTCTCTCGAGAAGAGAGGCGCTCCCCCAGGCCCACCCGGACC
CCCCGGACCACCTGGCCCCCCTGGTCCACCTGGCCCACCTGGCCCACCTGGACCACCAGG
ACCACCAGGTCCCCCCGGCCCTCCCGGTCCTCCCGGACCCCCTGGACCCCCTGGACCCGC
CGGCGCAAGTGGTCCAGGACAGCAGGGTAGGGGAGGTTACGGTCCCGGTCAACAAGGTCC
ATCTGGTCCTGGTTCCGGTGCTGCAGCAGCTGCCGCTGGAGGTGCTGGTCAGGGCGGATA
TGGTGGTCTTGGCGGTCAAGGTGCTGGAAGAGGTGGATTAGGAGGACAGGCCGCAGCCGC
TGCTGCCTCTGGTCCCGGAGGATATGGACCTGGACAGCAAGGTCCAGGCGGATATGGTCC
CGGTCAGCAAGGCCCCTCAGGACCCGGTTCTATGGCTGCAGCCGCTGCTGCAGCTATGGG
ACCCGGTGGTTACGGCCCTGGACAGCAAGGACCTGGCGGATACGGTTCTCCCGCTGGCGC
AAGTGGTCCAGGACAGCAGGGTAGGGGAGGTTACGGTCCCGGTCAACAAGGTCCATCTGG
TCCTGGTTCCGGTGCTGCAGCAGCTGCCGCTGGAGGTGCTGGTCAGGGCGGATATGGTGG
TCTTGGCGGTCAAGGTGCTGGAAGAGGTGGATTAGGAGGACAGGCCGCAGCCGCTGCTGC
CTCTGGTCCCGGAGGATATGGACCTGGACAGCAAGGTCCAGGCGGATATGGTCCCGGTCA
GCAAGGCCCCTCAGGACCCGGTTCTATGGCTGCAGCCGCTGCTGCAGCTATGGGACCCGG
TGGTTACGGCCCTGGACAGCAAGGACCTGGCGGATACGGTTCTCCCGCTGGCGCTCCCCC
AGGCCCACCCGGACCCCCCGGACCACCTGGCCCCCCTGGTCCACCTGGCCCACCTGGCCC
ACCTGGACCACCAGGACCACCAGGTCCCCCCGGCCCTCCCGGTCCTCCCGGACCCCCTGG
ACCCCCTGGACCCGCCGGTGGATGATAAGCTTGGATGATACTTTAATTGATGCATGTGAA
GAATTGAAGAAACAGTAATCTAGTAAAATATAAGAAAGTCCTGAAAGTAGTTCATTTAGT
AAGCTATGAACCCAAGTTGGGGTAACCGATTTTTCAGGTAAACAGCTGTAGTGAGGAGAC
ATCTCAAATCGCCTGAGAAGAAAAATATCTATGAGACAACTACAATGCGAAGAATGTGTT
CAAATGCTCTATTATTCACTAGTTTACCTATTCAATGACCAACTCCTGGGAGGCTTCGTT
ACCGATAATTGGTAACTTAACTAGTACGTCCTTCATCAACTCGTAGACAGTGAAAGTGAC
GGCTTGTCCCGGAGCAACACGCATGATACGCGGCGTGATACCTTTGTACAGAGCACGCAT
ACCTTCCTCTTGGATCAATCTCTTGGAAATACGAAGAATACGGGCCACACCAGACTCATT
GGTGACACCAGTCTCTCTCTGCATACGAGTCTTGATGGTATCTAGAGGAGCCGCTGGAGC
ATGGACATGGAGGCCCAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGGGCA
TGATGTGACTGTCGCCCGTACATTTAGCCCATACATCCCCATGTATAATCATTTGCATCC
ATACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTACGGCTCCTCGCTGCAGACCTGC
GAGCAGGGAAACGCTCCCCTCACAGACGCGTTGAATTGTCCCCACGCCGCGCCCCTGTAG
AGAAATATAAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTTCATATACTTCCTTTTAAA

ATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAACAACCATGGGTACCACTC
TTGACGACACGGCTTACCGGTACCGCACCAGTGTCCCGGGGGACGCCGAGGCCATCGAGG
CACTGGATGGGTCCTTCACCACCGACACCGTTTTTCGCGTCACCGCCACCGGGGACGGCT
TCACCCTGCGGGAGGTGCCGGTGGACCCGCCCCTGACCAAGGTGTTCCCCGACGACGAAT
CGGACGACGAATCGGACGACGGGGAGGACGGCGACCCGGACTCCCGGACGTTCGTCGCGT
ACGGGGACGACGGCGACCTGGCGGGCTTCGTGGTCGTCTCGTACTCCGGCTGGAACCGCC
GGCTGACCGTCGAGGACATCGAGGTCGCCCCGGAGCACCGGGGGCACGGGGTCGGGCGCG
CGTTGATGGGGCTCGCGACGGAGTTCGCCCGCGAGCGGGGCGCCGGGCACCTCTGGCTGG
AGGTCACCAACGTCAACGCACCGGCGATCCACGCGTACCGGCGGATGGGGTTCACCCTCT
GCGGCCTGGACACCGCCCTGTACGACGGCACCGCCTCGGACGGCGAGCAGGCGCTCTACA
TGAGCATGCCCTGCCCCTAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAACTT
GTCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCAAATGTTAGCGTGAT
TTATATTTTTTTTCGCCTCGACATCATCTGCCCAGATGCGAAGTTAAGTGCGCAGAAAGT
AATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTGGCTTCATACCCAGCATT
GACCTTTGGTATGAGCATCTGAAAAACAACCAGGTGTTGCAAAGTTAAACATCCTTCTTT
GTTCATATAGAACCCACTATTCATGGTACTCCCCAATCGAATTTCACATTCTGGTTTTGA
AATTACACACCACGTTAGCTTATAAGATTTCATATAACTTATTGATATACGGTTTCCATT
GTTCGAATAGTTGAGGTTGTATGTAATTCGATTGAAGGGGCCATTTTTGTTTCCTACTTT
TCCTGGGAGCTTATCCGATGCGCTTCAAAGCTGGAATTGTAAATATAGAGAAAAAGAAGG
ATGTTGTTTTATTCTTGAAAGAGTATAATTTTACTTCTAGCAACTCTCCCACTTCGCTTG
ACTTCATTTATTTCTTGGGCACATAGGCGTAGTAATCTAGACCAACAGATAATTTGCCGG
AATGATATAGCGATTGGAAAATGAACTGAAATTTTTTGCTGTCTTTCAATTTGACGGGCA
GTTCATCAGGCGCGCCGTGACCGACCATATAAATACGTTGAGAATGTTATTCTTCCTCGT
AGTTGAAGTGGCTTCATAATTTCAGAACTCAATAGATAAACTAGGATGTTTTAAAGCAAT
TAATGCTCACAAGTAAGGAGCGACTCTCTTGCTTTTCGAATACTAAAAGTATCGTCCCAA
CCCAGAAAAAAAGACCTCTTAACTGCAAAATAAACTCTATATATTTCTTCTAAAACAGTT
TCAGGTTGGATAGTATCGCATTCTCATCACTTCTAACTAGTAGGCCATGAGATATATTAA
CGTTTACTTGAGTTCTAAGTTCTCCGAATTAGATGCACAGCACAAACAAGATTAGGTTTC
ACTTGGTACAAAATACGAACAGAGTTTAAGGTCGTAATTTCATTTCGTTATTGATCCCCA
CAATCTATTCTTATCACAGTCATCAGATAGTCGCGAAAAAGCATGCAGAAAAGGGGGTCG
TCCCTATCTAAGTTGTAGCATTACAACAAATATGACCGCTGTGAGCAAAAGGCCAGCAAA
AGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTG
ACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA
GATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGC
TTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAC
GCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAAC
CCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGG
TAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGT
ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAA
CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCT
CTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGA
TTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTGGAGCG
CAAAAAACCCCGCCCCTGACAGGGCGGGGTTTTTTCGCGATCGG

Methods

1. Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) is commonly used to amplify a certain DNA fragment. In our project we mainly used the PCR for the creation of our inserts.

In a first step the reaction mix was prepared in PCR-tubes. Every sample also needed a negative control, entailing the same ingredients but ddH_2O instead of the template DNA.

1.1.Reaction Mix (for 1 reaction):

- 10µL 5x Q5 Reaction Buffer
- 1µL 10mM dNTPs
- 10µL 5x Q5 High GC Enhancer (optional)
- 0,5µL Q5 High-Fidelity DNA Polymerase
- 1µL 10µM Primer Solution (1µL of each used primer was added, at least one forward and one backward primer are needed for a successful PCR reaction)
- 1µL template DNA 1pg/µL-1ng/µL (1µL of each used template was added, some reactions, only consisting of matching primers, did not need a template DNA)
- ddH_2O to a total volume of 50µL

The PCR-tubes were transferred to a thermocycler to go through the following temperature program.

1.2.Q5 Program:

- 30sec 98°C (initial denaturation)
- 30x:
    - 10sec 98°C (denaturation)
    - 40sec 72°C (primer annealing and extension)
- 2min 72°C (final extension)

2. Gel electrophoresis (Agarose gels)

In our project two different types of gel electrophoresis, using Agarose gels, were used. The first method was used to clean up the obtained PCR-products using gels with big pockets, able to fit all 50µL of the PCR-product. The second use of the gel electrophoresis, with Agarose gels resulted in an analytical gel, run after a control digest. Since no product needed to be gained out of these gels, smaller pockets only containing 10µL of the digested sample.

2.1. Recipe for Agarose gels (300mL=2gels)

· 3g Agarose

· fill with 1x TAE buffer to 300mL

· add 30µL Sybrsafe

3. PCR/Gel clean up

For the clean-up of PCR-products and samples of the gel-electrophoresis with agarose gels, the Wizard SV Gel and PCR Clean-Up System by Promega was used.

3.1. Gel Slice and PCR Product Preparation

· Following electrophoresis, excise DNA band from gel and place gel slice in a 1.5ml microcentrifuge tube.

· Add 10µl Membrane Binding Solution per 10mg of gel slice. Vortex and incubate at 50–65°C until gel slice is completely dissolved. Add an equal volume of Membrane Binding Solution to the PCR amplification.

3.2. Binding of DNA

· Insert SV Minicolumn into Collection Tube.

· Transfer dissolved gel mixture or prepared PCR product to the Minicolumn assembly. Incubate at room temperature for 1 minute.

· Centrifuge at 16,000 × g for 1 minute. Discard flowthrough and reinsert Minicolumn into Collection Tube.

3.3. Washing

· Add 700µl Membrane Wash Solution (ethanol added). Centrifuge at 16,000 × g for 1 minute. Discard flowthrough and reinsert Minicolumn into Collection Tube.

- Repeat Step 6 with 500µl Membrane Wash Solution. Centrifuge at 16,000 × g for 5 minutes.
- Empty the Collection Tube and recentrifuge the column assembly for 1 minute with the microcentrifuge lid open (or off) to allow evaporation of any residual ethanol.

3.4. Elution

- Carefully transfer Minicolumn to a clean 1.5ml microcentrifuge tube.
- Add 50µl of Nuclease-Free Water to the Minicolumn. Incubate at room temperature for 1 minute. Centrifuge at 16,000 × g for 1 minute.
- Discard Minicolumn and store DNA at 4°C or –20°C.

4. Concentration determination (NanoDrop)

- Switch on the NanoDrop computer. Select the NanoDrop program.
- Select "nucleic acid" the NanoDrop should now measure at a wavelength of 230nm.
- Clean the NanoDrop by pipetting 1µL MQ-H2O on the NanoDrop and clean it with a cleaning cloth (Always clean both sides of the NanoDrop).
- Put 1µL of 10mM Tris pH=7,5 nuclease free on the NanoDrop, close it and select "blanc" (if the DNA is diluted in a different solution use this solution for the blanc).
- Clean the NanoDrop with a cleaning cloth and pipet 1µL sample on the NanoDrop.
- Select "measure".
- Write down the obtained concentration
- Clean the NanoDrop like in step 3. Put a cleaning cloth between the two sides of the NanoDrop.
- Store the samples at -20°C.

5. Sequencing (Microsynth)

All samples used in this project were sequenced by the company “Microsynth”.

- Prepare for each sample two 1,5mL reaction tubes with a screw lid.
- Put 3µL 10µM forward primer in the first tube and 3µL 10µM backward primer in the second tube.
- Add 12µL sample to both reaction tubes, determine the concentration of the sample according to the table below.

Plasmid	40-100 ng/µL
PCR 1000-2000bp	30 ng/µL
PCR 500-1000bp	15 ng/µL
PCR 250-500bp	7,5 ng/µL
PCR 100-250bp	4 ng/µL

- Label the reaction tubes with a Microsynth label. Write the contents of the reaction tubes and the accompanying label numbers in your lab journal.
- Send the samples for sequencing.
- Fill out an order form on the website of the company „Microsynth“:
- Copy the expected sequence to the results and align the gained sequence to it.

6. Golden Gate

Golden Gate is the cloning methods used by our team for the preparation of our various plasmids. Using several steps, always including a new plasmid backbone allowed us to easily obtain different versions of our plasmids e.g. to test which of our two primers leads to the higher expression rates. For the different steps of the Golden Gate, different restriction enzymes were used to create the desired sticky ends. For this project the restriction enzyme BsaI was used to create plasmids, using the backbone 1. For our backbone 2 versions, the restriction enzyme BbsI was chosen. Since the backbone 2 step can be skipped, if only one expression cassette is needed in the complete plasmid, two different restriction enzymes were

used for the creation of plasmids, using the backbone 3 versions. If the backbone 3 was created directly out of the backbone 1, the restriction enzyme BbsI was used, whereas the restriction enzyme BsaI was taken, if the backbone 3 was built out of backbone 2 versions.

For each sample a reaction, according to the table below, and a negative control are prepared in PCR-tubes.

Golden Gate Reaction/sample	
	Volume [µL]
Insert 1 (40nM)	1,0
Insert 2 (40nM)	1,0
Insert 3 (40nM)	1,0
Further inserts can be added if needed	
Backbone (40nM)	1,0
CutSmart Buffer	2,0
ATP (20 mM)	2,0
Restriction enzyme	1,0
T4 Ligase (1:10)	2,5
dH2O or TRIS	*Depending on the number of inserts, fill to a total of 20µL*
total	20,0

The PCR-tubes were transferred to a thermocycler to go through the following temperature program.

Thermocycler	
5 min	37° C
2.5 min	16° C
45 repeats	
5 min	50° C
10 min	80° C

7. E.coli transformation

- thaw chemically competent E.coli (strain DH10b) on ice.
- pipette 100 µL of the E.coli cells into a 1,5 mL tube.
- add 10 µL of the Golden Gate assembly (always transform the negative control too).
- put the mix 10 minutes on ice.
- transfer the tube into a thermoblock: 42°C for 90 seconds.
- recover cells on ice for 2 minutes.
- add 900 µL LB liquid medium without antibiotics.
- transfer the tube into a thermomixer: 37°C for 1h and 700 rpm.
- pipette 1x 50µL and 1x 10µL of your sample onto a LB-Agar plate containing the desired antibiotic.

- take a spatula and distribute the sample (don't make a smear, you want to get single colonies in the end)
- incubate the plates over night at 37°C
- the next day compare the plate containing the cells with the plasmid with the negative control: you should count way more cells to know that the transformation was successful.

8. Mini-Prep

For this project the HiYield Plasmid Mini DNA Kit from the company “SGL” was used.

Part 1 (day 1)

- Select 12 clones of the 24h incubated E. coli plate from the E. coli-transformation.
- Prepare an eprouvette containing 2mL selective medium (depending on the resistance of the plasmid) for each selected clone.
- Prepare an Agar-Plate with the correct selective agar.
- Label the eprouvettes with the numbers of the clones, divide the Agar-Plate in sections and label them according to the number of clones this will serve as master plate.
- Use sterile tips to take some bacterial material of the selected clones of the E. coli plate, draw a line in the correct section of the master plate and put the tip in the correct eprouvette.
- Incubate for 24h, 37°C, shaking.

Part 2 (day 2)

- Wrap the master plate in parafilm and store it at 4°C
- Transfer 1mL of the E. coli cells of the eprouvette in a 1,5 mL reaction tube.
- Centrifuge: 15,000 x g, 1min.
- Discard the supernatant.
- Repeat Step 8-10 until all the E. coli culture is in the reaction tube.
- Add 200µL PD1 Buffer (make sure the RNAse A was added)

- Resuspend the cells by vortexing or pipetting up and down.
- Add 200µL PD2 Buffer + mix gently by inverting 10 times (do not vortex to avoid shearing the genomic DNA)
- Incubate at room temperature for at least 2min but do not exceed 5min.
- Add 300µL PD3 Buffer + immediately mix gently by inverting 10 times (do not vortex to avoid shearing the genomic DNA)
- Centrifuge: 15,000 x g, 3min
- Place a PD column in a 2mL Collection Tube.
- Add the supernatant of step 17.
- Centrifuge: 15,000 x g, 1min
- Discard the flow-through.
- Place the PD column back in the 2mL Collection Tube.
- Add 400µL W1 Buffer into the PD Column.
- Centrifuge: 15,000 x g, 1min
- Discard the flow-through
- Place the PD column back in the 2mL Collection Tube.
- Add 600µL Wash Buffer into the PD Column.
- Centrifuge: 15,000 x g, 1min
- Discard the flow-through.
- Place the PD column back in the 2mL Collection Tube.
- Centrifuge: 15,000 x g, 3min to dry the column matrix.
- Transfer the column into a 1,5mL reaction tube.
- Add 50µL of Elution Buffer in the centre of the column matrix.
- Incubate 2min at room temperature.
- Centrifuge: 15,000x g, 2min to elute the DNA.
- Measure the DNA concentration via NanoDrop and store at -20°C

9. Digest

In this project the restriction digest was primarily used as a control of the created plasmid as well as the linearization of the plasmids for the successful cloning into the yeast Pichia Pastoris.

For a control digest, a fitting enzyme must be chosen, ideally creating parts with different sizes that can be seen on an analytical gel after electrophoresis.

Digest mix:

· 0,5µL restriction enzyme

· 1µL sample

· 1µL CutSmart buffer (company NEB)

· 7,5µL dH_2O

The digest mix was incubated for 2 hours at the ideal temperature for the restriction enzyme, all enzymes used in this project had an ideal temperature of 37 degrees Celsius and loaded on an analytic gel afterwards.

10. Pichia Pastoris-transformation

· The completed plasmid was linearized, using the restriction enzyme AscI.

· 5-10µg of the linearized plasmid (the volume should not exceed 30µL) are added to 80µL of electrocompetent *P.pastoris*. As a negative control 30µL of dH_2O are used instead of the sample.

· The cell-mixture is then transferred to a chilled electroporation cuvette (2mm) and incubated on ice for 5min.

· The electroporation is performed by the “BioRad Gene Pulser” using 1,5kV for a period of 4ms.

· 1mL of ice-cold liquid YPD-media is added to the electroporated cells in the cuvette immediately after the electroporation. Then the mixture is transferred to a sterile reaction tube and incubated at 28 degrees Celsius for approximately 2 hours.

· 20µL, 200µL and the rest of each sample are plated on three YPD-Agar plates containing the antibiotic nourseothricin and incubated for approximately 72 hours at 28 degrees Celsius until distinctive cultures can be seen.

11. Cultivation of *Pichia pastoris*

Separation of the clones:

- Four clones of each *Pichia pastoris* plate are chosen and plated on a new YPD-Agar plates containing the antibiotic nourseothricin using the fractionated smear method and incubated for 72 hours.
- This procedure is repeated to ensure the purity of the selected clones.

Pre culture:

- 10mL of liquid YPG-Medium containing nourseothricin are inoculated with two purified clones of each sample and incubated at 28 degrees Celsius for 24 hours shaking.
- A wild-type control is also inoculated in YPG-Medium not containing any antibiotics.

Main culture:

- 20mL of BM-Medium containing the antibiotic nourseothricin are inoculated to an OD600 of 0,1, set to a glucose-concentration of 2% and incubated for 48 hours at 28 degrees Celsius shaking.
- After 12, 24, 36 and 48 hours 2mL samples are taking of each culture and the glucose-concentration is set to 1%.
- 1mL of the sample is used for the OD-measurement to control the cell growth, the other mL is centrifuged at 16,000xg for 5min and used for the protein conformation.

12. OD-Measurement

- The photometer is activated and set to the desired wavelength.
- 1mL of each sample is transferred to a plastic cuvette, put in the photometer and measured at the correct wavelength.
- 1mL of the solution the microorganism is cultivated in is used as a blanc for the measurement.

SDS-Page with Coomassie staining:

Sample preparation:

- 9,75µL Sample

- 3,75µL Loading dye
- 1,5µL reducing agent
- Incubate at 95 degrees Celsius for 10min.

SDS-Page:

- 5µL Ladder and 15µL sample are loaded in each well of the 10% bis-tris-gel.
- The gel is run for approximately 30min at 200V.

Coommassie blue staining:

- Stain the gel in coommassie stain solution for at least 4 hours.
- Wash the gel with dH_2O until the bands appear.

EPFL - HESTIA

Design

Thinking Process

Silk proteins demonstrate interesting mechanical properties such as toughness, strength, lightweight, biodegradability and the possibility to produce different morphologies (fibers, foams, capsules, films) (ref). In addition to that, these comprise a high percentage of the amino acids glycine, serine and alanine which have an intermediate hydrophobicity (ref). As we already decided to leverage the dual-insulative and sustainable properties of the aerogel, we were investigating a way to protect it from water since the cellulose aerogel is well-known for its hydrophilicity. We wanted to tackle this issue using synthetic biology and carrying the sustainable aspect of our new material. Therefore, working with silk proteins appeared to us as the ideal solution for protecting the cellulose aerogel from humidity present between walls.

After looking into many different kinds of silks, we were hesitating between Green Lacewing and spider silk. Spider silks are usually used for foraging or as a lifeline. Green Lacewing insects produce two types of silk : one produced by the larvae (cocoon) and the other by adult females (egg-stalk) . The second one has a protective shelter and structural support function in nature, which are two interesting advantages for a waterproof coating for our aerogel (ref). Hence, we decided to work with Green Lacewing silk. This became an opportunity for us to bring even more innovation to our material since this type of silk is not well-known and studied compared to spider silk.

In the species *Mallada signata*, two serine- and glycine-rich proteins (Ma1XB1 and Ma1XB2) have been identified (ref). They both have highly repetitive core domains and small terminal domains. The core domain’s structure is rich in β-sheets with an approximative sheet-length of four amino acids followed by turns. These form repeating structural units constituting β-helices which have a significant positive correlation with the proteins’ surface hydrophobicity (ref). We discovered that a consensus motif for the core domain of Ma1XB2 (named [AS]) was already generated. Besides, a recombinant protein constituted by 8 repetitions of the [AS] module and both natural termini of Ma1XB2 (named N[AS]8C) had also already been expressed in *E.Coli* (ref). As this artificial egg-stalk protein carried similar properties as Ma1XB2, we chose to work with it.

N[AS]8C’s characteristics were further tested in various morphologies such as films, capsules, hydrogels and foams (ref). When $N[AS]_8C$ was forming a film, its hydrophobic properties were enhanced. Therefore, we decided to engineer a unique hydrophobic biofilm that would be able to bind to cellulose to create a protective waterproof coating for our final material.

Plasmid Design (SnapGene)

We designed our plasmids in accordance with our approaches (that you can find below) to create a modular silk biofilm. In case you wish to express our green lacewing silk in a different context, feel free to reuse parts of our plasmids as needed. All of the plasmids presented below have a unique code (01a, 01b, etc.) referring to our approaches.

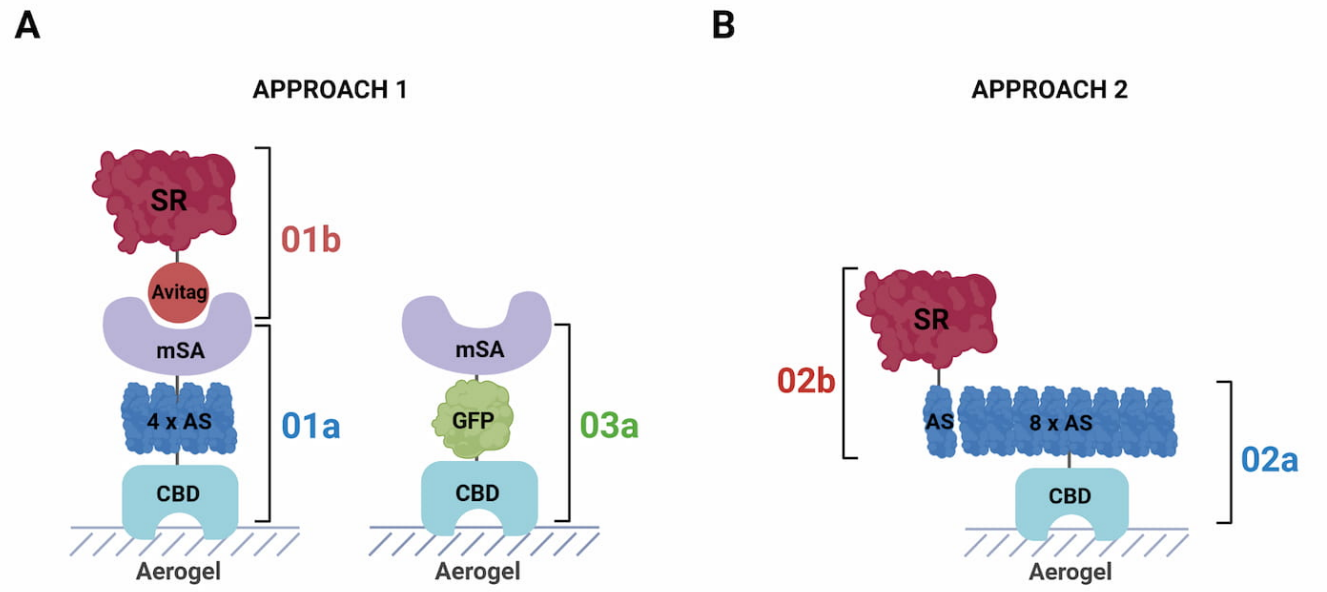


Figure 1: Diagrams describing our approaches to creating a modular silk biofilm. *The [AS] domains are silk modules. The 03a approach serves as a control.*

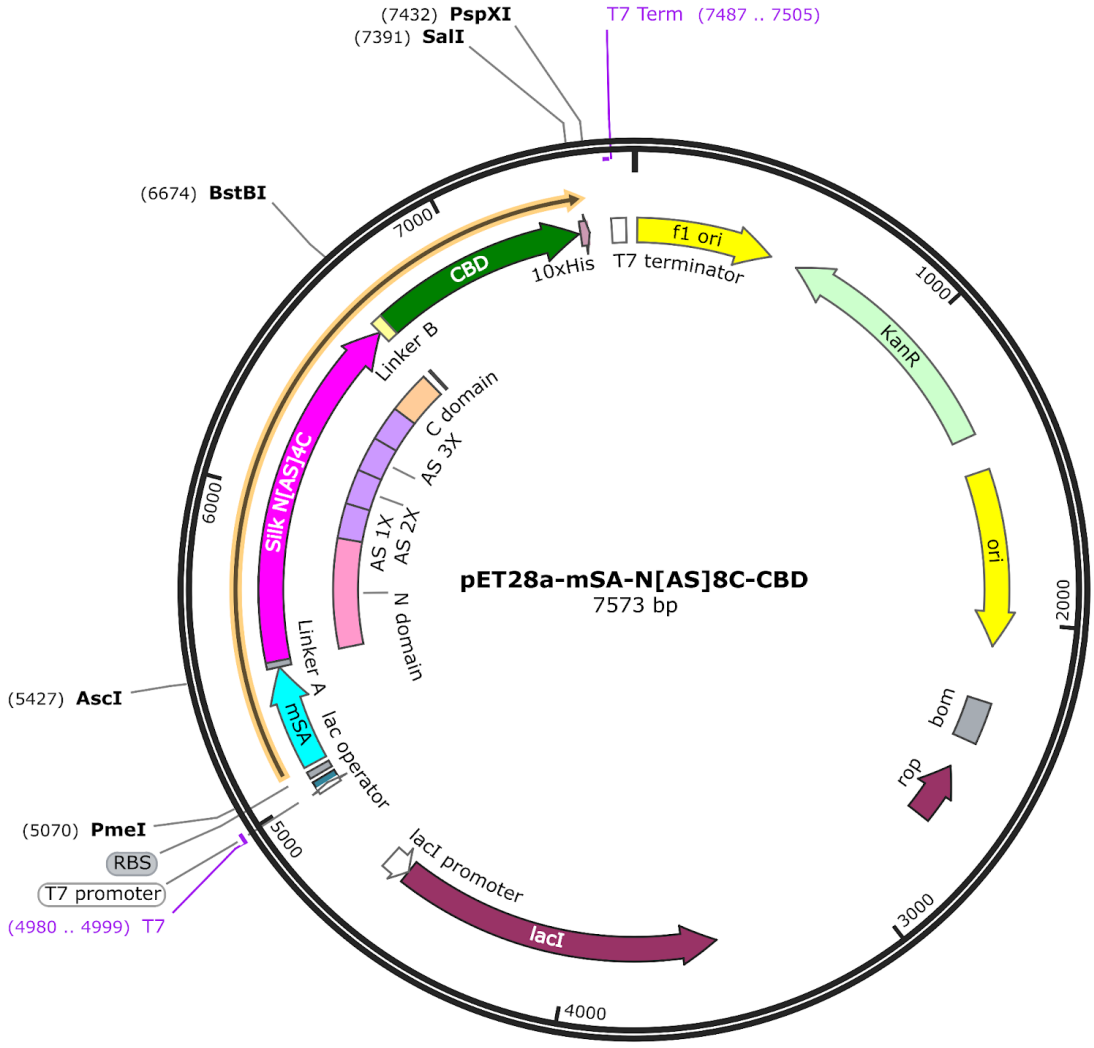


Figure 2: 01a-pET28a-mSA-N[AS]4C-CBD-10His plasmid map. *The 4x[AS] modules serve as our silk. The backbone of this plasmid is pET28a, the plasmids were expressed in E.coli DH5α and in E.coli BL21(DE3).*

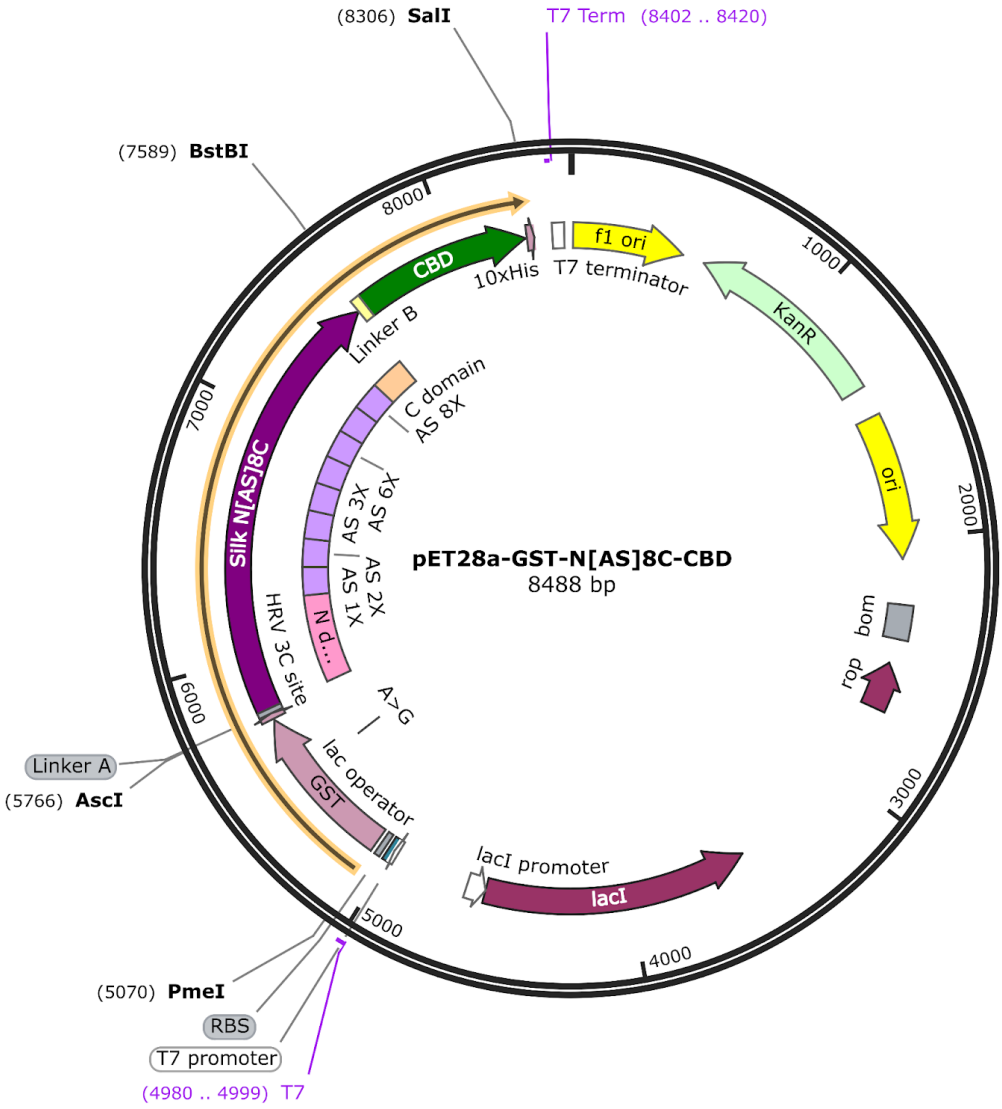


Figure 3 | 02a-pET28a-GST-N[AS]8C-CBD-10His plasmid map. *The 8x[AS] modules serve as our silk. The backbone of this plasmid is pET28a.*

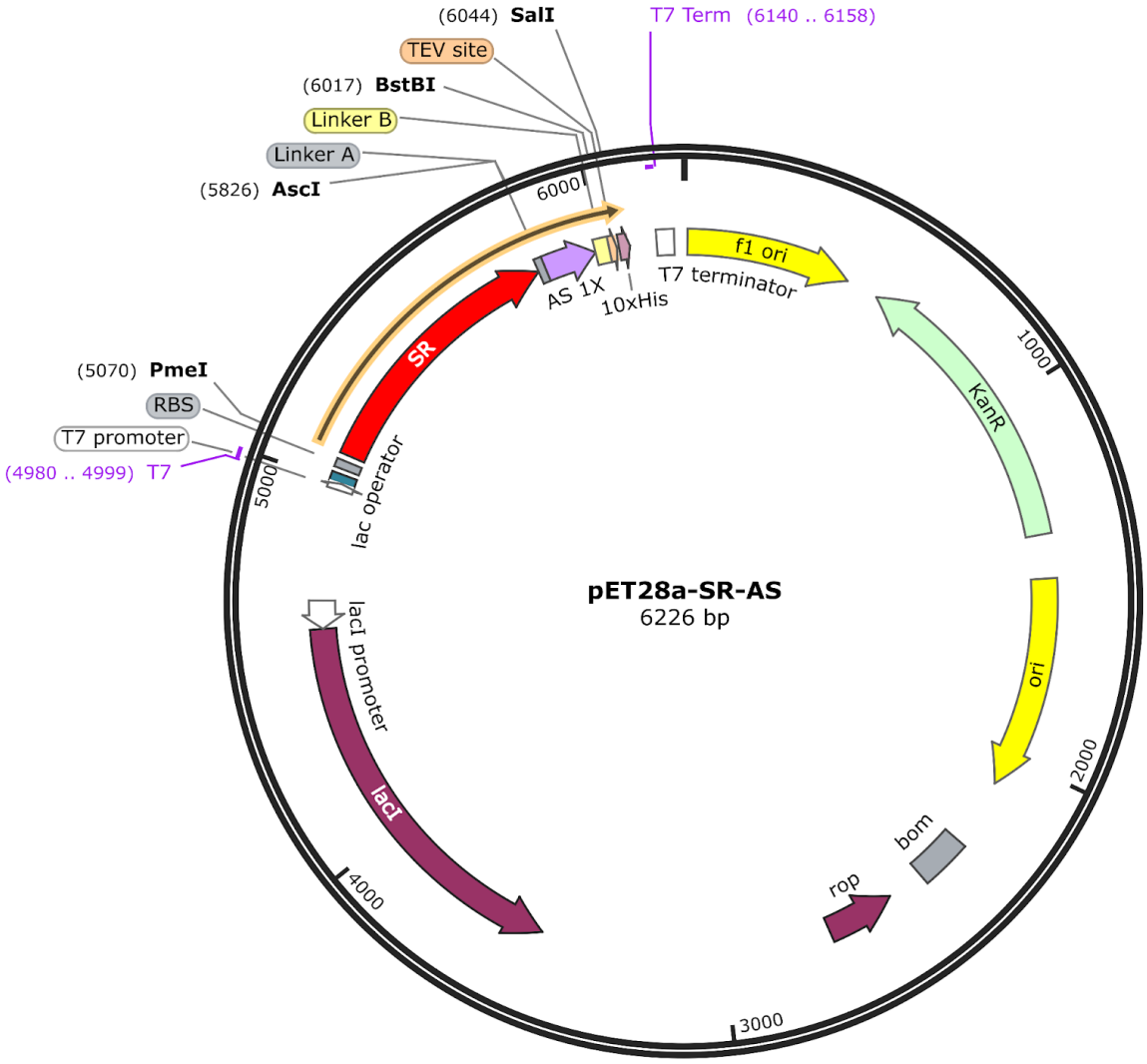


Figure 4 | 02b-pET28a-SR-AS-TEV-10His plasmid map. *The [AS] module was added for the polymerisation of our modular protein (added on top of the biofilm) with the silk. The backbone of this plasmid is pET28a.*

Characterization

Table 1 | Short characterization of the Green Lacewing silk used in HESTIA’s project.

Origin	*Mallada signata* (Chrysopidae) Green Lacewing
Type of expression	Recombinantly expressed in E.Coli BL21(DE3) cells Cytoplasmic expression No post-translational modifications
Structure	**Length:** 592 aa **Molecular weight:** 53 kDa **Secondary structure:** small β-sheets with an approximative sheet-length of 4 aa followed by turns. **Tertiary structure:** succession of β-helices formed with 2-3 small β-sheets.
Activity/function of the protein	Strength Extensibility Biocompatibility Biodegradability Surface hydrophobicity
Usefulness in our project	Form a biofilm being a hydrophobic coating for our aerogel of cellulose

3D model

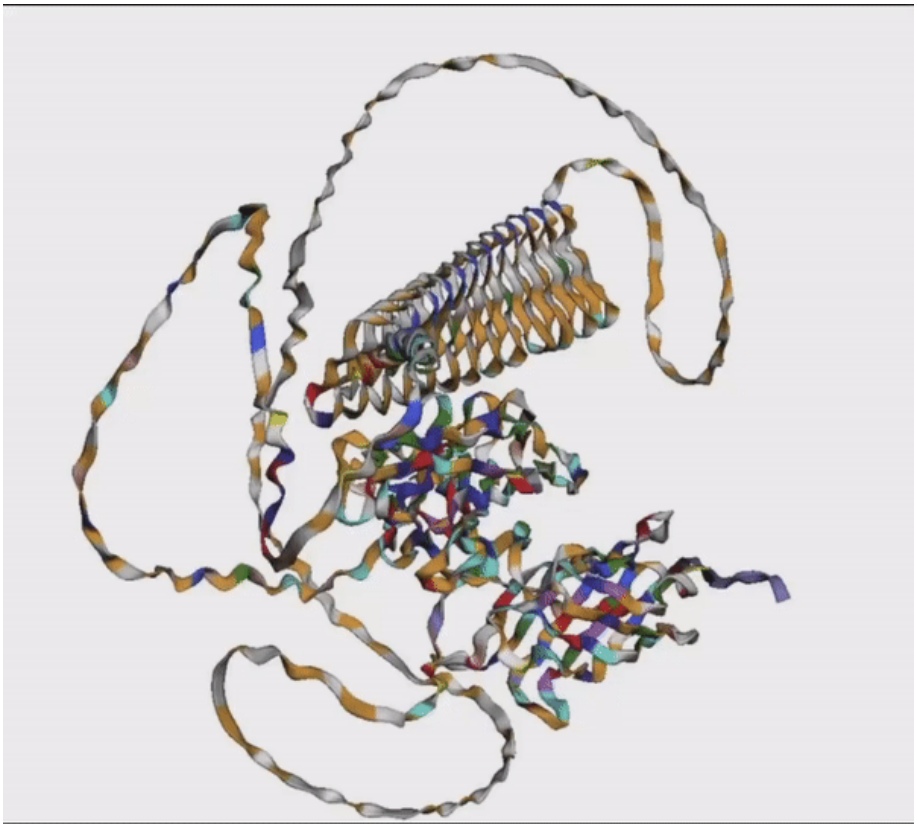


Figure 5 | 3D model of our 01a construct made using AlphaFold and Pymol. The repetitive domain on top corresponds to the [AS] modules of the silk.

Sequences

Nucleic acids’ sequence

“aactataacatttggagcaacgtgaacgcgcatccgaccgattgcggcaacagcggcggcagcagcggcagcagcgcggcgagcggc
gcggcgagcagcagcggcagcgcggcgggcagcggcgcggcgagcggcagcggcgcggcgagcggcagcggcgcggcgagcggca
gcggcgcggcgagcggcagcggcgcggcgagcggcagcggcgcggcgagcggcagcggcgcggcgagcggcagcggcgcggcgag
cggcagcggcgcggcgagcggcagcggcgcggcgagcggcagcggcgcggcgagcggcagcggcgcggcgagcggcagcggcgcg
gcgagcggcagcggcgcggcgagcggcagcggcagcgcgagcggcagcggcagcgcgagcggcagcggcagcagcagcggcagcg
gcagcagcggctgcggcagcggcggcagcgcgggcgcgagcagcaacggcagcagcgcgaccgcgagcaaaggcagcgcgggcgc
gaccagcaacggcagcaccgcggtggcgagcaaaggcagcgcgggcgcgagcagcggcaacagcaccgcgagcgcgaccaaaggc
agcgcgggcgcgagcagcaacggcagcagcgcgaccgcgagcaaaggcagcgcgggcgcgaccagcaacggcagcaccgcggtgg
cgagcaaaggcagcgcgggcgcgagcagcggcaacagcaccgcgagcgcgaccaaaggcagcgcgggcgcgagcagcaacggcag
cagcgcgaccgcgagcaaaggcagcgcgggcgcgaccagcaacggcagcaccgcggtggcgagcaaaggcagcgcgggcgcgagc
agcggcaacagcaccgcgagcgcgaccaaaggcagcgcgggcgcgagcagcaacggcagcagcgcgaccgcgagcaaaggcagcg
cgggcgcgaccagcaacggcagcaccgcggtggcgagcaaaggcagcgcgggcgcgagcagcggcaacagcaccgcgagcgcgacc
aaaggcagcgcgagcgcgagcagcgatggctttagcgcggcgtgcgatagcggcgaaagcgaagcggtggataaagcgaacctggcgg
cgattgcgaacattgcggcggcggcgggcaaaccggcggcgtgcggcagcgcgccgccgagcgatgattattatgattatggctgcggc”

Figure 6 | Exact nucleic acids’ sequence (modified from Bauer et. al (2012)). Retrieved from a.a seq using EMBL-EBI.

Amino acids’ sequence

NYNIWSNVNAHPTDCGNSGGSSGSSAASGAASSSGSAAGSGAASGSGAASGSGAASGSGAA
SGSGAASGSGAASGSGAASGSGAASGSGAASGSGAASGSGAASGSGAASGSGAASGSGAAS
GSGSASGSGSASGSGSSSGSGSSGCGSGGSAGASSNGSSATASKGSAGATSNGSTAVASKGSAG
ASSGNSTASATKGSAGASSNGSSATASKGSAGATSNGSTAVASKGSAGASSGNSTASATKGSA
GASSNGSSATASKGSAGATSNGSTAVASKGSAGASSGNSTASATKGSAGASSNGSSATASKGSA
GATSNGSTAVASKGSAGASSGNSTASATKGSASASSDGFSAACDSGESEAVDKANLAAIANIA
AAAGKPAACGSAPPSDDYYDYGCG

Figure 7 | Modified amino acids’ sequence (modified from Bauer et. al (2012)).

N-terminus – As Module x4 – C-terminus
1x AS Module: GSAGASSNGSSATASKGSAGATSNGSTAVASKGSAGASSGNSTASATK

HESTIA’s Silk Protocols

Purification Protocol using the MagneHis Kit from Promega

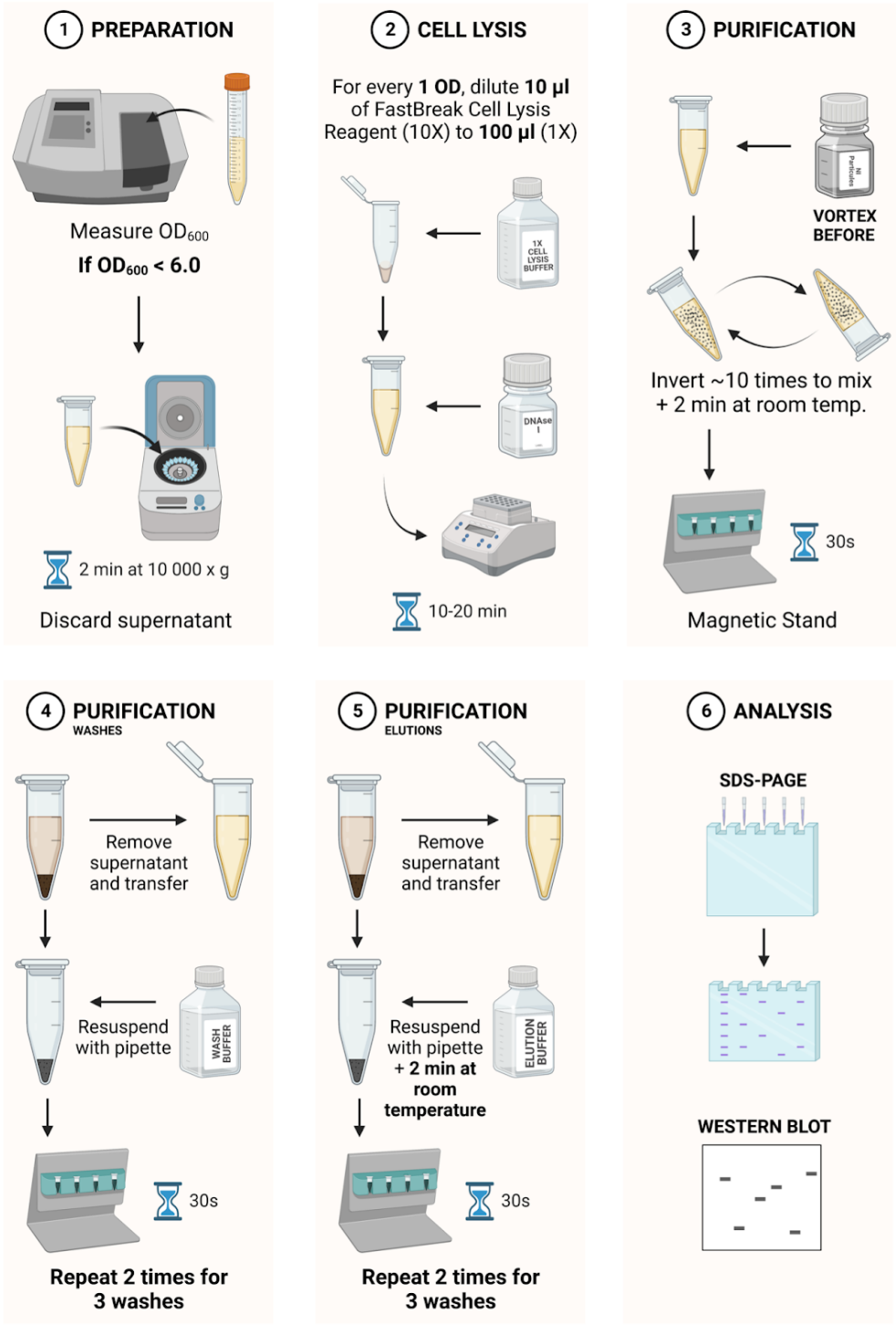


Figure 8 |Diagram of Procedure

Materials

- 10X FastBreak™ Cell Lysis Reagent
- 1M imidazole solution
- Lyophilized DNase I
- MagneHis™ Ni-Particles
- 5M NaCl solution or solid NaCl
- Magnetic stand
- MagneHis™ Binding / Wash Buffer
- MagneHis™ Elution Buffer
- Bacterial culture induced with IPTG for GFP production

Protocol

Step 1: Bacterial Cell Lysis

- ☐ Determine the OD_{600} for the fresh bacterial culture. *The cell culture should have a final OD_{600} < 6.0 for efficient processing.*
- ☐ Centrifuge 1ml of bacterial culture at 10,000 x g for 2 minutes in a microcentrifuge. Remove the supernatant completely. *We used a paper to get rid of the final supernatant residu. No pipetting is needed.*
- ☐ For every 1 OD_{600}, dilute 10 µL of FastBreak™ Cell Lysis Reagent, 10X, to 100 µL (1X) by adding 90 µL of double-distilled water. Do not process more than 1 OD_{600} of culture per 100 µL of 1X FastBreak™ Cell Lysis Reagent.
- ☐ Resuspend the cell pellet in 1X FastBreak™ Cell Lysis Reagent.
  After this step, always vortex and put the sample on ice after pipetting !
- ☐ Resuspend lyophilized DNase 1 as indicated on the vial if not already done (80 µl of water added to the DNase 1), and add 1 µL to the lysed bacterial culture. Store the resuspended DNase 1 in aliquots at -20°c for long term or at 4°C for up to one week.
- ☐ Incubate with shaking for 10-20 minutes at room temperature on a rotary mixer or shaking platform.

Step 2 : Purification of Polyhistidine proteins from 1 mL of bacterial culture using MagneHis™ Ni-Particles

- ☐ First of all, do not forget to label all the tubes.
- ☐ Vortex the MagneHis™ Ni-Particles to a uniform suspension. *It is very important to vortex the beads because they always clump together.*
- ☐ Add 30 µL of MagneHis™ Ni-Particles either to cell pellet resuspended in 1X FastBreak™ Cell Lysis Reagent.
  Note: You may need to increase the amount of MagneHis™ Ni-Particles used for high-expressing proteins.
- ☐ Invert tube to mix (approximately 10 times), and incubate for 2 minutes at room temperature. Make sure the MagneHis™ Ni-Particles are well mixed.
- ☐ Place the tube in the appropriate magnetic stand for approximately 30 seconds to capture the MagneHis™ Ni-Particles. Using a pipette, carefully remove the supernatant, and put it in a new 1.5 mL tube.

- ☐ Remove the tube from the magnetic stand. Add 150 µL of MagneHis™ Binding/Wash Buffer to the MagneHis™ Ni-Particles. *Resuspend up and down by pipetting the beads with the buffer.*
- ☐ Place the tube in the appropriate magnetic stand for approximately 30 seconds to capture the MagneHis™ Ni-Particles. Using a pipette, carefully remove the supernatant, and put it in a new 1.5 mL tube.
- ☐ Repeat the wash step 2 times for a total of 3 washes. *Always resuspend up and down by pipetting the beads with the buffer.*
- ☐ Remove the tube from the magnetic stand. Add 100 µL of MagneHis™ Elution Buffer, and pipet to mix (or vortex).
- ☐ Incubate for 1–2 minutes at room temperature. Place in a magnetic stand to capture the MagneHis™ Ni-Particles. Using a pipette, remove the supernatant containing the purified protein and put it in a new 1.5 mL tube.
- ☐ Analyze the samples by SDS-PAGE or by functional assay.

Purification Protocol using Ni-NTA Beads from the Protein Production and Structure Core Facility of EPFL

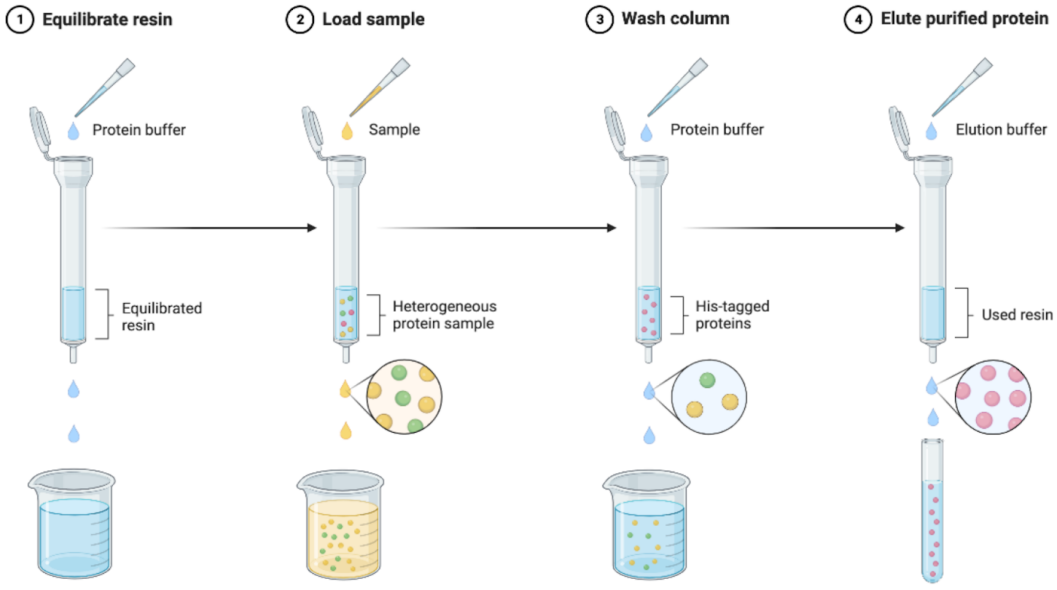


Figure 9 |Diagram of Procedure

Materials

- 2.5M imidazole stock solution
- 5M NaCl solution
- Protease inhibitor tablet
- Turbonuclease
- Sonicator
- 1M HEPES

- HisPur Ni-NTA beads in 20% ethanol
- Disposable plastic column
- Centrifuge going to 20 000 xg
- 0.30 µm filters
- Tube of bacterial culture induced with IPTG

Protocol

Step 1 : Prepare buffers

- ☐ Wash Buffer A (1L) **:**
  - ☐ 300 mM NaCl
  - ☐ 20 mM HEPES pH 7.5
  - ☐ 20 mM imidazole (for non specific binding)
- ☐ Elution Buffer B (250 mL)
  - ☐ 300 mM NaCl
  - ☐ 20 mM HEPES pH 7.5
  - ☐ 500 mM imidazole pH 7.5

We use stocks of 5M NaCl, 1M HEPES pH 7.5 and 2.5M imidazole pH 7.5.

Step 2 : Preparation of our sample

- ☐ Defrost your tube of pellet in water, or obtain a cell pellet by centrifugation at 5000xg for 10 min at 4°c.
- ☐ Dilute the pellet with the wash buffer (ideally we need big volumes). We resuspent our cell pellet into 20 mL. *Bigger the volume is, the better. However, be careful to not completely drown the pellet with the wash buffer.*
- ☐ Add some bonus stuff which are useful :
  - ☐ Add 1 protease inhibitor tablet. As previously, just approximate and don’t pipette it.
    Protease inhibitor protects the proteins sensitive to degradation and inhibit their destruction. Quite expensive.
  - ☐ Add 5 µL of turbonuclease. Pipette the precise volume.
    Turbonuclease cleaves DNA to avoid the formation of jelly that is due to the presence of undestroyed DNA or RNA. Therefore, it ensures a better fluidity for our sample.
- ☐ Mix well and vortex quickly, it makes the sample more homogenous. We can be harsh to cells without fragilising them.
  WARNING : After this step do not vigorously shake your sample !
  After cell lysis, proteins are not within the protective environment of the cell.
- ☐ Put the cultures on a shaker at 4°c for 20min.
- ☐ Lyse cells with a sonicator at 70% power, 10 secs on and 10 secs off pulse cycle over 2.5 minutes (in total, set the machine for 5 minutes). The samples will get hot. Then, put the sample on ice.

Place the sample on ice in a beaker. Wash the tip with alcohol and water before and after.
Sonication uses sound waves to explode the cells.

- ☐ Centrifuge the lysed cells (lysate) at 20 000 x g for 30 minutes (1 hour or 2 hours could still be fine).
  For the balances for the centrifuge, take a flask and fill it approximately with the same volume of the sample with water.
  To open the machine, push the cover and pull it out. Stay next to the machine till it shows that it has reached the wanted speed. The sample should be less viscous after the centrifugation.

Step 3 : Resin preparation

- ☐ If needed, pour out a bit of ethanol from the Ni-NTA beads in solution.
- ☐ Equilibrate 2-5 mL Ni-NTA beads into the wash buffer (this removes the 20% ethanol in which the beads are stored), 2 mL for low expressing protein constructs, 5 mL for high expressors.
- ☐ Measure out 2-5 mL of settled beads (*we measured 25 mL of the beads in ethanol which corresponded to 5 mL of settled beads)*
- ☐ Empty (as much as possible) the ethanol without letting go of beads.
- ☐ Repeat 3 times :
  - ☐ Add the wash buffer to around 40-50 mL, mix a bit.
  - ☐ Centrifuge the tube at 500g for 5 minutes (do not forget to balance the centrifuge with a tube of water).
  - ☐ Empty (as much as possible) the supernatant.

Step 4 : Incubation of lysate with Ni-NTA resin

- ☐ When the centrifuge is finished, carefully transfer the supernatant into a beaker on ice. Keep some in the tube for the SDS-PAGE.
- ☐ With a syringe, aspirate the supernatant CAREFUL NOT TO MAKE BUBBLES.
- ☐ *If you are using a machine : filter the buffers using a 0.45 µm filter.*
- ☐ Add imidazole to the filtered supernatant to 25 mM. This reduces non-specific binding of endogenous His-rich *E.coli* proteins.
- ☐ Decant the excess liquid from the washed beads.
- ☐ Add the supernatant to the beads.
- ☐ Put the beads with the supernatant rotating overnight at 4°c.

Step 5 : Purification

- ☐ Setup a stand with a plastic disposable column.
- ☐ Transfer the beads to a disposable plastic column.
- ☐ Wash the beads in sequential order to remove loosely bound proteins before elution. The washes contain increasing amounts of imidazole to remove unbound, unspecific

proteins from the resin. Do not forget to label everything and to keep everything to know where the proteins are. Keep all your fractions on ice.

- ☐ Wash 1 - 20 mM imidazole - 30 mL
- ☐ Wash 2 - 50 mM imidazole - 30 mL
- ☐ Elute 1 - 250 mM imidazole - 30 mL
- ☐ Elute 2 - 500 mM imidazole - 30 mL

  If needed:
- ☐ Elute 3 - 1 M imidazole - 20 mL
- ☐ Elute 4 - 2.5 M imidazole - 10 mL
- ☐ Elute 5 - 5 M imidazole - 10 mL

We had to do more elution steps than planned because the control (GFP protein) which contains both the CBD domains and the mSA didn’t eluted at the normal elution concentration.

In what fraction the protein elutes depends on the size of the His-tag, where the tag is (N or C) and the 3D context of the target (is the tag accessible to the resin).

Step 6 : Concentration

- ☐ Concentrate the protein using a filter of 30kDa.
- ☐ Try to remove the imidazole excess by puting the concentrated proteins under dialysis. The dialysis buffer was composed of 300mM NaCl and 20 mM HEPES.
- ☐ To see if there are protein aggregations, we centrifuged the result after dialysis. *If a pellet appears, it means that there is an aggregation of protein which may mean that the proteins are not functional anymore.*
- ☐ Finally, measure the protein’s concentration after dialysis.

Silk Biofilm Protocol

General description of the method in [Development of an artificial silk protein on the basis of a lacewing egg stalk protein]

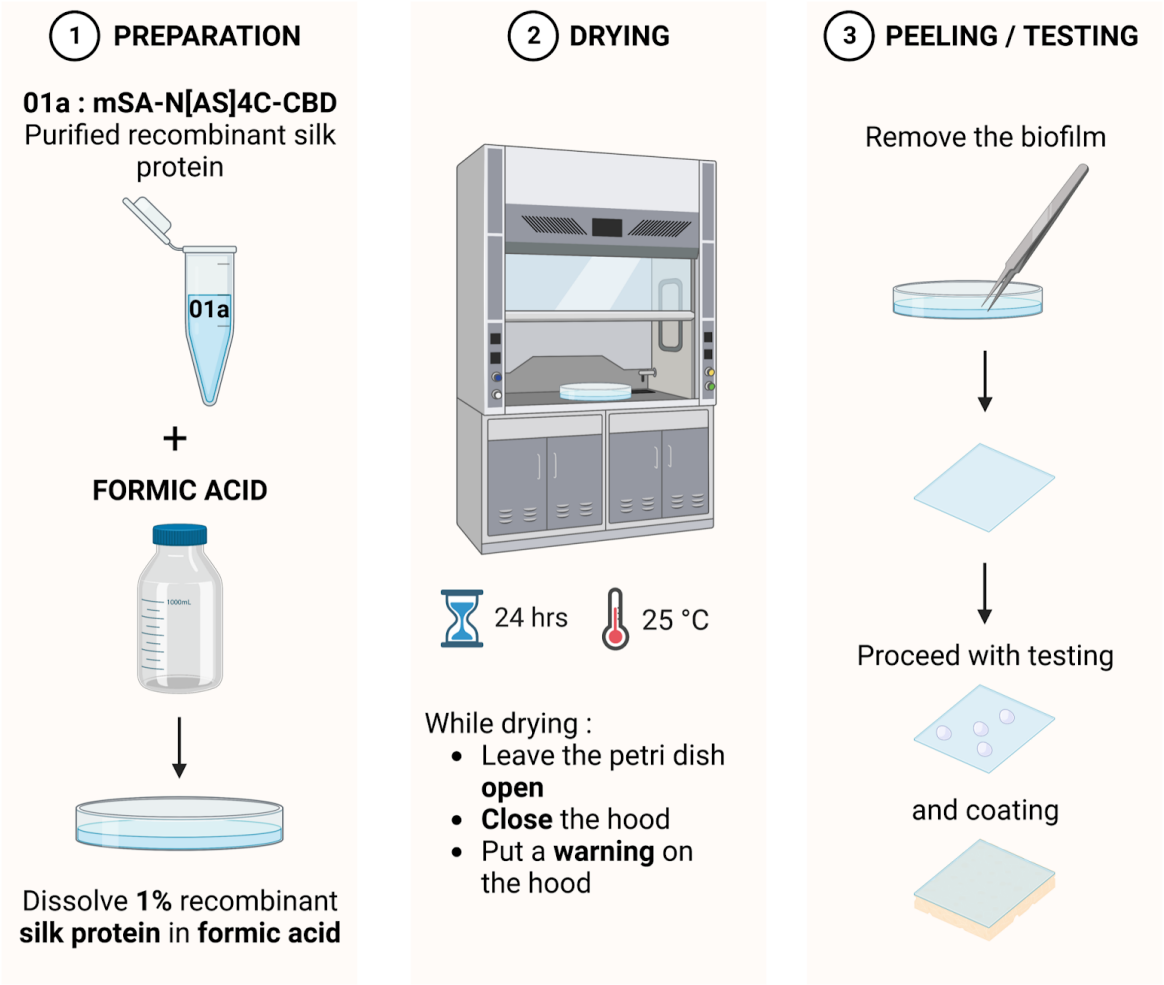


Figure 10 |Diagram of Procedure

Materials

- Polystyrene petri dish
- Formic acid
- Our recombinant silk protein 01a (mSA-N[AS]4C-CBD) purified

Protocol

Biofilm preparation

- ☐ Dissolve 1% of (w/v) mSA-N[AS]4C-CBD (recombinant silk protein) in formic acid.

To determine the amount of protein required depending on the volume of formic acid (or vice versa), the online [Percent Solutions Calculator from PhysiologyWeb] can be used.

We have a concentration of XXX mg/mL for the silk protein.
We have at our disposition XXX mL of the silk protein solution, which corresponds to m_silk = XXX mg/mL x XXX mL = XXX mg.
So if we want a 1% (w/v) solution of silk protein in formic acid, we need the following volume of formic acid:
V_formic_acid = 100 [µL/mg] x m_silk [mg] = XXX µL
Reminder: the units of % (w/v) are in kg/L, or g/mL or mg/µL.

- ☐ Cast formic acid on polystyrene (petri dish).

Drying step

- ☐ Dry the film with an airing chamber for 24h (let the lid of the petri dish open).

Recover the biofilm

- ☐ Peel off the substrate

Green Lacewing silk ressources :

[Development of an artificial silk protein on the basis of a lacewing egg stalk protein]