Ultra-high throughput evolution of a bacterial endolysin as a new antimicrobial agent

Vilnius-GMC Lab Notebook

In this document, we provide a timeline of events and experiments that we took while participating in iDEC 2021 competition. Since we used protocols that are widely accepted by the scientific community and we are familiar with the methods, we did not include them once more in this lab book. The protocols used are documented in the report and under the „Protocols“ section on this wiki page. Here we document data and thinking processes, some of which are not included elsewhere. Each part of the work has a timestamp and who conducted it: LC – Lorenzo Camisi, JJ – Jonas Juozapaitis.

2021-05-25 JJ **Peptide synthesis companies**

Starting to look for companies able to synthesize FRET D, L peptides.
Contacting: Lifetein, JPT Peptide Technologies, GenScript, BioCat, Pepscan.

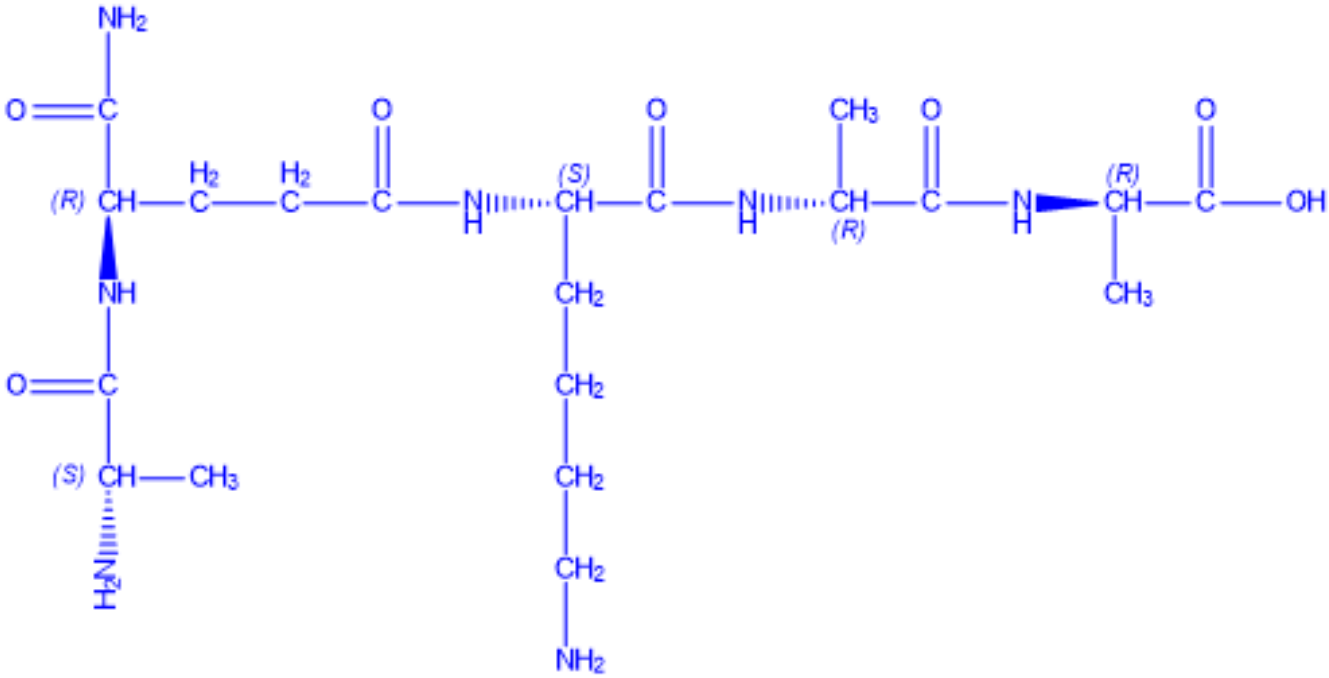
2021-05-25 – 2021-06-28 JJ&LC **Target proteins selection**

Bacillus subtilis **LytE** https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2958.2008.06211.x

Enterococcus faecium **SagA** https://pubs.acs.org/doi/10.1021/acs.biochem.0c00755

2021-06-20 JJ&LC **Substrate peptide design**

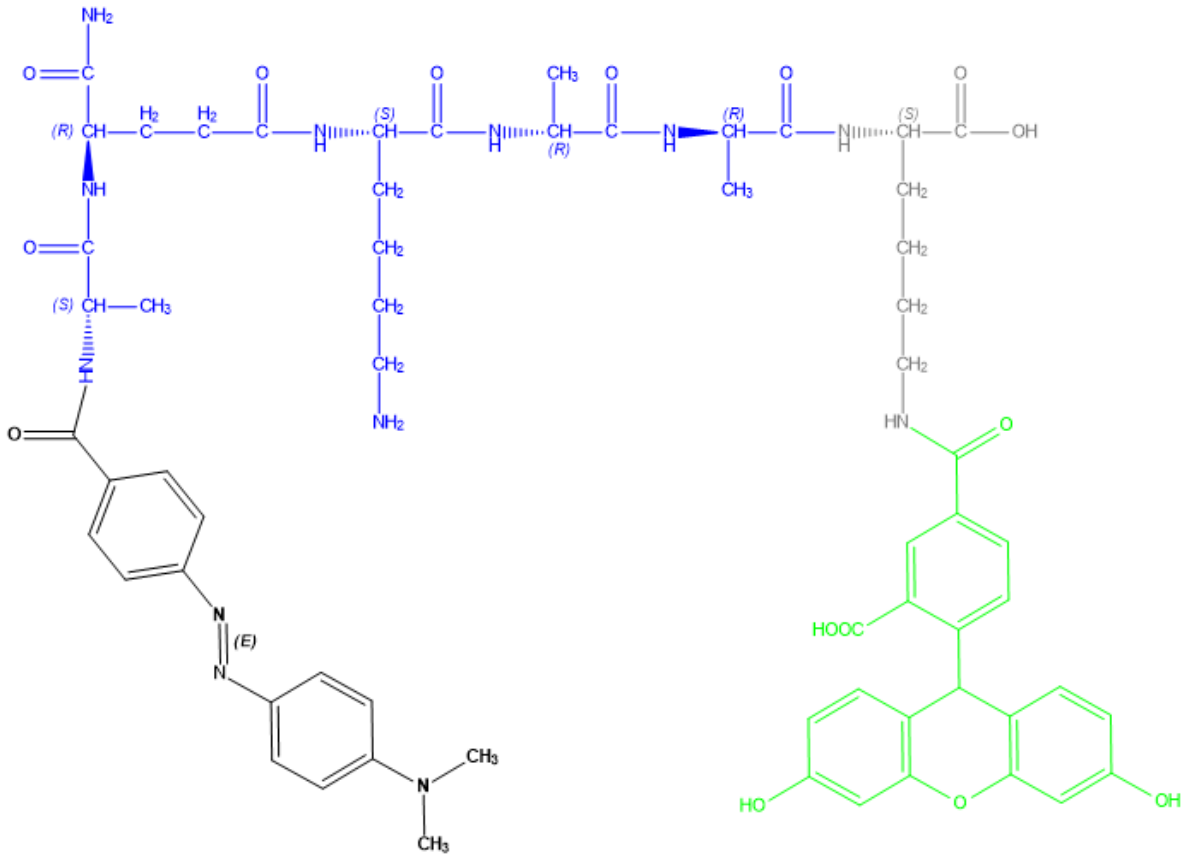
Since LytE and SagA are cutting *S.aureus* peptidoglycan linking peptide Ala-{D-isoGln}-Lys-{D-Ala}-{D-Ala}, we selected to use this peptide as a substrate for the enzyme.



2021-06-22 JJ **Ordering target peptide**

The peptide will be synthesized by LifeTein, because of the affordable price and ability to make the peptide of interest. Also, the peptide contains a 6th amino acid lysine which is required to attached 5‘ FAM fluorophore.

Final peptide structure:



2021-06-28 JJ **Looking for S. aureus and BSL-2 laboratory**

JJ Found *S. aureus* DGCC7710 lab strain we could borrow. Finding out in which laboratory we can work with BSL-2 organisms.

2021-06-28 LC **Plasmid maps**

Creating a Benchling project for iDEC, importing plasmids, looking up gene sequences

2021-06-29 JJ **Verifying protein sequences**

SagA protein has some X instead of amino acids in the annotated protein sequence. We are not able to find the full protein sequence, only the NlpC/P60 domain in the paper (Espinosa et. al, 2020) and cited previous works.

Creating multiple sequence alignment (explained in detail in the report and methods) of highly similar SagA homologs and looking up what amino acids are in the place of X. SagA homologs of *Enterococcus* are highly similar. Decide what amino acids to use and report back to LC.

2021-06-29 LC **Plasmid design**

FInalyzing plasmid sequences, adding restriction and silent mutations to remove restriction sites of enzymes that will be used for cloning sites to synthetic DNA fragments, designing correct RBS and promoter sequences. Using Benchling.com online tools and Whatcut for analysis (http://watcut.uwaterloo.ca/template.php).

We will use pKTCTET vector. It is a high copy number vector with kanamycin resistance. Genes will be cloned using standard restriction-ligation. Isolated NlpC/P60 domains will be cut and cloned into the same vectors. When generating mutant libraries, NlpC/P60 domain coding DNA will be amplified by error-prone PCR and ligated back to the vector. This way mutagenesis will be performed only on the NlpC part of the proteins.

Also, we clone protein genes while also fusing them with C-terminal His tags just to be able to verify protein expression by Western blotting (if needed) or for easy purification of proteins.

Protein expression from this plasmid is induced by IPTG.

Plasmid maps can be found under the section „Supplementary files“

2021-06-29 LC **Ordering genes**

LC is ordering the genes from “Twist Bioscience”

Gene sequences:

LytE:

ctacatacatATGAAGAAACAGATAATTACAGCAACCACCGCCGTGGTATTGGGTAGTACTCTCTTTGCCGGGGC
CGCATCGGCGCAGTCAATTAAAGTTAAAAAAGGCGATACTCTTTGGGACCTGAGCCGTAAGTATGATACCACT
ATTTCTAAGATCAAATCAGAAAATCATCTGCGTAGTGATATTATCTATGTCGGACAGACACTGAGCATTAATGG
TAAATCCACTAGCTCCAAATCCAGCAGCTCAAGCAGCTCTAGTAGTACCTACAAAGTGAAATCGGGGGATTCG
CTGTGGAAGATTTCGAAAAAGTACGGTATGACCATCAACGAGCTCAAAAAGCTGAACGGCCTGAAATCTGAC
TTACTGCGGGTTGGCCAAGTGTTAAAGCTGAAAGGTAGTACCTCCAGTTCCTCTAGTTCAAGCTCCAAAGTCA
GCTCTTCGAGTACGAGTACGTACAAGGTTAAGTCCGGTGATTCCTTGTCCAAAATCGCTTCTAAATATGGCACC
ACCGTCAGCAAACTGAAATCACTCAACGGTCTTAAAAGCGATGTGATTTACGTTAATCAAGTCTTAAAAGTCA
AAGGCACCAGCACaTCGAGCAGCAAACCAGCTTCAAGCTCTTCCTCGTCGTCAAGCAAAACCAGCAGCACCTC
ACTGAATGTTTCCAAACTTGTGAGTGATGCAAAAGCGCTAGTGGGAACACCGTATAAATGGGGAGGTACGAC
GACGTCTGGTTTCGATTGCAGTGGCTTCATCTGGTACGTACTGAATAAGCAGACGTCTGTGGGGCGCACTAGC
ACAGCGGGGTATTGGAGTTCTATGAAATCAATTGCGTCACCTTCAGTGGGCGACTTTGTATTCTTTACGACGTA
TAAATCTGGCCCCTCGCAcATGGGCATCTATATTGGGAACAACAGCTTTATCCACGCGGGTAGCGACGGCGTT
CAGATATCCTCGTTGAATAACTCGTACTGGAAACCGCGCTATCTGGGCGCCAAACGATTTctcgagcaccacc

SagA:

agatatacatATGAAAAAGTCTTTAATCAGCGCTGTTATGGTTTGTAGTATGACTCTCACCGCAGTCGCCTCCCCG
ATCGCCGCCGCCGCTGATGATTTCGACTCCCAGATTCAGCAGCAAGATCAGAAAATTGCGGACCTCAAAAACC
AGCAAGCGGACGCGCAAAGCCAGATAGATGCGCTGGAATCGCAGGTATCGGAGATTAATACCCAGGCACAG
GATCTGCTTGCCAAACAAGATACCCTACGCCAAGAAAGCGCTCAGCTGGTTAAAGATATCGCTGATCTGCAAG
AACGCATCGAAAAGCGTGAGGACACTATCCAGAAACAAGCGCGCGAGGCACAGGTGAGCAACACCAGTTCC
AACTACATCGATGCGGTTTTAAATGCAGATTCCCTGGCCGATGCGATCGGTCGTGTTCAGGCGATGACCACCA
TGGTGAAAGCCAACAACGATTTGATGGAACAACAGAAACAAGACAAAAAGGCCGTTGAAGACAAAAAAGCG
GAAAATGACGCTAAACTGAAAGAGCTGGCAGAGAACCAGGCAGCGCTGGAGAGTCAAAAAGGCGACCTTCT
GAGCAAACAGGCGGATCTGAACGTCCTGAAAACCTCGCTTGCGGCCGAGCAGGCTACGGCTGAGGACAAAA

AAGCCGATTTAAATCGTCAGAAGGCCGAAGCGGAAGCCGAACAGGCACGCATTCGCGAACAGCAGCGTTTG
GCAGAACAGGCGAGACAGCAAGCTGCACAGGAAAAGGCAGAGAAAGAAGCGCGTGAACAGGCTGAAGCCG
AGGCACAAGCAACCCAAGCCTCAAGTGCAGCCCAGTCCTCCGCTACCGAAGAATCGTCGGCGGCGCAATCCT
CTACCACCGAGGAAAGCAGCAGCGCAGCACAGTCATCAACCACCGAAGAATCGACGACGGCGCCGGAAAGC
TCCACCACAGAAGAAAGTACGACGGCACCGGAATCGAGTACTACCGAAGAATCAACCACTGCCCCTGAAAGC
AGCACTACGGAAGAGAGCACGACCGTGCCGGAATCATCCACTACTGAAGAGTCAACTACGGTCCCGGAAAGC
TCAACAGAAGAGTCGACGACACCGGCTCCGACAACGCCTTCTACTGATCAGTCTGTGGATCCaGGtAACAGTA
CaGGaTCGAATGCGACGAACAATACCACTAATACAACCCCTACCCCAACAACAAATACAACACCAACGCCGAC
GCCTAGCGGGAGTGTGAACGGTGCCGCGATTGTCGCGGAAGCTTATAAATATATAGGGACGCCtTATGTCTG
GGGAGGTAAAGACCCCAGTGGATTCGACTGCAGTGGCTTTACGCGGTACGTGTACCTCCAGGTAACCGGGCG
AGATATTGGTGGTTGGACCGTACCCCAAGAGAGCGCGGGCACCAAGATTTCTGTGTCTCAGGCGAAAGCCGG
CGATCTGCTGTTTTGGGGATCTCCGGGCGGTACGTATCACGTAGCCATTGCCTTGGGCGGTGGGCAATATATT
CATGCGCCCCAGCCAGGCGAGAGCGTGAAGGTTGGTTCTGTGCAGTGGTTCGCGCCGGATTTTGCGGTGTCA
ATGCTCGAGCACCACC

2021-06-29 LC **Ordering primers**

Primer pair will be used for amplifying synthetic DNA fragments and for Sanger sequencing of cloned genes:

T7 fwd GATCCCGCGAAATTAATACGACTCACTATAGG

T7 rev CAGCAGCCAACTCAGCTTCCTTTCG

Primers will be used for cutting NlpC/P60 domains from LytE and SagA and cloning them into the same vectors. This way plasmids coding only the NlpC/P60 domains will be created. At this point in the project, it is unknown which constructs will be suitable for our experiments.

SagA_P60_flankig_NdeI_fwd CGTTATACATATGCCaGGtAACAGTACaGGaTC

LytE_P60_flanking_NdeI_fwd GGTTATACATATGAGCAAAACCAGCAGCACCTCA

2021-06-30 LC&JJ **Writing abstract for the project**

LC also filled in the Responsible research form.

2021-08-05 LC **Substrate peptide arrived**

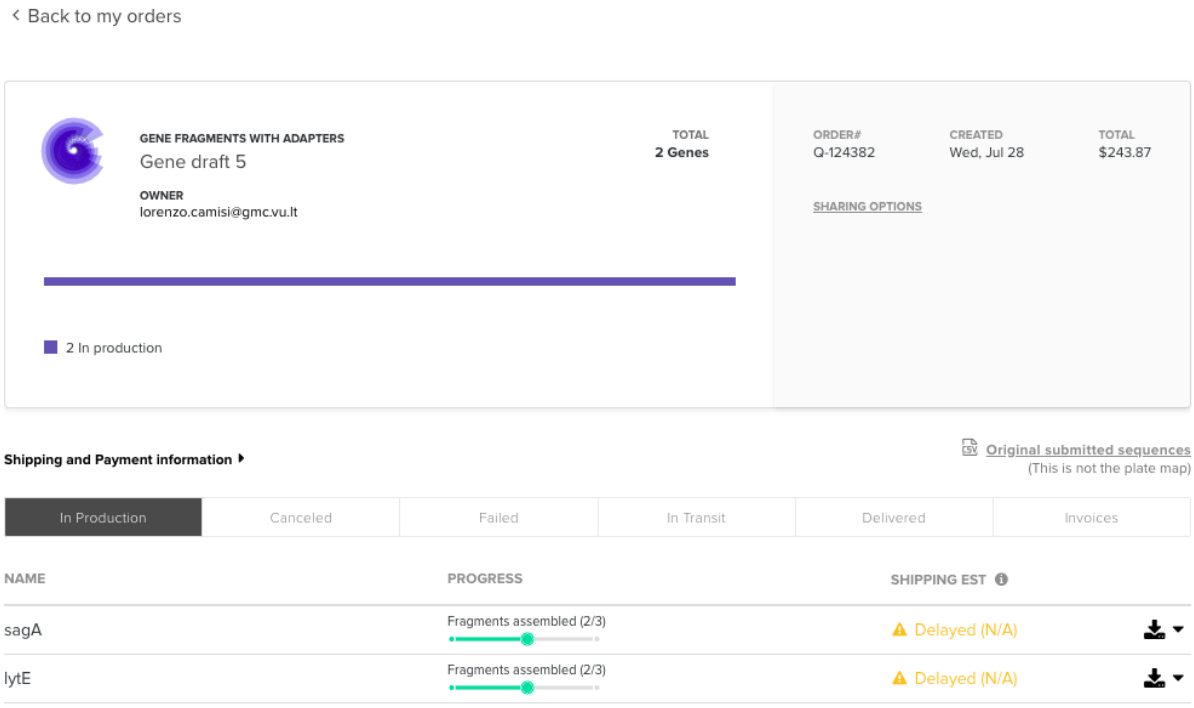
2021-08-11 **TwistBioscience delays synthetic DNA arrival**

The peptide is dissolved in water (500 ng/µl) and aliquots are frozen. Tubes are covered from light with tin foil.

2021-08-12 JJ **Preparation of competent *E. coli***

Preparing chemically competent *E. coli* for transformation and protein expression. For transformation, we will use well-known DH5α cells and for expression Bl21 (DE3). 50 ml of bacteria are grown, and competent cells are prepared according to the protocol. This yields 50 aliquotes of competent cells that are frozen az glycerol stocks.

2021-08-22 **Another delay at TwistBioscience**



2021-09-06 **Genes finally arrive**

2021-09-07 LC **Cloning of full-length SagA and LytE**

Gene fragments are dissolved in 10 µl of water. Vector backbone and fragments were digested according to the protocol and ligated at ratio (vector: insert 1:3).

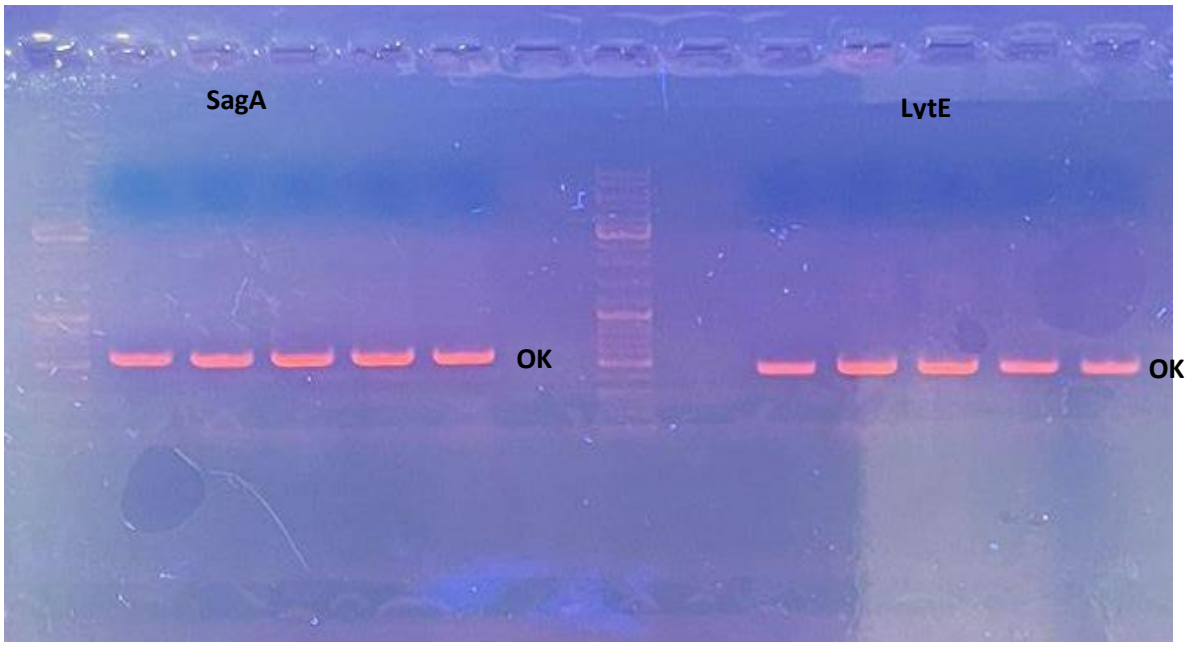
Fragment concentration after purification: Vector: 25 ng/ µl, SagAP60 5 ng/µl, LytE 6 ng/µl.

This calculator is used for calculating the volume of ligation mix: https://nebiocalculator.neb.com/#!/ligation.

E. coli DH5α is transformed with the ligation mixtures according to the protocol and left overnight for the colonies to form.

2021-09-08 LC **Colony PCR of SagA and LytE**

The next morning colonies are formed. Colony PCR is performed according to the protocol. All the colonies tested have the correct insert sequence. Plasmids were amplified and purified for sanger sequencing.



2021-09-11 JJ **Sanger sequencing**

Plasmids and primers were aliquoted and sent for Sanger sequencing. 2 variants in total: SagA and LytE.

2021-09-12 LC ***E.coli* BL21 transformation**

Full-length gene variants were transformed into BL21 strain for protein expression, while still waiting for Sanger sequencing results. In the case of correct cloned sequences, this would save precious time.

2021-09-16 LC **Cloning of NlpC/P60 fragments.**

Only the NlpC/P60 fragments were amplified using the primers described in the previous chapter. Restriction enzymes were added to the PCR reactions, incubated for 1 hour, and later purified and cloned according to the protocol.

2021-09-17 LC **Colony PCR of SagAP60 and LytEP60**

Colony PCR is performed according to the protocol. All the colonies tested have the correct insert sequence. Plasmids were amplified and purified for Sanger sequencing.

2021-09-13 LC **Night culture for protein expression**

Single colonies are picked and 4 ml of Lb with kanamycin is inoculated.

2021-09-14 LC **Protein expression**

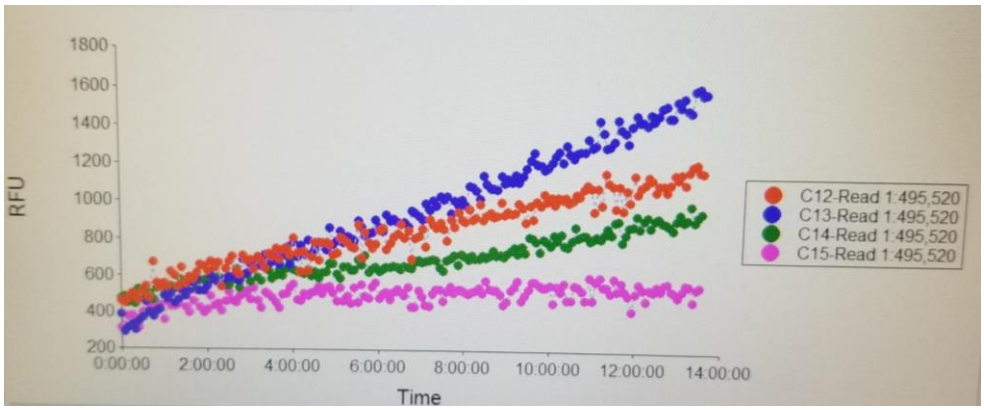
In the morning 1 ml of night cultures were added to the 30 ml of fresh Lb medium supplemented with antibiotic and incubated until OD_{600} reached 0.5. Cells were induced, later collected (1 ml of OD_{600} 1.0 of cells or equivalent, depending on the sample).

2021-09-14 LC **Microtiter plate assay**

Cells were sonicated as described in „Protein expression“ and „Protein purification“ protocols. Lysates were centrifuged for 1 hour on tabletop centrifuge at 13.2 rpm. Reactions for microtiter plate assay were prepared as described in the protocol and 1/20 reaction volume of lysate was used for the reaction. Reaction in the microtiter plate assay was incubated overnight.

2021-09-15 LC&JJ **Microtiter plate assay results**

This is our first indication, that at least one of our proteins of choice is able to digest the FRET beacon peptide. This was the first rough proof-of-concept experiment that encouraged us to continue this work. At this point, we realized that lysates is not an ideal form of enzyme introduction into the reaction. We also wanted to test the assay using all 4 variants of enzymes (SagA, LytE, SagAP60, and LytEP60).



Red - control with cell lysate, Blue – saga, Green – LytE, Pink – buffer and substrate peptide.

2021-09-16 LC ***E.coli*** **BL21 transformation**

NlpC/P60 variants were transformed into BL21 strain for protein expression, while still waiting for Sanger sequencing results. In the case of correct cloned sequences, this would save precious time. We also decided to use a control protein retro-aldolase (indicated RA). A good control would be SagAP60 active site mutant, but it takes some time to clone it, and LC already had another protein in the same plasmid. This protein will be used as a control.

2021-09-20 LC **Night culture preparation**

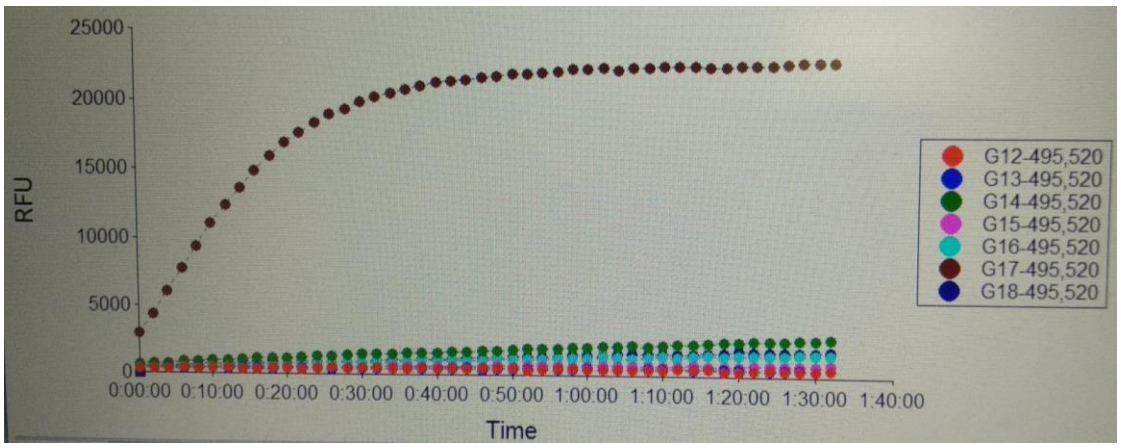
Single colonies are picked and 4 ml of LB with kanamycin is inoculated (the 4 variants for which we wait the sequencing results).

2021-09-21 LC **Protein expression**

In the morning 200 µl of night cultures were added to the 5 ml of fresh LB medium supplemented with antibiotic and incubated until OD_{600} reached 0.5. Cells were induced, later collected (1 ml of OD_{600} 1.0 of cells or equivalent, depending on the sample was used for cell lysis).

2021-09-21 LC **Microtiter plate assay**

Cells were sonicated as described in „Protein expression“ and „Protein purification“ protocols. Lysates were centrifuged for 1 hour on a tabletop centrifuge at 13.2 rpm. Reactions for microtiter plate assay were prepared as described in the protocol and 1/20 reaction volume of lysate was used for the reaction.

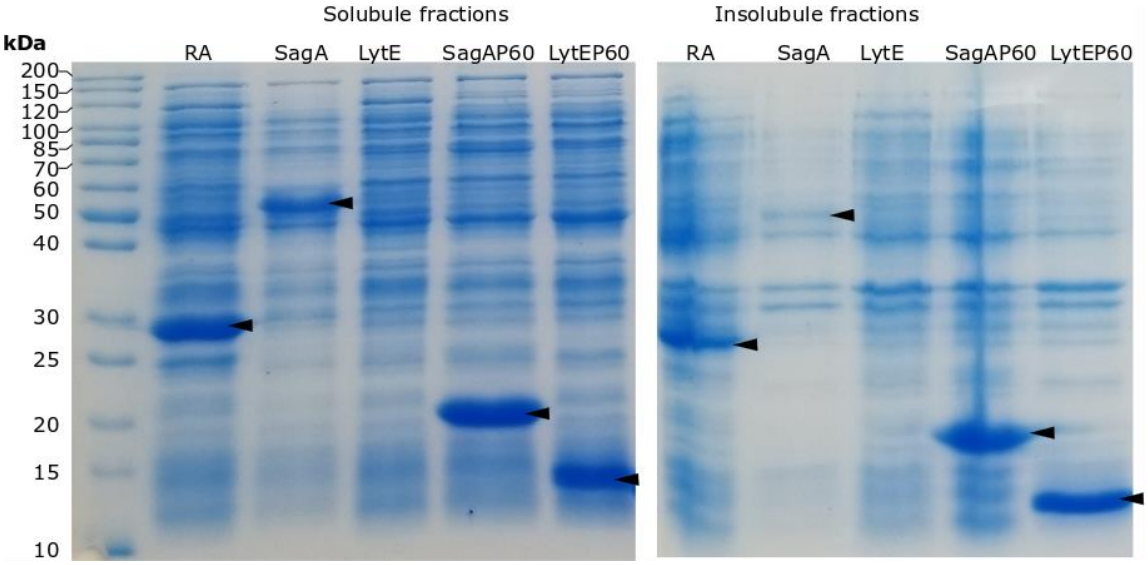


Red – buffer with substrate, Blue – control uninduced cell lysate, green – RA, Pink – SagA, Cyan – LytE, Brown – SagAP60, dark Blue – LytEP60.

We learned that SagAP60 fragment (brown) is active and cuts the target peptide thus increasing the fluorescent signal. We also observe that the fluorescent signal in the case RA is higher than the background (Red).

2021-09-22 JJ **Protein expression assay**

Protein expression (of the lysates that were used in the assay) was tested by the SDS-PAGE as described in the „Protein expression assay“ protocol.

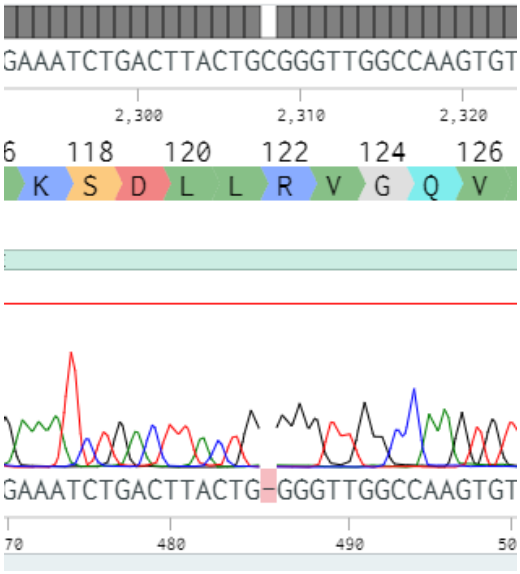


This revealed, that LytE is not expressed (activity of LytE lysate was also not observed). We also decided to repeat the assay using spin-column purified proteins, since this would remove much of the other cellular proteins which also may be able to digest the reporter peptide (for example in the case of RA control, which was more active than the background in the previous assay). Purification would enable us to normalize protein concentration. Expression SDS-PAGE also confirmed that a big part of the overexpressed protein is soluble, so we decided to use the same expression conditions.

Night culture was also inoculated for protein expression in 30 ml flasks for small-scale purification.

2021-09-23 JJ&LC **Sanger sequencing results**

Sanger sequencing revealed that LytE gene has a frameshift mutation.



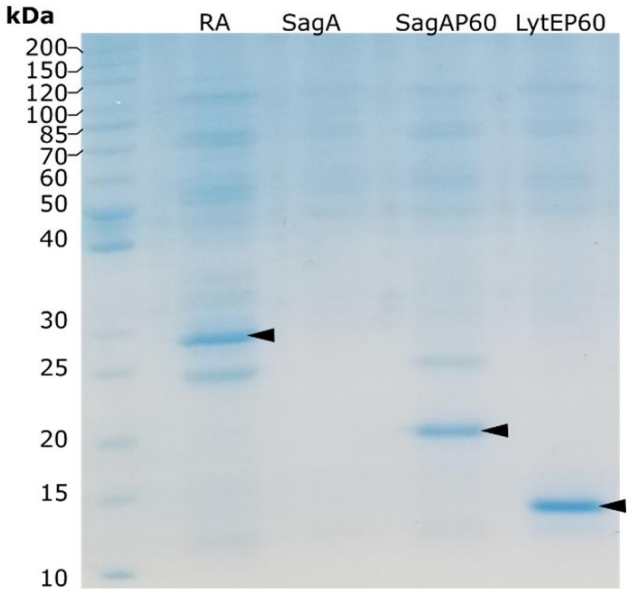
This explained why LytE was not expressed. All other sequences were correct. LytE will not be used for further experiments at this time.

2021-09-23 JJ **Protein expression and purification**

In the morning 1 ml of night cultures were added to the 30 ml of fresh LB medium supplemented with antibiotic and incubated until OD_{600} reached 0.5. Cells were induced, later collected, and purified by following the small scale purification protocol. Protein samples were run on SDS-PAGE.

Protein concentrations after dialysis:

Sample ID	User ID	Date	Time	mg/ml	260/280	A280 10mm	E 1%	Ext. Coeff x10e3	Mol. Wt. kDa	Cursor Pos.	Cursor abs.
SagA	Default	2021-09-24	08:03	3.55	1.47	3.555	10.00	NaN	NaN	280	3.555
SagA_P60	Default	2021-09-24	08:04	6.39	1.20	6.386	10.00	NaN	NaN	280	6.386
SagA_P60	Default	2021-09-24	08:05	6.85	1.27	6.851	10.00	NaN	NaN	280	6.851
RA	Default	2021-09-24	08:06	5.99	1.16	5.986	10.00	NaN	NaN	280	5.986
lytE_P60	Default	2021-09-24	08:08	6.83	1.30	6.829	10.00	NaN	NaN	280	6.829

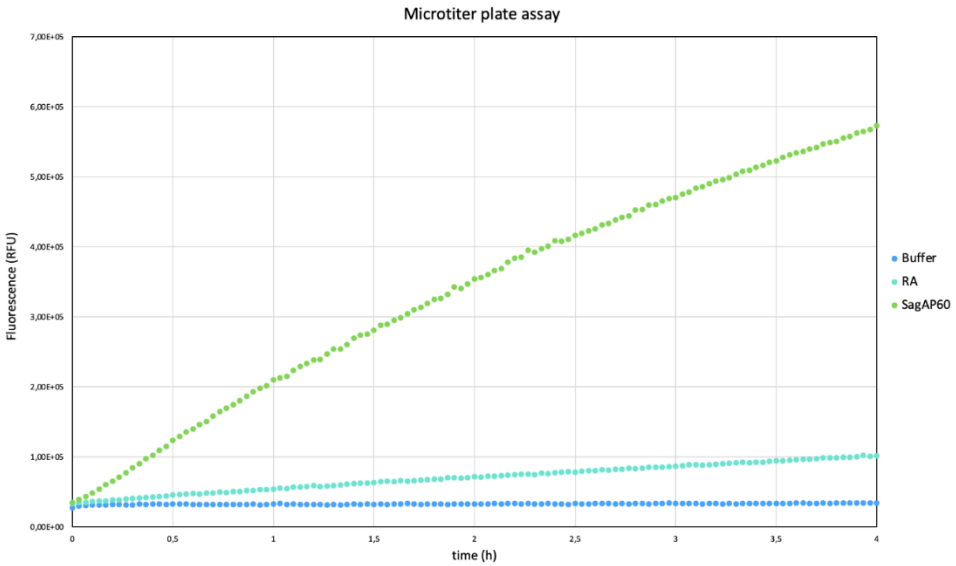


SagA was not purified using this method. This could be explained by unstable protein or handling errors. Concentrations of target proteins were measured spectrophotometrically, purity was estimated using ImageJ.

We will use only RA as control and SagAP60 for further experiments. This is mainly because SagAP60 showed the highest activity in the microtiter relate experiment with the lysates.R A and SagAP60 lanes look similar, also LytE gene sequence is not correct and full-length SagA was not purified successfully.

2021-09-23 LC **Microtiter plate assay**

Cells were sonicated as described in „Protein expression“ and „Protein purification“ protocols. Lysates were centrifuged for 1 hour on a tabletop centrifuge at 13.2 rpm. Reactions for microtiter plate assay were prepared as described in the protocol and 1/20 reaction volume of lysate was used for the reaction. In this assay, the same concentration of RA and SagAP60 proteins were used.

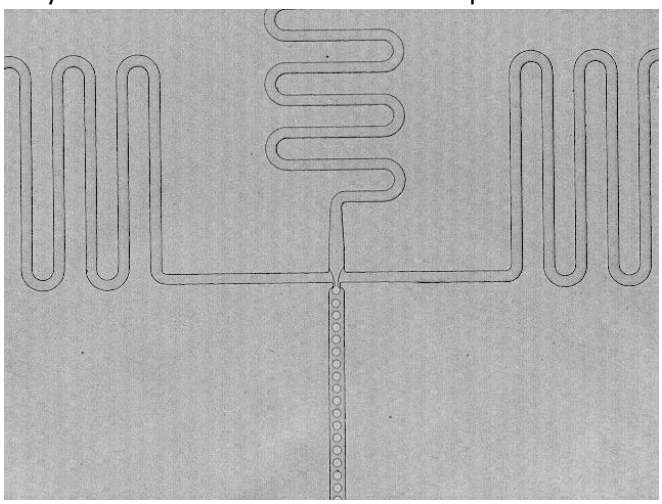


S. aureus night culture is inoculated in a sealed 15 tube with 4 ml of LB medium for tomorrow's experiment with live cells.

2021-09-25 LC **Endopeptidase assay in droplets**

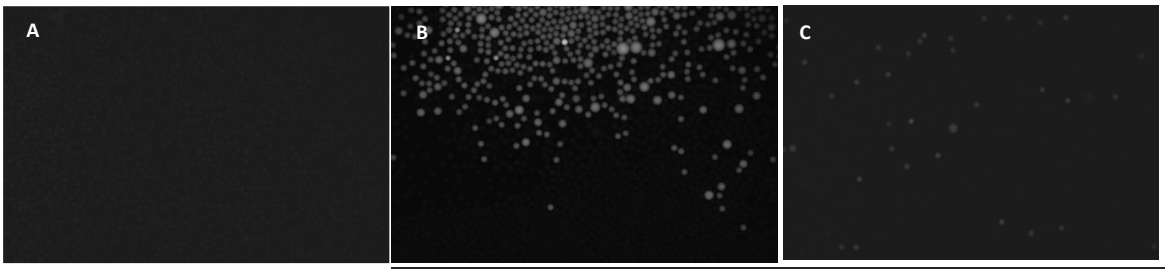
Following the test in a microtiter plate, a solution of assay buffer (HEPES 25 mM, NaCl 100 mM, substrate 25 µM, pH 7.5) containing the enzyme and a control solution of assay buffer without the enzyme were encapsulated to verify the functionality of the enzyme assay in buffer and imaged after overnight incubation at 37°C.

Enzyme and substrate are mixed in droplets:



2021-09-27 LC **Endopeptidase assay in droplets results**

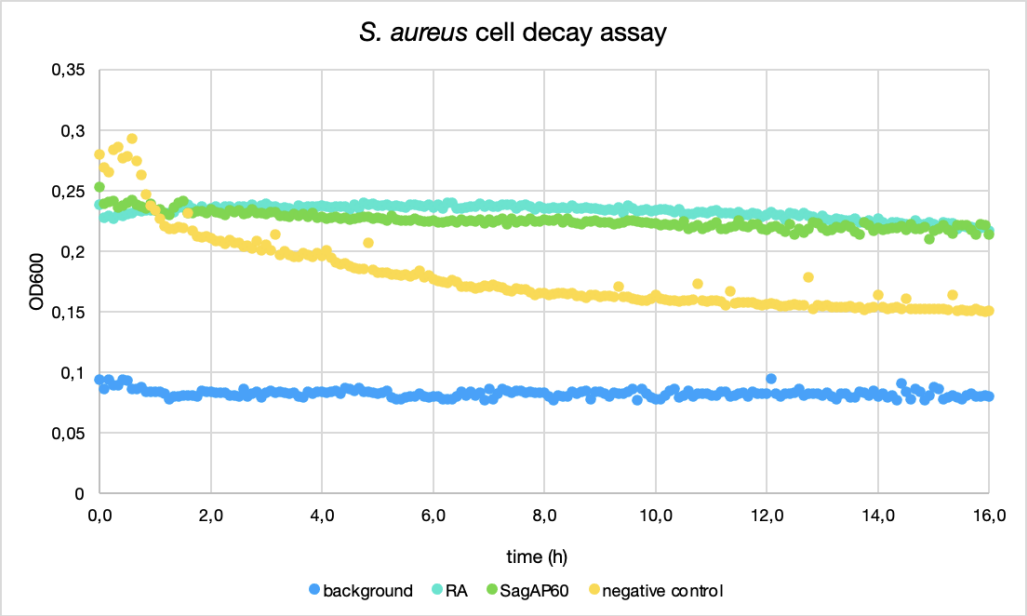
To verify that the long-term retention of the fluorescent signal in the positive droplets, 10 µL aliquots from both emulsions were mixed and imaged after incubation at 37°C for 2 days. No appreciable fluorescein exchange between positive and negative droplets was detected.

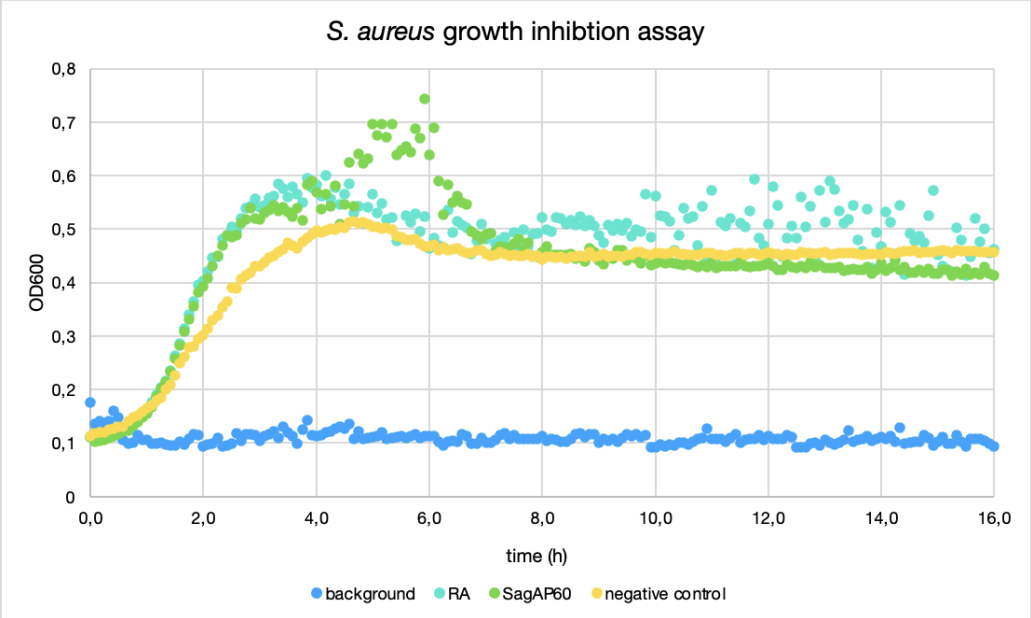


Assay in droplet. A – positive sample after encapsulation (time 0); B – positive emulsion containing the enzyme (top) next to control emulsion (bottom); C – mixed emulsion of positive and control sample: no leakage is observed after 2 days incubation.

2021-09-27 LC&JJ **Lysis of *S. aureus* cells**

Purified SagAP60 and RA control were used on a suspension of *S. aureus* to test the starting antimicrobial activity. The test was carried out as cell decay assay, by observing the decay over time of OD measured at 600 nm (OD_{600}) of cells suspended in buffer, and as cell growth assay, by observing the increase of OD_{600} of cells suspended in LB medium.
From the results it appears that the added proteins have an apparent positive effect on growth compared to negative control: we believe this is because they were added to the assay as glycerol stocks, providing an additional carbon source we haven't accounted for, with a 4% glycerol final concentration in the well. For this reason, this test will be repeated using a protein sample without glycerol.





To repeat this assay, we are going to purify a larger amount of RA (control) and SagAP60 proteins and do not use glycerol in the protein storage buffer. This time we are going to use a chromatographer. We expect a higher protein yield and purity.

2021-10-06 LC **Night culture for large scale purification**

E. coli BL-21 harboring RA and SagAP60 proteins expressing plasmids were used to inoculate 30 ml of night cultures.

2021-10-06 JJ **Protein expression**

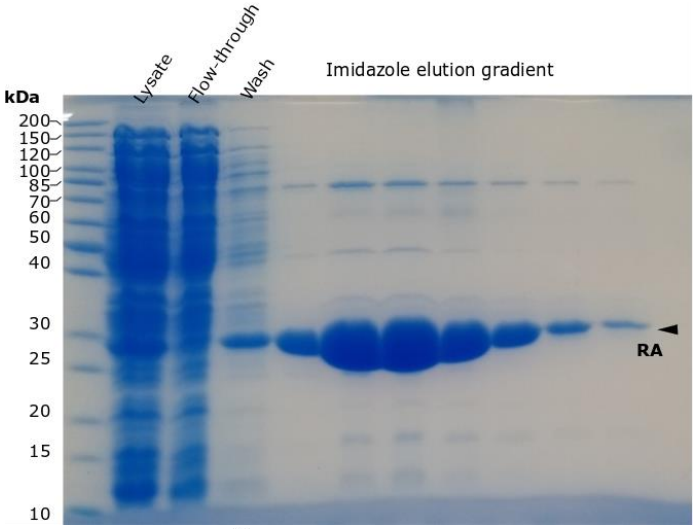
5 ml of night cultures were used to inoculate 400 ml of LB medium in 2 l flasks. For each protein, 6 flasks were inoculated in total. Biomass was grown according to the protocol. After induction bacteria were harvested by centrifugation and frozen at -20 °C.

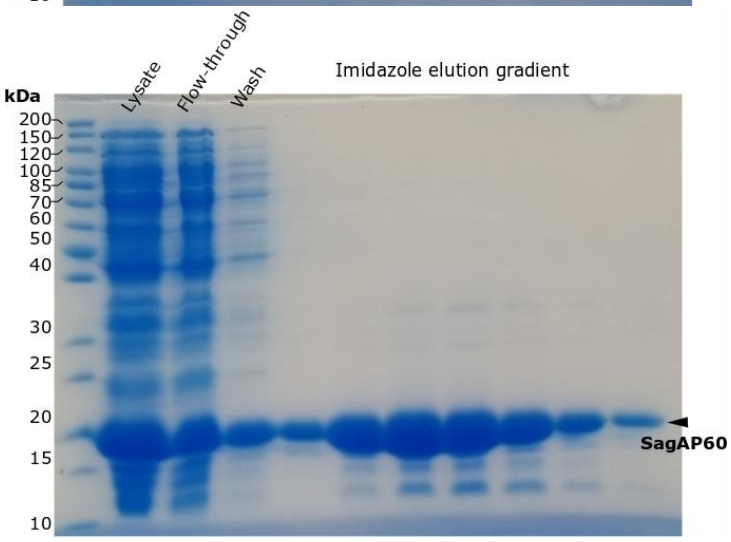
2021-10-10 JJ **Protein purification**

Proteins were purified according to the protocol for large-scale purification.

2021-10-11 JJ **SDS-PAGE of elution fractions**

Elution fractions of Ni ion affinity chromatography were run on SDS-PAGE. Protein purity is ~95 %. Proteins can be used for repeating lysis experiments of *S. aureus* cells.



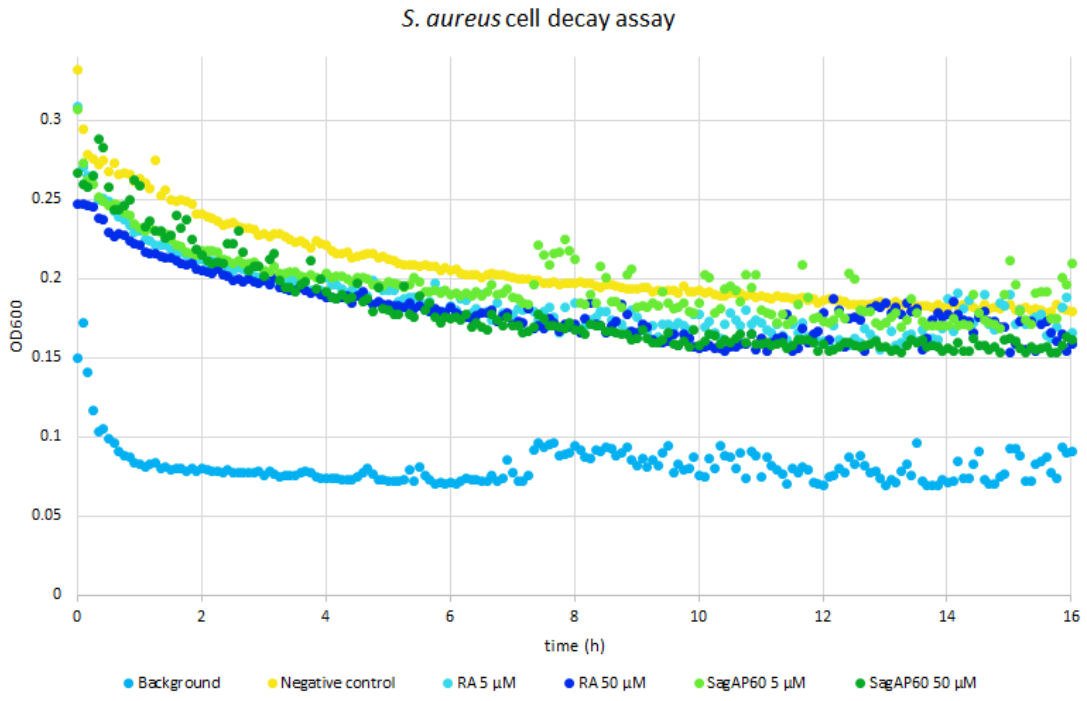


S. aureus night culture is inoculated in a sealed 15 tube with 4 ml of LB medium for tomorrow's experiment with live cells.

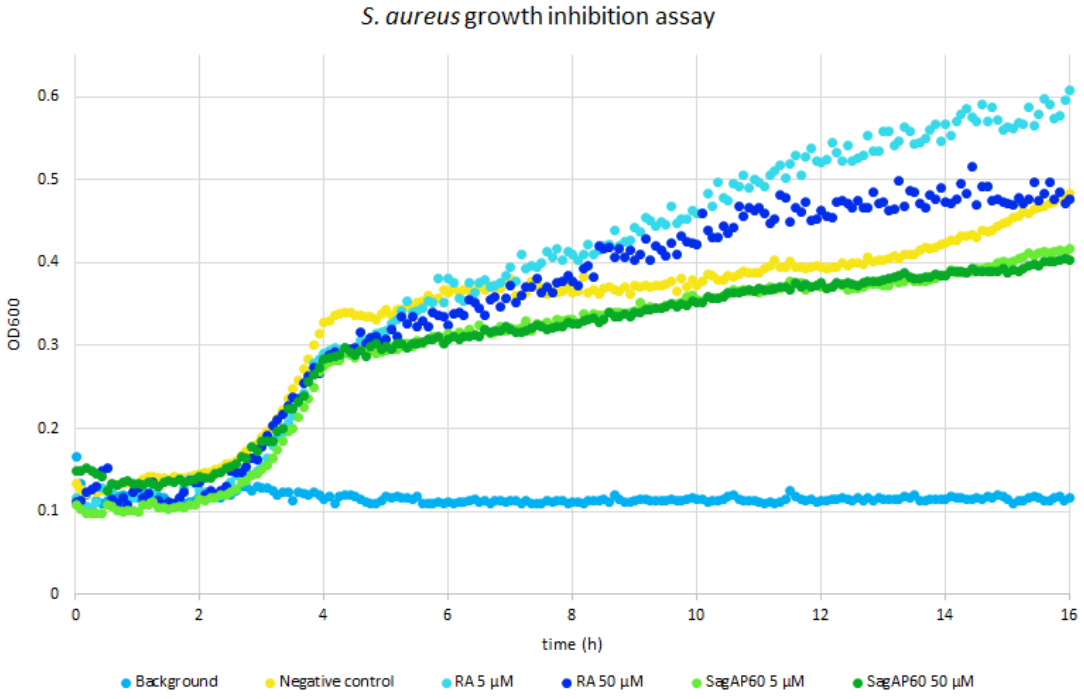
2021-10-11 JJ **Lysis of *S. aureus* cells**

Every condition is tested in 2 replicates, while the background and negative control (no enzymes added) was repeated 4 times. Averaged results are shown here. It is still difficult to observe the sagAP60 effect on *S. aureus* by cell decay assay. In all cases OD600 is decaying, also the error of the instrument used sometimes is bigger than the difference between conditions tested. Since we were

able to purify a larger amount of proteins we also tested a 50 µM concentration of proteins in the assay.



S. aureus growth inhibition assay seems to be a better method as cell growth is different and distinguishable between different conditions tested. In all 4 experiments where SagAP60 was added to the growing cells (both 5 and 50 µM) cell growth was lower than the negative control and RA control. However, the difference is too small. Also, cells, where RA was added, seem to grow faster. It is unclear why this is observed. Could be that RA is impacting cell growth positively in some unknown mechanism. At this point, it is clear that for evaluating lytic activity on live cells an active site mutant control will be needed. Also, it is apparent, that SagAP60 does not lyse *S. aureus* cell wall efficiently. This is probably explained by the low activity of enzymes which may be increased by directed evolution.


S. aureus growth inhibition assay
OD600
0.6
0.5
0.4
0.3
0.2
0.1
0
0
2
4
6
8
10
12
14
16
time (h)
Background
Negative control
RA 5 µM
RA 50 µM
SagAP60 5 µM
SagAP60 50 µM