UPPSALA
UNIVERSITET

Expression of FLAG-tagged argonautes in Dictyostelium Discoideum

Zozek Abdul Rahman

Degree project in biology, Master of science (2 years), 2022
Examensarbete i biologi 30 hp till masterexamen, 2022
Biology Education Centre and Department of Cell and Molecular Biology, Uppsala University
Supervisors: Fredrik Söderbom and Bart Edelbroek
External opponent: Jonas Kjellin

Abstract

Argonautes are conserved RNA-binding proteins that can regulate gene expression post transcriptionally through a process known as RNA interference (RNAi). This is done through the use of small RNAs, e.g. sRNAs that act as a guide for the argonautes, allowing for sequence-specific binding to the target site. This interaction has been studied in many organisms, one of which is the model organism *Dictyostelium discoideum*. *D. discoideum* is an amoeba that has been used extensively in genetic experiments due to its unique lifestyle, and ease of use. Being a eukaryotic, unicellular organism, it proves to be a great tool for the study of regulatory systems in eukaryotes, allowing us to study this argonaute-sRNA interaction in detail. By analysing which RNAs bind to the argonautes, we can better understand which genes these proteins regulate and what role RNAi has in the organisms as a whole.

In this study, I investigate three of the five argonautes found in *D. discoideum*, namely *agnA*, *agnC* and *agnE*. By transforming FLAG-tagged versions of these genes into the amoeba, I successfully express two of these modified proteins in *D. discoideum* and verified expression by using antibodies designed specifically to recognise the FLAG-tags. This opens up the possibility for the characterisation of the argonaute proteins to better understand their role and function in the regulation of genes.

1. Introduction **4**
1.1 Gene Regulation 4
1.2 RNAi 4
1.3 Argonaute Proteins 4
1.4 Dictyostelium discoideum 5
1.5 Characterization of argonautes 5
1.6 Aims of this work 6

2. Results **7**
2.1 Plasmid design 7
2.2 Transforming XL10-gold 7
2.4 Transforming D. discoideum cells 16
2.5 Western blot 18

3. Discussion **19**
3.1 The argonaute constructs 19
3.2 Transformation into Dictyostelium cells 19
3.3 Expression of FLAG-tagged proteins 20
3.4 Things left undone 20

4. Materials & Methods **21**
4.1 Argonaute Plasmids 21
4.2 Transformation of E. coli 21
4.3 Bacterial cultures 22
4.4 Plasmid purification 22
4.5 Digestion 22
4.6 Fragment purification 23
4.7 Ligation 23
4.8 PCR 23
4.9 Gel electrophoresis 23
4.10 Sanger sequencing 23
4.11 D. discoideum strains 23
4.12 D. discoideum cultures 24
4.13 SorMC buffer 24
4.14 SM agar 24
4.15 HL5-C medium 24
4.16 Transformation of D. discoideum 24
4.17 Counting of cells 25
4.18 Western Blot 25
4.19 FLAG-IP buffer 25

Acknowledgements **26**

References **26**

Supplementary material **27**

1. Introduction

1.1 Gene Regulation

All living organisms, even single celled organisms, have many functions and traits that can change in activity very quickly due to factors such as changes in the environment (Dolinoy 2007) or starvation (Sharma & Chatterji 2010) and this change is the result of the regulation of gene expression. Gene regulation is very important for the communication between cells (Asghar *et al.* 2012) but also within a single cell (Rosenfeld *et al.* 2005).

The process of gene regulation can occur at several points, with the most common being transcriptionally through the binding of transcription factors to the DNA to enhance or repress transcription (Hobert 2008) or post-transcriptionally through the use of regulatory RNA-binding proteins (Glisovic *et al.* 2008). More recently, a lot of focus has shifted to RNAs, specifically, small RNAs (sRNAs) and microRNAs (miRNAs) which are now thought to be a major factor contributing to the regulation of gene expression in a process called RNA interference (RNAi) (Valinezhad Orang *et al.* 2014).

1.2 RNAi

RNAi is a complex process, in which a double-stranded miRNA or siRNA molecule binds in with a multiprotein complex called the RNA-induced silencing complex (RISC) where one strand is discarded (in the case of miRNA) or cleaved (in the case of sRNA) and the other strand (guide strand) used for sequence-specific- (sRNA) or incomplete complementary (miRNA) binding to the target mRNA, inducing cleavage of it (sRNA & miRNA), or inhibiting translation leading to degradation of the mRNA (miRNA only), and thus causing a silencing effect (Grishok 2005, Engels & Hutvagner 2006, Lam *et al.* 2015). This binding is very specific, due to the often perfect complementarity to the target sequence (sRNA). The RISC complex contains several proteins, one of which is the argonaute protein.

1.3 Argonaute Proteins

The argonaute protein is a key player in the RISC and is the protein responsible for binding the single-stranded sRNA or miRNA and using it as a guide to finding the target site (Grishok 2005, Lam *et al.* 2015). Argonautes, while there are some differences between them, are highly conserved within most organisms when considering their structure (Wu *et al.* 2020). The majority of PIWI domain containing proteins fall in the Ago superfamily of proteins, due to their shared PIWI domain, although there are still a lot of things separating the different argonautes from each other as well. (Cerutti & Casas-Mollano 2006, Swarts *et al.* 2014). Argonautes have been shown to affect the development of several organisms such as plants, where mutation of *ago1*, which is expressed throughout development, results in plant architectural deformities as it develops (1998 p. 1, Lynn *et al.* 1999). Similarly, argonaute

mutants have also been shown to affect the development of animals, such as in *Drosophila melanogaster*, where mutations in the argonaute gene *aubergine* is required for posterior body patterning (Wilson *et al.* 1996). Another eukaryote in which Argonautes are conserved is the social amoeba *Dictyostelium discoideum*. Within *D. discoideum*, 5 different argonaute proteins have been identified: AgnA, AgnB, AgnC, AgnD and AgnE (Cerutti & Casas-Mollano 2006, Boesler *et al.* 2014), however, the gene coding for AgnD is thought to be a pseudogene.

1.4 *Dictyostelium discoideum*

D. discoideum is an amoeba belonging to a large group of "social amoeba", a name reflecting their developmental lifecycle (Annesley & Fisher 2009), where they live as unicellular organisms, but can at later stages, due to starvation or other stress factors, aggregate together and adopt a multicellular lifestyle and crawl around looking for a better place to live or develop (dubbed the slug phase) and eventually forming a fruiting body, where the cells at the top form spores while the cells further down die in favour of the survival of the species (Loomis 2012).

This complex lifecycle in a eukaryote where the cells undergo many changes relatively quickly requires a lot of gene regulation which is the reason, together with their ease of use for biochemical analysis, why *D. discoideum* has become a model organism for such studies, and similarly, why it is of interest in this study.

1.5 Characterization of argonautes

As RNAs bind to the argonaute proteins during the gene regulatory process of RNAi, we can investigate and better understand which genes each of the argonaute proteins potentially regulates. While we know that there are multiple argonaute proteins, with each of these proteins potentially being responsible for the regulation of different genes, we can get a better picture of the whole RNAi process by characterising each of these proteins. Due to this, in combination with the fact that argonaute proteins are highly conserved between organisms, it is of high interest to learn more about how exactly these proteins work, and what targets they have in the cells. While argonautes are important for RNAi, it is not very clear as to how exactly this affects the development of *D. discoideum*, however, it has been observed that *agnE* is required for proper development, while *agnB* and *agnC* control cell growth through the regulation of transposon mobilization (Liao *et al.* 2018b). The same group also observed that depletion of AgnB and AgnC results in slower growth through cell division, while the deletion of *agnE* results in faster growth (Liao *et al.* 2018a)**.**

1.6 Aims of this work

To further investigate the functions of argonautes in *D. discoideum*, in this study, my aim was to transform and express a plasmid in *D. discoideum* cells that can be used to identify its expression. This plasmid could then be used to identify whatever RNA has managed to bind to the expressed protein through the use of a 3xFLAG-tag, thus being able to pull these proteins down and characterise them. This is done in hopes of eventually being able to characterise exactly what targets each of the argonaute proteins found in the amoeba has, and what roles they may hold.

Out of the five argonaute genes, one, *agnD*, is thought to be a pseudogene and may not be expressed as a protein, while another, *agnB*, has already been cloned and expressed as a tagged version in the amoeba (Edelbroek B., unpublished). This leaves 3 argonautes, namely AgnA, AgnC and AgnE, which was my goal to express and identify.

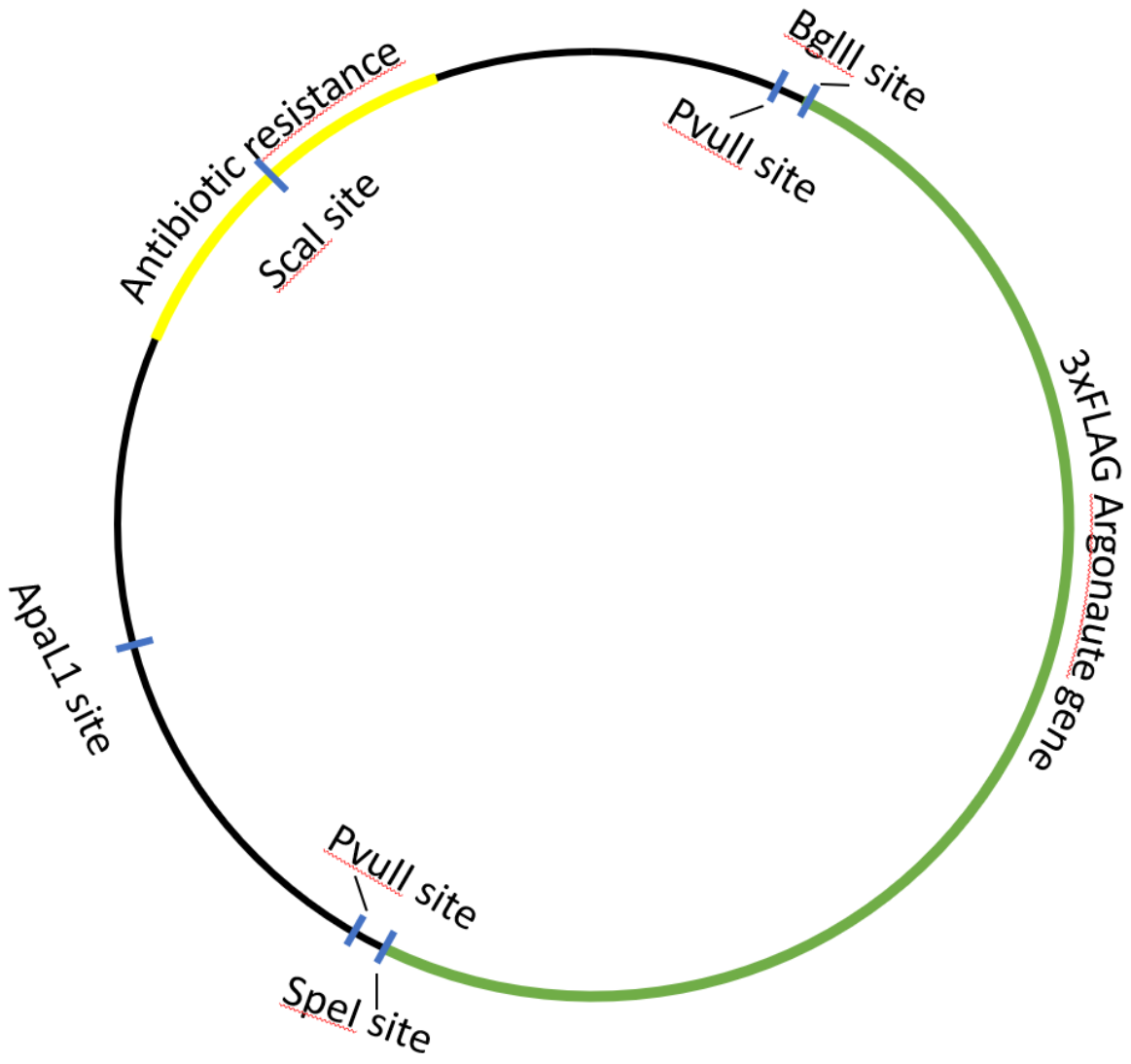


Figure 1: Schematic showing the plasmids containing the different argonaute genes. The restriction enzymes BglII and SpeI were positioned at each end of the argonaute genes (green) for all argonaute plasmids, and ScaI in middle of the antibiotic resistance (yellow) for AgnC and AgnE while ApaL1 was used for AgnA. PvuII was used initially for AgnC, but was later substituted for ScaI.

2. Results

2.1 Plasmid design

Prior to the start of this study, three plasmid constructs were designed and constructed, harbouring one each of the three argonaute genes (Fig. 1, Fig. S1-S3). These constructs have been designed in such a way that their introns have been removed, restriction sites added to each side flanking the gene and, most importantly, a 3x FLAG-tag sequence added to the start of the gene of interest, resulting in a 3xFLAG-tag at the N-terminus of the final protein. Other than the gene itself, different antibiotic resistance genes are also part of the plasmid, as to allow us to select for the plasmids with the use of antibiotics. Among the three plasmid constructs, one has a kanamycin resistance gene (AgnA) while the other two contain ampicillin resistance genes (AgnC & AgnE).

A fourth plasmid is also used, namely the pDM304 vector (Veltman *et al.* 2009) (Fig. S4), which was used as a destination vector, as it contains the rest of the elements needed for successful replication and expression in the amoeba cells.

2.2 Transforming XL10-gold

Before we can express the FLAG-tagged argonaute genes in the desired species *D. discoideum*, we must first put them into the appropriate plasmid which contain additional elements that are necessary for the genes to be expressed correctly and can replicate within *D. discoideum*. The plasmid pDM304 has been chosen for this purpose and has been reported to work well for the expression of many different proteins (Veltman *et al.* 2009, Hiraoka *et al.* 2020). To do this, however, we need a much higher amount of each argonaute plasmid (AgnA, AgnC and AgnE) than what was ordered. We can easily do this by first transforming *Escherichia coli* bacteria with our desired argonaute plasmids (pDM304 was already transformed prior to this study) in order to amplify the plasmids.

For this purpose, the *E. coli* strains DH5α and XL10-Gold cells were used, and both strains were used to transform an argonaute plasmid. Two strains were used in case one strain would prove problematic for the transformation of our plasmids. To prevent un-transformed cells from growing on our plates, the plates contained antibiotics corresponding to the resistance in our desired plasmid (Ampicillin for all but argonaute A, which holds kanamycin resistance). The transformations were successful, and cells of each strain bearing each of our plasmids grew successfully.

To keep things consistent for our future experiments, some of these cells were re-streaked, so that we could conduct multiple experiments using the same, identical clones. For the purification of the plasmids, the concentrations attained tended to be very low and to circumvent this, many different plasmid purification conditions were tried. In the end, the plasmid purification was optimized by using much larger cultures. This meant, however, that the plasmid Miniprep Kits used for purification would no longer be able to sustain such large volumes. Hence, this larger culture was split into several smaller samples, which were purified in parallel. Furthermore, to avoid the previous problems of low concentration in high volume, the plasmids were eluted in a much lower volume at the end. This workaround worked very well, and the total yields were increased four- to six-fold.

To make sure that we had the right plasmids in our cells and that the restriction enzymes work, we cleaved the plasmids using BglII and SpeI restriction enzymes and performed gel electrophoresis to check their sizes (Fig. 2). Due to using two enzymes to cleave, we expect fragments of sizes 3027 bp for AgnA, 3714 bp for AgnC and 3689 bp for AgnE (bands marked on the figure). We also expect a band corresponding to the rest of the plasmids with sizes 2 278 bp for AgnA and 2341 bp for both AgnC and AgnE. It should be noted that BglII cleaves twice on the AgnA plasmid, causing one additional band to appear. We also see a larger band in all lanes, which is due to uncleaved plasmid still present in the samples.

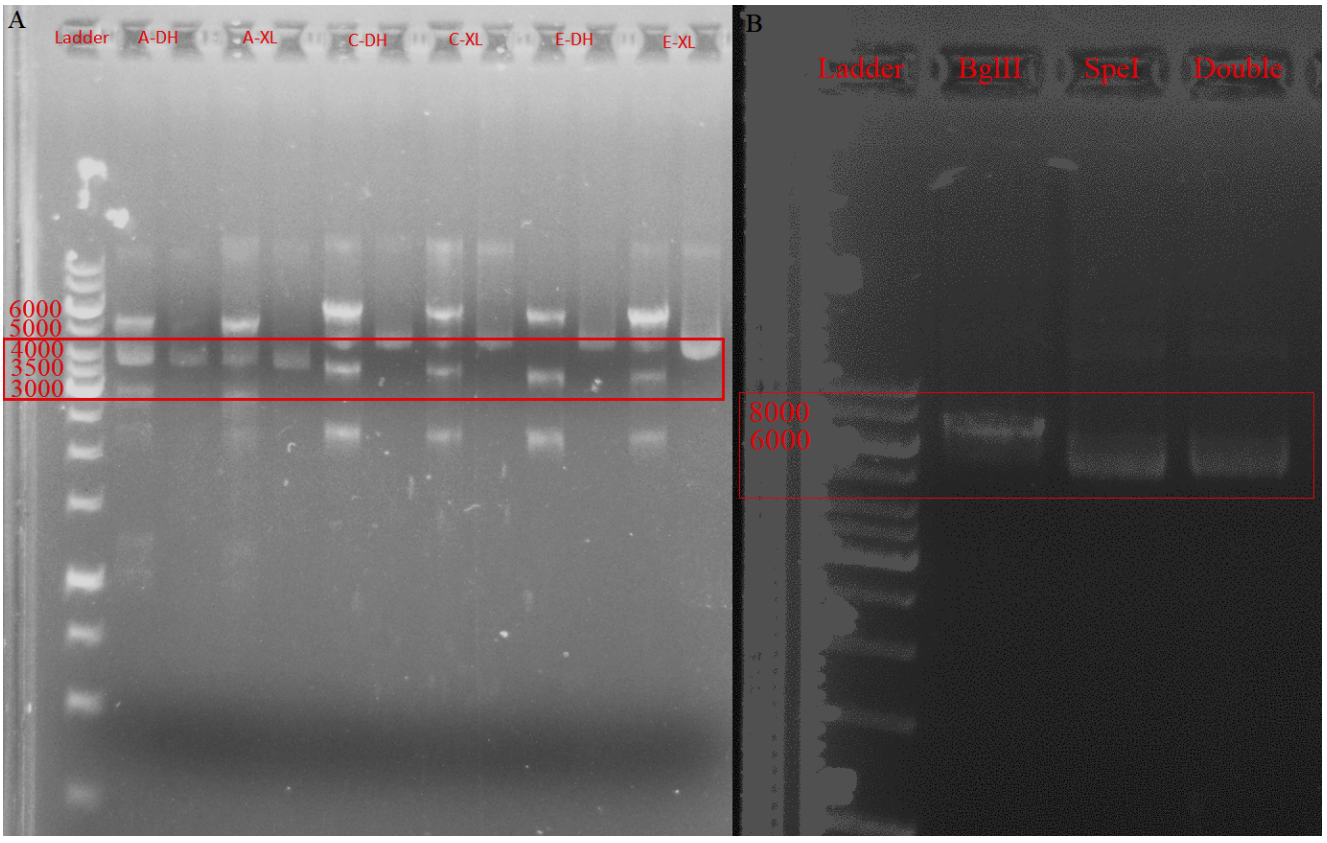


Figure 2: Agarose gel showing single and double digests of each plasmid, including pDM304. A-DH: AgnA plasmids grown in DH5α cells. A-XL: AgnA plasmids grown in XL10-Gold cells. C-DH: AgnC plasmids grown in DH5α cells. C-XL: AgnC plasmids grown in XL10-Gold cells. E-DH: AgnE plasmids grown in DH5α cells. E-XL: AgnE plasmids grown in XL10-Gold cells. **A.** Double digests for each plasmid in each strain using

BglII and SpeI restriction enzymes. Every digest is followed by a negative control, containing an uncleaved plasmid of the same type. **B.** Single and double digests of pDM304 destination vector using BglII and SpeI restriction enzymes.

2.3 Creating the argonaute construct

Now that we had pure plasmids to work with, we wanted to move the argonaute genes over to our destination vector pDM304. To do this, we used restriction enzymes to cut both plasmids to be able to then ligate the pieces using a ligase. Each of the argonaute genes had been designed to be easily cut out by using two restriction enzymes, namely BglII and SpeI. Due to the fact that the two cleavage sites used (for BglII and SpeI) are right next to each other on the pDM304 vector, it was impossible to deduce, from the gel electrophoresis, if one or both restriction enzymes had cleaved the plasmid. To avoid re-ligation of the plasmid when only one restriction enzyme had cleaved, the cleaved plasmid was treated with a phosphatase. This phosphatase dephosphorylates the DNA ends after digestion, preventing them from being ligated. Once the ligation had been finished, the plasmids were transformed into XL10-Gold cells. Only one strain of *E. coli* was used to reduce the workload.

Many transformants grew on each plate, and to make sure that we had the correct insert, selected colonies were re-streaked and from those, colony PCRs were performed to amplify the argonaute inserts and part of the pDM304 backbone. Subsequently, the samples were analyzed by gel electrophoresis to check for the amplified insert (Fig. 3).

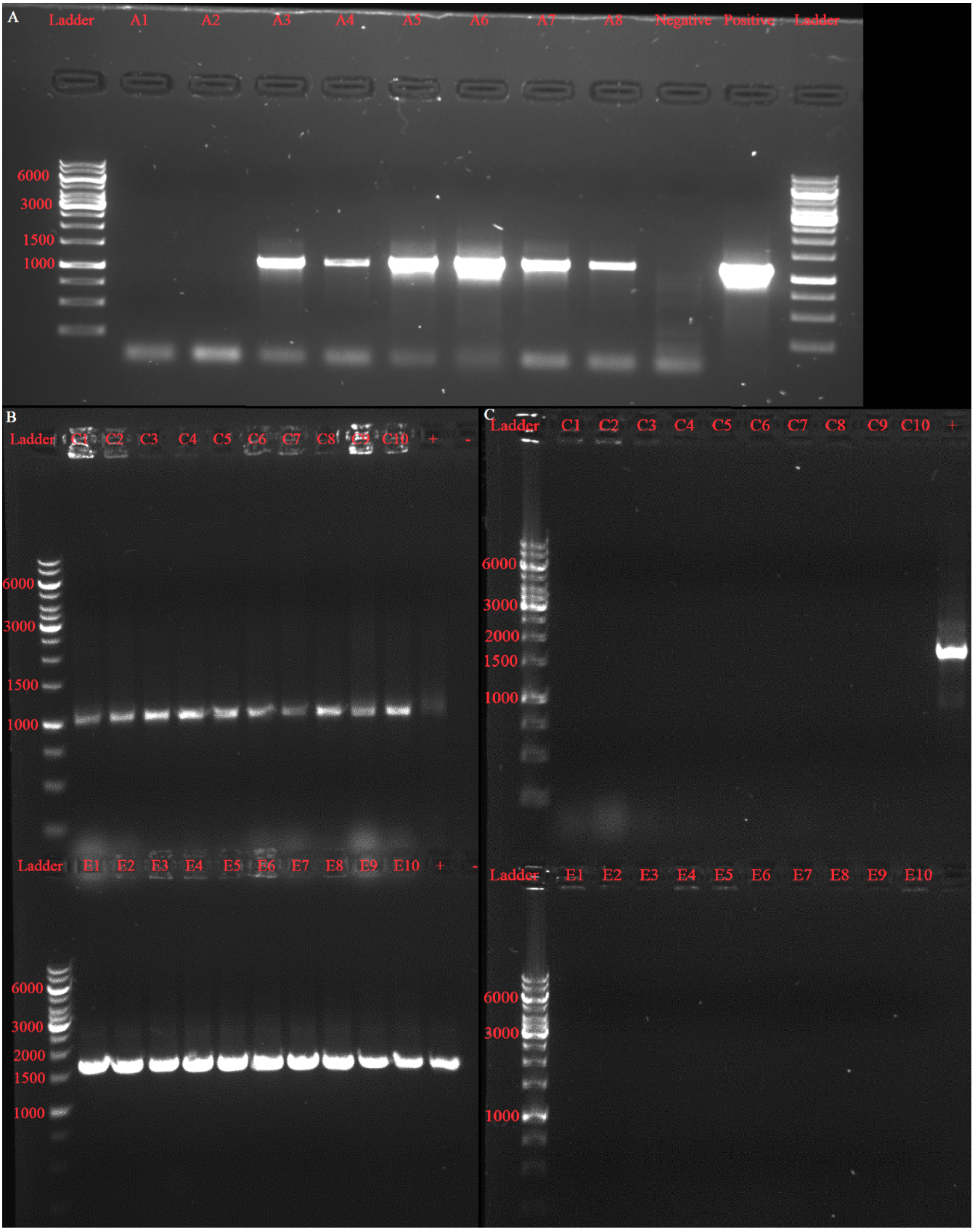


Figure 3: Gel electrophoresis showing the results of the colony PCR of XL10-Gold cells transformed with each ligated argonaute-pDM304 construct. **A.** PCR with primers checking for the *agnA* gene insert with an expected size of 1191 bp in 8 samples denoted with numbers 1-8 as well as a positive control, denoted with 'positive', and negative control, denoted with 'negative'. **B.** PCR with primers checking for the *agnC* (top) and *agnE* (bottom)

gene inserts with expected sizes of 1066 bp and 1786 bp respectively. 10 samples were used for each argonaute gene, denoted with numbers 1-10 as well as a positive control, denoted with '+' and negative control denoted with '-'. **C.** PCR with primers checking for the pDM304 backbone in both *agnC* (top) and *agnE* (bottom) plasmids with an expected size of 1730 bp. 10 samples were used for each argonaute gene, denoted with numbers 1-10 as well as a positive control, denoted with '+'.

AgnA showed positive results for the gene insert in 6 out of 8 of the samples (Fig. 3A), but unfortunately, while both AgnC and AgnE do show positive results for the gene insert (Fig. 3B), they do not show a positive result for the pDM304 backbone, as there are no bands visible on that gel (Fig. 3C), signifying that the gene is still sitting on the old plasmid in which they originated from and have not ligated into pDM304. PCR analysis for pDM304 was not performed for the AgnA for reasons explained below. The reason they could grow on antibiotic seeded plates was the fact that both the destination vector and the origin vector contain ampicillin antibiotic resistance genes (Fig. S2-S4), and thus we cannot select for only one of the plasmids by using ampicilin, as they would both be resistant to it. This is where the bulk of the problems with the project came in. This is not a problem that exists for AgnA since that plasmid has a different antibiotic resistance on the origin plasmid (KanR) (Fig. S1) compared to the destination vector (AmpR). Thus, the AgnA transformation succeeded, while the AgnC and AgnE transformations did not.

In an attempt to ensure that we get the ligations we wanted for the remaining two argonaute genes, an extra restriction enzyme was used to digest the origin plasmid, thus creating three fragments and decreasing the chances that it would self-ligate back (since only the destination vector pDM304 was treated with phosphatase), as that would require all three fragments to come together in the exact way that they were cut, which is unlikely. Normally, we would expect all of the plasmids to be digested in a digestion reaction, but since they were clearly not getting fully digested, having an additional enzyme cleaving the plasmids would also mean that more plasmids in total would get digested, and potentially reducing the number of intact plasmids. The restriction enzyme chosen was initially PvuII for AgnC and the method was tested for that plasmid only to start out (Fig. 4). For the BglII and SpeI single digests, we expect one band with the full plasmid length (6055 bp), while we expect an additional band for the PvuII digest due to there being two restriction sites for that enzyme (2 bands, one with ~4000 bp, and the other ~2000 bp). For the triple digest, we expect three bands (sizes 3712 bp, 1476 bp and 657 bp), even though the enzymes together would cleave five times. This is due to two of the fragments being so small that it would most likely pass through the gel before we stop it (sizes of 66-152 bp).

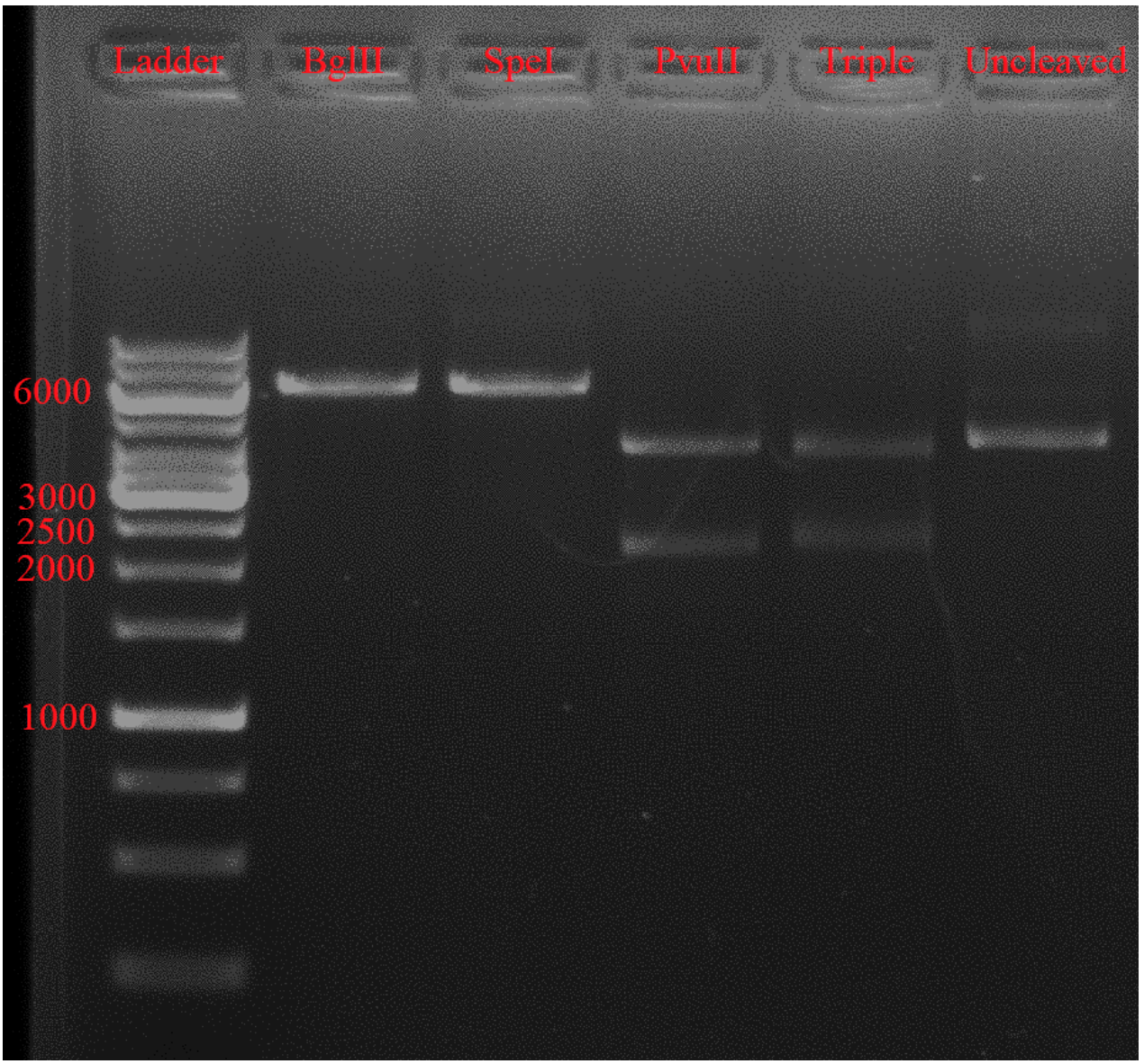


Figure 4: Restriction digest of *AgnC* plasmid using BglII, SpeI and PvuII restriction enzymes.

The single- and double digests looked as expected with the correct number of bands of the expected sizes. The triple digest looked good, with only one band missing from the triple digest, however, this may be due to the band being so small. Nonetheless, it was used in another attempt to ligate its argonaute gene into the pDM304 destination vector and transform it into competent XL10-Gold cells. The ligation was analyzed as before (se above). However, the results were still the same, with positive inserts but not positive for the backbone (data not shown).

It was theorised that the reason for negative colonies was the fact that the original argonaute plasmids might not have digested fully and hence undigested original plasmid was transformed. A workaround for that would be to allow the digestion reactions to run for much longer, up to an hour (instead of the 20 minutes incubation time used before). This should not be problematic, as the protocols for each restriction enzyme do state that digestions for up to 16 hours should work fine. To see if this could be the case, a digestion reaction was run for a

total of 1 hour, with aliquots loaded on a gel every 10 minutes (Fig. 5). This time, both AgnC and AgnE were digested, and to simplify the procedure, the choice for the third restriction enzyme was changed to ScaI for this reaction, since it cleaves both plasmids in the same ampicillin resistance gene (Fig. 1), leading to bands of good sizes that can be seen on the gel for both plasmids (compared to PvuII that only really cleaved AgnC plasmid well). Here we expect one full-size band for time point 0 (~6000 bp) and three bands for the rest, with each subsequent time point showing a clearer band (3700 bp (insert), 809 bp and 1542 bp).

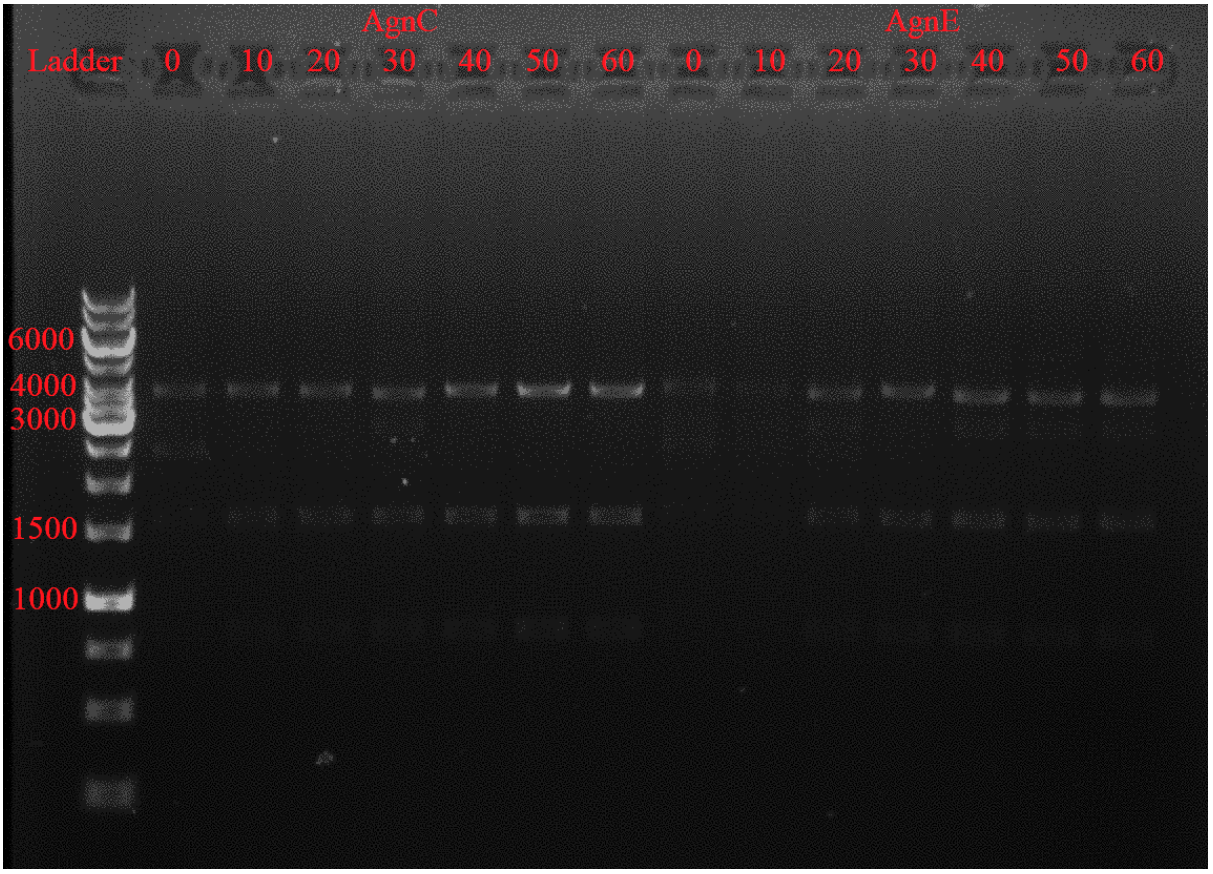


Figure 5: Digestion reaction over time using BglII, SpeI and ScaI restriction enzymes. Each well shows the result from the digestion reaction 10 minutes after the previous one. Left side shows the results for AgnC and the right side shows the results for AgnE.

Due to each of the bands getting brighter and more clear for each subsequent time point, it does seem that running the digestion reactions for an extended time does ensure that there are much fewer plasmids in the mixture (optimally free of plasmids), and might ensure that positive colonies show up after a transformation. We do, however, see multiple bands on time point 0, which was not expected. This may be due to the fact that the time point 0 was measured with the start of incubation, while the digestion reaction might have already started earlier when the mixture was made (reacting slowly). To further ensure that the mixtures would be plasmid-free, the digested bands corresponding to the insert (largest, top band) were cut out from the gel using a razor and purified, before being used in the ligation reaction (Fig. 6).

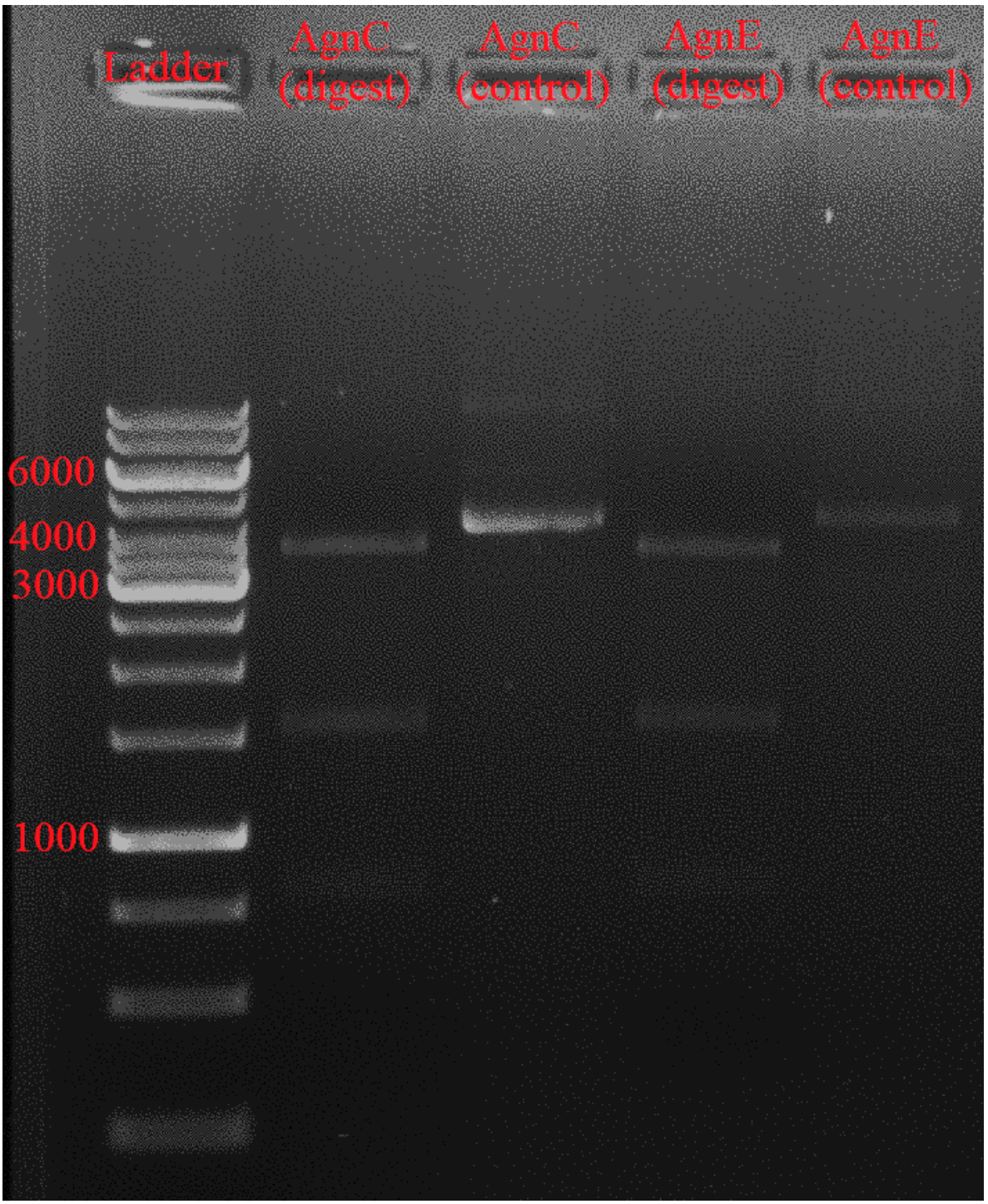


Figure 6: Triple digestion reactions for AgnC and AgnE with uncleaved controls. Restriction enzymes used were BglII, SpeI and ScaI for all triple digest reactions.

This time, there were barely any intact plasmid remnants visible on the gel, and a transformation was conducted in hopes that any small number of undigested plasmids that might still be present, would not interfere.

After ligation into the destination vector, pDM304, they were transformed into XL10-Gold cells. Each transformant, other than the negative controls had colonies growing (a total of 8 colonies for each argonaute transformant), where 4 of these colonies were initially randomly

picked and checked for the appropriate insert and backbone (Fig. 7A) and, after observing these to be negative for the backbone, the remaining 4 colonies were grown in cultures, their plasmids purified and then digested to check the sizes of the plasmids (Fig. 7B) by digesting them with one restriction enzyme only (in this case BglII) and analyzing them by gel electrophoresis. The reason why the size of the plasmid was analyzed the second time was due to the fact that a successfully ligated plasmid would be significantly larger than the un-ligated plasmid (~10 500 bp as opposed to ~6 000 bp). As the restriction digest is significantly faster than the PCR reaction, we could quickly check if the rest of the colonies were negative before running the proper PCR reaction. For the size check, the expected size would thus be the full pDM304 plasmid size including the argonaute inserts (~10 500 bp for both AgnC and AgnE).

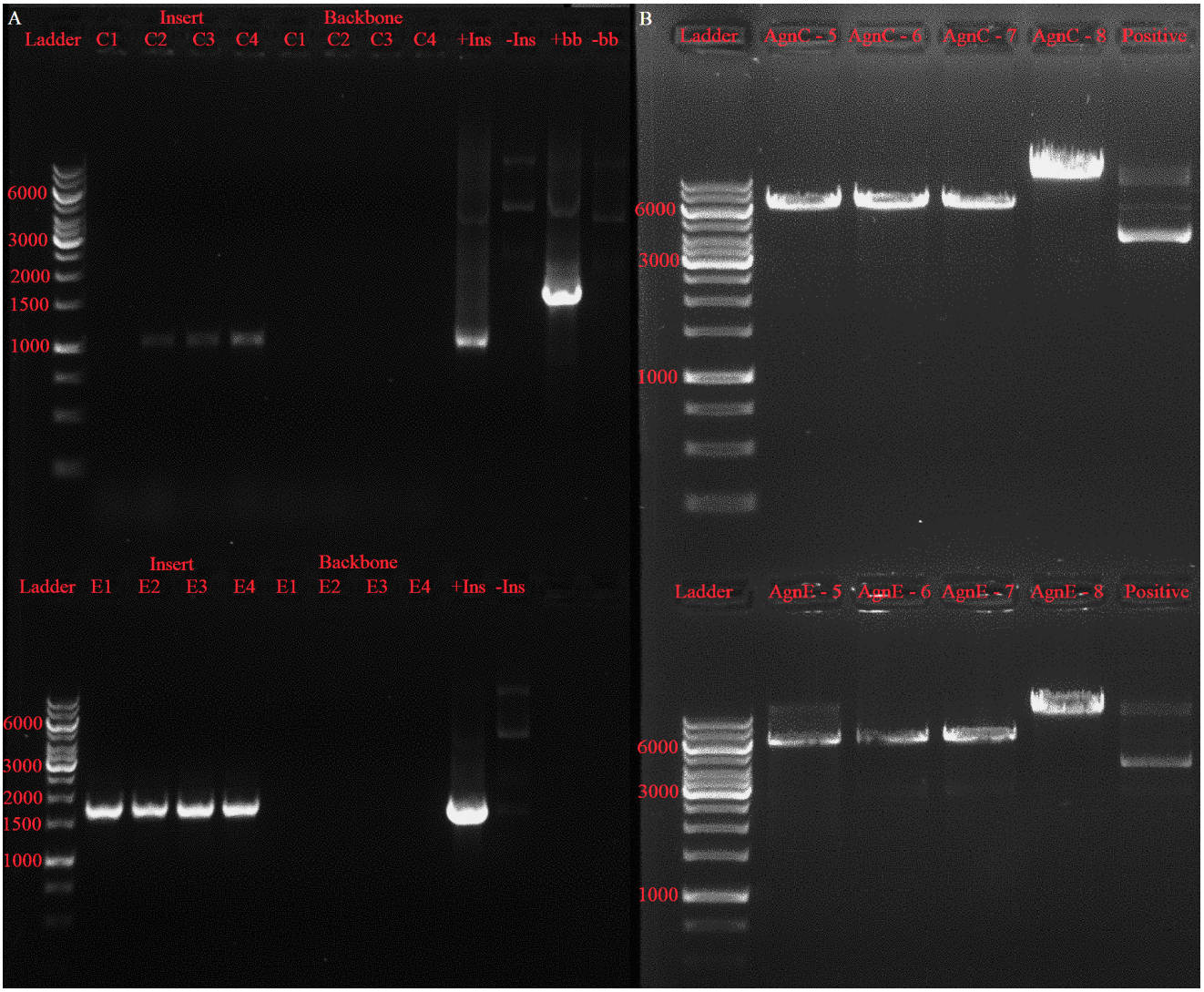


Figure 7: A. PCR reaction checking for AgnC and AgnE insert and backbone, using primers with an expected size of 1066 bp for AgnC insert and 1789 bp for AgnE insert and 1730 bp for pDM304 backbone. 4 colonies were used for each argonaute gene denoted by numbers 1-4. **B.** Digestion reaction of purified plasmid DNA from XL10-Gold transformants only using BglII restriction enzyme. 4 colonies were used for each argonaute, denoted by numbers 5-8 along with positive control, denoted by 'positive'.

Despite the longer digestion incubation time, most cells were negative, except for one colony from each argonaute (AgnC-8 and AgnE-8), by chance (Fig. 7B). It appears that the hour-long digestion was still not enough to fully digest all plasmids in the mixture and a longer incubation time is required, spanning over multiple hours to get a consistent positive transformation. The gel purification, although most likely helping, also did not guarantee that there would be no undigested plasmid. This might be because of the fact that supercoiled plasmids might have migrated down further by chance, following the gel piece that was cut. However, the two positive colonies were used in subsequent experiments. Before continuing, however, a colony PCR was performed, checking for both insert (argonaute gene) and backbone (pDM304) using the appropriate oligos (STable 1), to ensure that those two colonies truly were positive (Fig. 8). The result of this gel showed faint bands for AgnC both for insert and backbone (Fig. 8 lanes 2 & 3), but not as clear for AgnE backbone. However, this might be due to the lower concentration of plasmids used, and since we clearly had the right size (Fig. 7B), we decided to run with it for the transformation of *D. discoideum*.

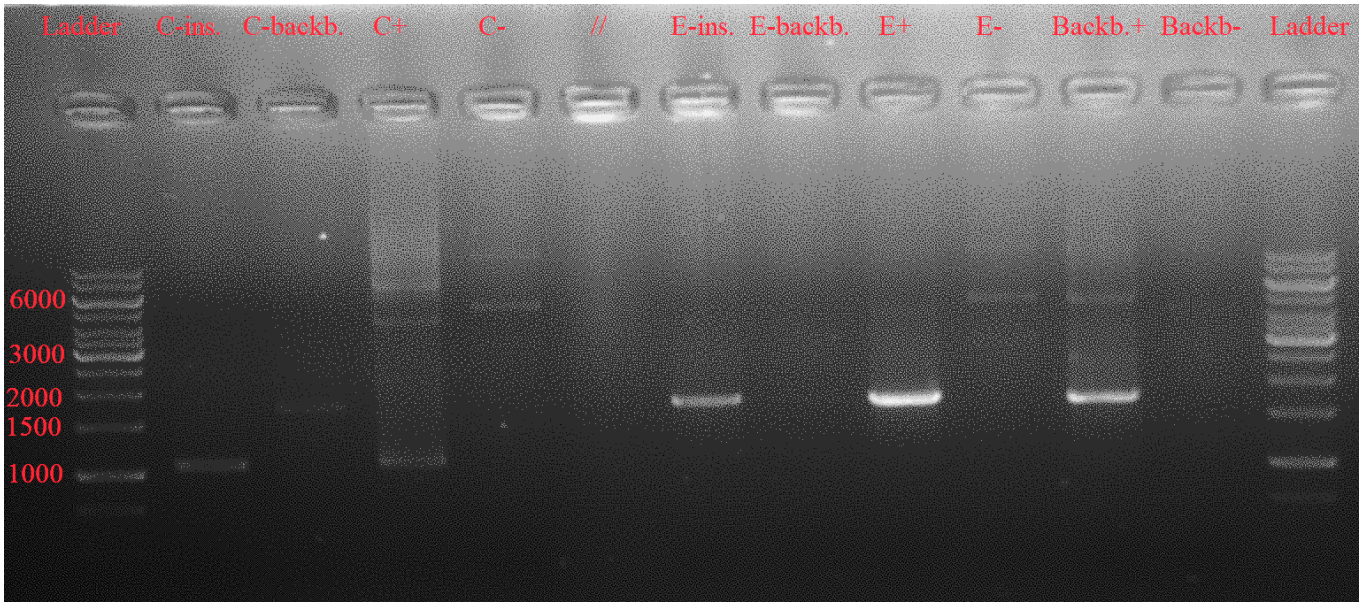


Figure 8: PCR reaction with primers checking for AgnC and AgnE inserts (C-ins. & E-ins.) with expected sizes of 1066 bp and 1786 bp respectively. Primers were also used to check for pDM304 backbone (C-backb. & E-backb.) with an expected size of 1730 bp. C-ins: AgnC checking for gene insert. C-backb.: AgnC checking for pDM304 backbone. C+: AgnC positive control. C-: AgnC negative control. E-ins: AgnE checking for gene insert. E-backb.: AgnE checking for pDM304 backbone. E+: AgnE positive control. E-: AgnE negative control. Backb.+: pDM304 backbone positive control. Backb.-: pDM304 backbone negative control.

2.4 Transforming *D. discoideum* cells

For each of the positive colonies, a culture was made and their plasmids were purified. With the correct plasmids at hand, it was time to transform these into *D. discoideum* cells. To start, only AgnA and AgnC were electroporated into *D. discoideum* cells, as to decrease the workload, and these were then allowed to grow in *Klebsiella* - SorMC suspension with G418 antibiotics. To check whether the plasmid had transformed successfully, a colony PCR was performed on the growing cells, for both insert (argonaute gene) (Fig. 9A & 9C) and

backbone (pDM304) (Fig. 9B & 9D) with two concentrations, one with low concentration and another with high concentration. This resulted in both AgnA colonies (A1 & A2) being positive for both colonies, visible only for the higher concentrations (Fig. 9A). For the AgnC colonies, only the wt (C2) was positive for both insert and backbone (Fig. 9C & 9D). Thus, the absolute majority do seem to be positive, except for the AgnC knockout strain, in which a PCR checking for 3 additional colonies was performed (Fig. 9E).

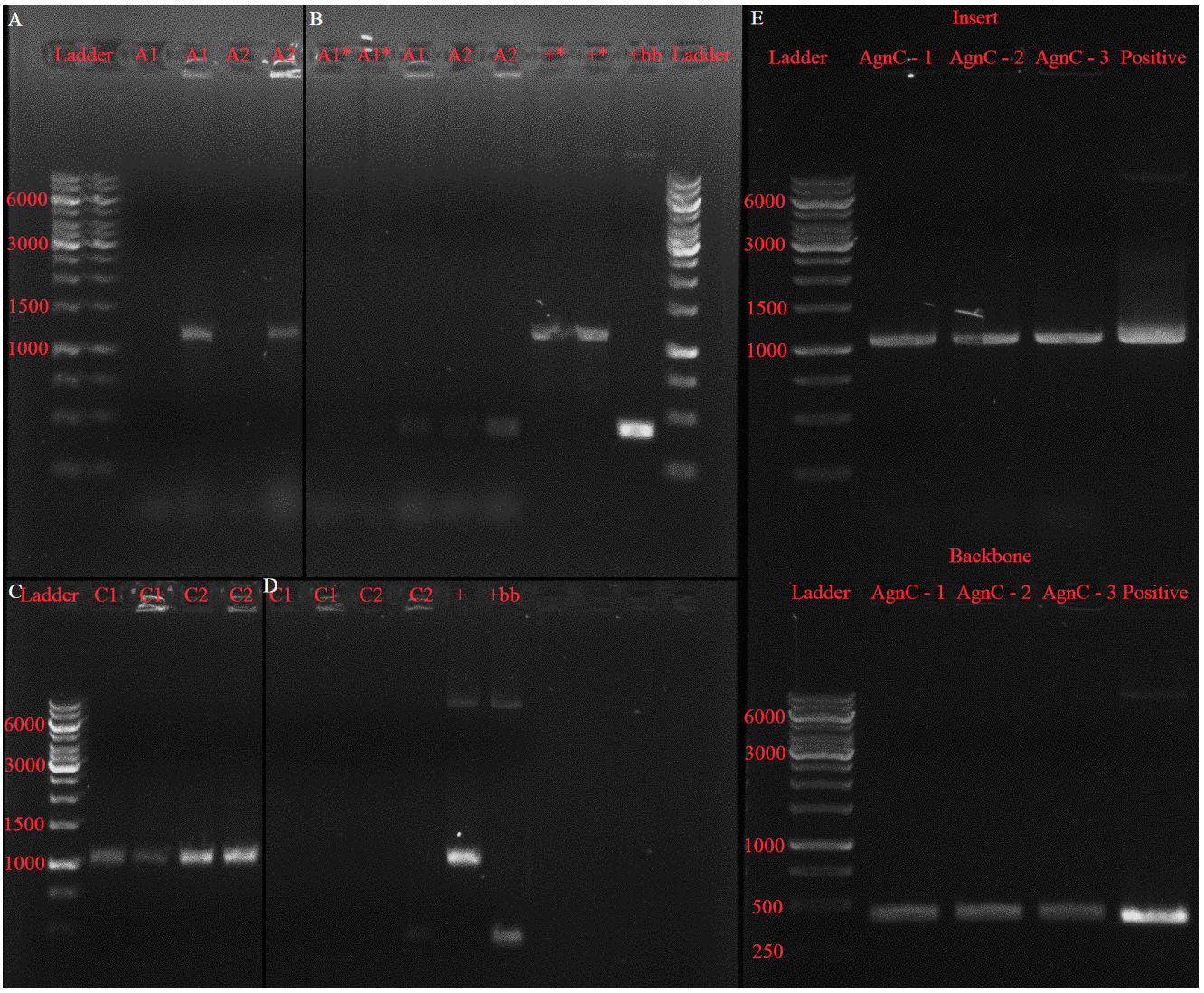


Figure 9: A. PCR reaction with primers checking for the *agnA* insert in two different *D. discoideum* colonies (expected band size of 1191 bp), one with *wt* background (A1) and one with *agnA* knockout background (A2). Each colony was run with two different concentrations: one with a lower concentration, followed by a higher concentration of plasmid. **B.** PCR reaction with primers checking for the pDM304 backbone in the same colonies as in A, with an expected band size of 435bp. Here, positive controls are included, denoted with '+' for insert and '+bb' for pDM304 backbone. **C.** PCR reaction with primers checking for the *agnC* insert (expected band size of 1066 bp) in two different *D. discoideum* colonies, one with *wt* background (C1) and one with *agnC* knockout background (C2). Each colony was run with two different concentrations: one with a lower concentration, followed by a higher concentration of plasmid. **D.** PCR reaction with primers checking for the pDM304 backbone in the same colonies as in C, with an expected band size of 435 bp. Here, positive controls are included, denoted with '+' for insert and '+bb' for pDM304 backbone. **E.** PCR reactions of three additional agnC- colonies, checking for both the *agnC* insert and pDM304 backbone, with expected sizes of 1066 bp and

435 bp respectively. Here, only a positive control is included for both insert and backbone, denoted with 'Positive'. * marked wells indicate wells that fused together, thus containing the same sample.

2.5 Western blot

One confirmed transformant for each strain was then analyzed by western blotting, to see that the FLAG-tagged proteins were indeed being expressed within the cells (Fig. 10). Expected sizes for the proteins were 114 kDa for AgnA and 140 kDa for AgnC. A band of size ~130 can be seen for AgnA and a band ~180 kDa can be seen for AgnC. As the 180 kDa and 130 kDa ladder bands are very close to each other, it can be hard to get an exact size reading in that region, and the bands that we see are most likely those that we are looking for, despite them seeming to be larger than expected. This is further supported by the expected and visible size difference between AgnA and AgnC proteins. Bands look similar across both wt strains and strains missing the native argonaute plasmids, and their level of expression does not seem to differ.

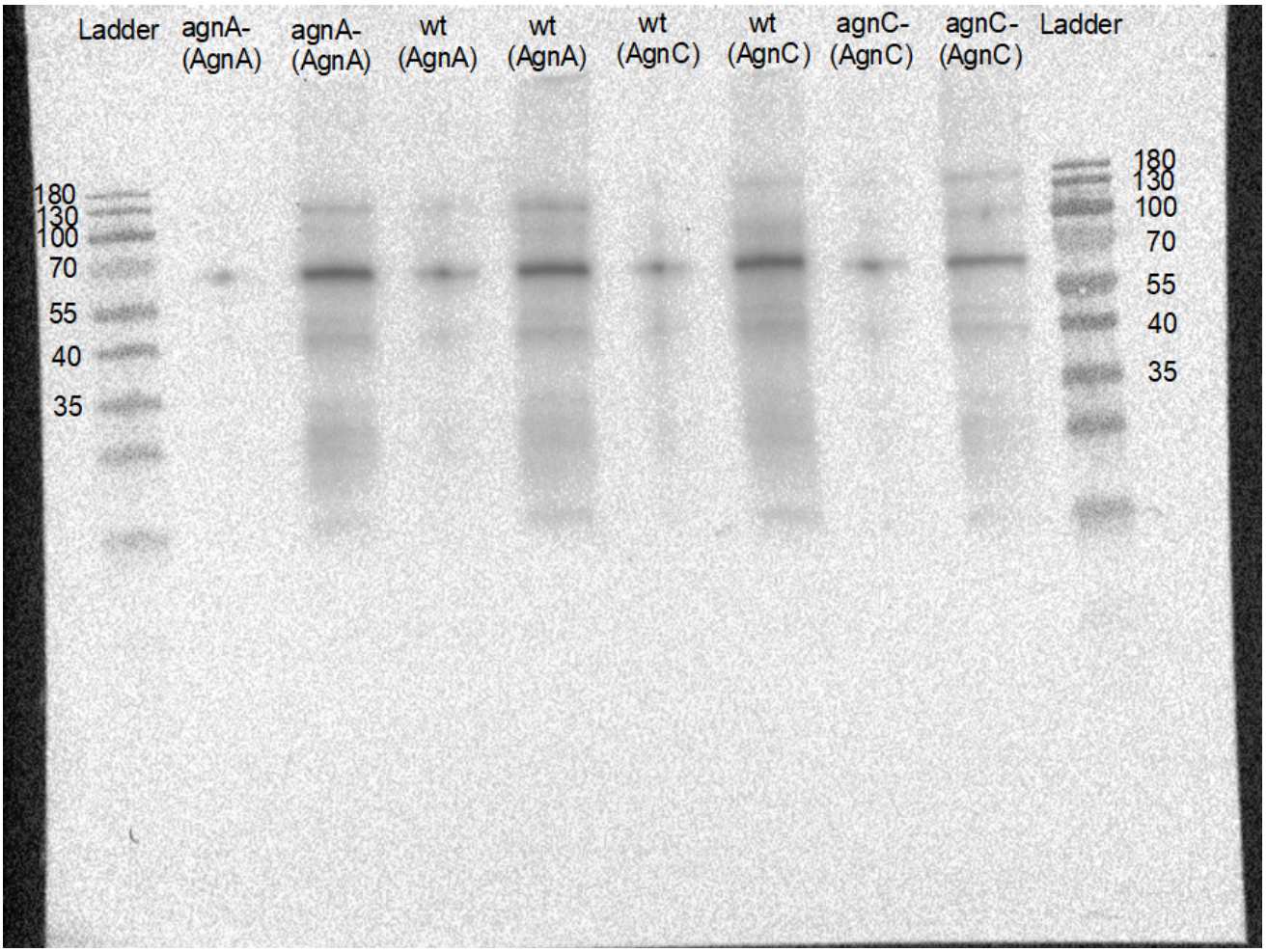


Figure 10: Western blot for AgnA *wt*, *agnA*-, AgnC *wt* and *agnC*-, each with a lower concentration (2 µL) followed by a higher concentration (20 µL). Expected sizes are 114 kDa for AgnA and 140 kDa for AgnC. *agnA*- is the *agnA* knockout strain. wt is the wildtype strain, used for both AgnA and AgnC. *agnC*- is the *agnC* knockout strain. Parentheses under the strain name indicate which plasmid was transformed in.

To finish up this analysis, AgnE was also transformed and allowed to grow, but no colonies showed up on that plate, hence either the transformation or selection had failed. Thus this transformation was repeated with better success. This also failed to yield any colonies. Due to time constraints, the AgnE transformation attempts were dropped.

3. Discussion

3.1 The argonaute constructs

The transformation of the different plasmids into the bacteria went quite well, and the bacterial transformations themselves never failed. However, due to the difficulty in selection for AgnC and AgnE, it was harder to distinguish the positive colonies from the negative ones growing on the plate. This was not as much of a problem for AgnA, where the cells, even though few, would always be positive. Due to this, most of the time of this study went to the digestion and ligation experiments, to try and get the correct plasmid ligations growing before moving on.

Different approaches were taken, such as using different volumes of plasmids and restriction enzymes. The only working solution was to allow the digestion reaction to run for much longer than normal, up to ~1 hour in combination with a gel purification and addition of a third restriction enzyme. Even better results can most likely be achieved if the digestion reaction is allowed to run for even longer, perhaps 5-10 hours.

3.2 Transformation into Dictyostelium cells

The transformation into the amoeba cells also went quite well, and at that point, only one antibiotic selection was needed (G418) so selection problems did not occur. However, the AgnE transformation was not as successful as the other two transformations. This transformation was repeated, as we thought that perhaps the transformation might have failed due to chance, or that the antibiotic at that point had been thawed and frozen several times, and might have lost its potency. Thus, for the second try, a new stock of antibiotics was used, but with the same results.

While we know that *agnE*- cells tend to grow faster, it is possible that overexpression of this argonaute, in comparison to the others, can cause a toxic effect. This was never investigated in this study, but might be an interesting thing to look out for in the future, by possibly using strains with mutated *agnE* and transforming inducible plasmids into the cells, and observing what happens as expression increases to different levels.

3.3 Expression of FLAG-tagged proteins

From the two transformants that did work, the FLAG-tagged proteins were successfully expressed (Fig. 10), which is especially clear when looking at the higher concentrations used. There is also another very clear and consistent band with a much smaller size. For our argonaute proteins, the expected sizes are 114 kDa for AgnA and 140 kDa for AgnC. The much darker band is of a much smaller size, around ~65 kDa. This has been observed before, in wt cells which have not had any plasmids transformed into them (Edelbroek B., unpublished).

Another thing of note is that the bands for the AgnC proteins are slightly larger than expected, being very close to 180 kDa, (expected is size is 140 kDa). However, the bands for 180 kDa and 130 kDa are very close to each other, and the protein identified is very likely the argonaute protein we are looking for. This could be identified much better had the gel been able to run for longer. An alternative would be to use a less dense gel, allowing the proteins to migrate faster.

Another thing to note is that the bands are very faint. This is most likely due to a low concentration of proteins (as we can see the difference of band visibility between the low and high concentration sampels). A possible solution to this would be to use many more cells in the western blot (as a higher volume would not work due to overloading), leading to a higher concetration of proteins. Another, more likely reason for this might simply be due to the low expression rate of those two specific argonautes (AgnA & AgnC) that have both previously been reported to have a lower expression than the other argonautes, often being undetectable by RT-PCR (Zhang 2006), which are in agreement with our results. However, our results were analysed through western blotting, as compared to RT-PCR, as well as using a FLAG-tagged construct instead of the native argonaute proteins. The plasmid might express the proteins to a higher degree than what would occur normally in the *D. discoideum* cells. Another similar study (Åström 2021) that used western blotting for AgnB, shows a relatively clearer band for it compared to our results for AgnA and AgnC.

Comparing AgnA and AgnC to each other, we can clearly see a size difference as we would expect. Due to this, along with the fact that the bands are around where they are supposed tro be, I am confident that the desired proteins are expressed in the cells and we can conclude that FLAG-tagged proteins can be expressed in *D. discoideum* with the use of plasmids, and we can learn more about gene regulation by using this construct to characterise these argonaute proteins and their function further.

3.4 Things left undone

Unfortunately, due to time constraints, I could not continue with the study to its much later stages. What these results show us is that the FLAG-tagged argonaute construct (except for AgnE) with the pDM304 backbone has been created, transformed into *D. discoideum* and successfully expressed. The next step would then be to characterise the argonaute proteins by

pulling them down, and also checking which RNA molecules have bound to them. This would tell us a lot about exactly what function the different argonautes might have in *D. discoideum* specifically, but also potentially in other organisms. A similar experiment has been done previously for AgnA, in which the argonaute was fused with GFP and the growth of the *D. discoideum* cells was observed in wild type and overexpression strains (Zhang 2006). They found that, during vegetative growth, there were no differences between the two strains, but that during development, the overexpressed strain is delayed compared to its wild type counterpart. It would be interesting to see if a FLAG-tagged transformant would show similar results for the growth and development of *D. discoideum*.

Furthermore, I never succeeded in transforming AgnE into the amoeba, and further tries, potentially with different approaches could be made. If unsuccessful, then it could also be interesting to investigate if AgnE overexpression does in fact have a toxic effect on the cells, and why.

4. Materials & Methods

4.1 Argonaute Plasmids

Three separate plasmids were used for the three FLAG-tagged Argonaute genes, two of which contain an *ampR* gene (*agnC & agnE*) while only one of them contains a *KanR* gene (*agnA*). The Argonaute genes sitting on each plasmid were designed by my supervisor Bart, and the plasmids were created and provided by Thermo Scientific's Invitrogen GeneArt Gene Synthesis.

4.2 Transformation of E. coli

For the entire thesis, only two bacterial strains were used, both from the *E. coli* species: DH5α & XL10-Gold. These cells were available before the start of the thesis. For transformation of DH5α and XL10-Gold competent cells, cells were thawed on ice followed by incubation with plasmid DNA on ice (volume of plasmid not exceeding 10% of total cell volume). Cells were then heat shocked at 42℃ for 30 seconds, followed by recovery on ice for 2 minutes. The complete volume was then spread on an LA plate with the appropriate antibiotic for selection (100 µg/mL for ampicilin and 50 µg/m for kanamycine).

If too many transformants are growing on the plates, an outgrowth can be created before the plating, resulting in 2 plates with different cell concentrations. This can be done by adding 1 mL LB medium to the cells after the heat shock, plate 50 µL of the cell volume, and spinning down the rest at 3000 x g for 5 minutes. The LB is then to be removed by decanting (leaving a small amount left in the tube) and the cells resuspended in the remaining LB medium, allowing for the plating of a higher concentration of cells.

4.3 Bacterial cultures

All bacterial cultures (with the exception of *Klebsiella*) were grown on LB medium (with the addition of the appropriate antibiotic). *Klebsiella* was grown in LB medium at 37℃ overnight and then stored in OD_{600} = 100 SorMC suspension. For smaller cultures, ~7 mL LB medium was used in a 50 mL Falcon tube, while larger cultures were inoculated in 50 mL LB medium in 250 mL flasks. Due to all three argonaute genes being very AT-rich, the temperature at which the cells were allowed to grow was 30℃. On certain occasions, the cells were allowed to grow at 37℃ for 3 hours before being moved to 30℃ overnight. This is to kick-start the cell growth.

4.4 Plasmid purification

Due to the midi-kits not being available at the time, all plasmid purifications were done using Thermo Fisher's GeneJET Plasmid Miniprep Kit, using an elution volume of 20 µL water instead of 50 µL elution buffer.

As the volumes for most DNA preps exceeded the maximum volume of the miniprep kits, most purifications were done in multiple rounds using the Miniprep Kit and pooling the purified plasmids together.

4.5 Digestion

Single, double and triple digestions along with uncleaved controls were all performed during the course of the thesis. Depending on which type of digestion was performed, the reaction volumes differed. For all digests, the total FD enzyme volume should not exceed 10% of the total volume, and in the case of multiple enzymes, this volume was split evenly. For single digests, ~800 ng of plasmid sample was used in a total volume of 10 µL of 1x FD buffer along with 1 µL of enzyme. For double digests, ~4 µg of plasmid sample was used (if the fragment would be used for ligation, otherwise ~800 ng) in a total volume of 100 µL of 1x FD buffer along with 5 µL of each enzyme. For triple digests, ~4 µg of plasmid sample was used in a total volume of 150 µL along with 5 µL of each enzyme. FastAP was considered an enzyme regarding total volumes, and thus a double digest that also included FastAP would use volumes corresponding to a triple digest.

All of the volume measurements were based on Thermo Scientific's guidelines for the use of the FastDigest restriction enzymes used in this thesis.

4.6 Fragment purification

Digestion fragments were either purified using Thermo Scientific's GeneJET PCR Purification Kit or their Gel Purification Kit depending on the experiment.

4.7 Ligation

For each ligation reaction, a maximum of ~100 ng of destination vector was used in a total volume of 20 µL 1x T4 ligase buffer along with 0.2 µL T4 DNA ligase (1% of the total volume). The ratio of insert:vector varied from 1:1 to 7:1 depending on the availability of plasmid DNA. If a higher volume would be used, then the amount of destination vector has to be scaled up accordingly.

4.8 PCR

For all PCR reactions, the 2x OneTaq master mix from NEBioLabs was used, along with a forward and reverse primer. The following pipetting protocol was followed, adapted from NEBio's protocol for 2x OneTaq master mix:

Template DNA added varied based on the concentration of the sample (up to a maximum of 10 ng) when liquid samples were used, or a single colony, if bacterial colonies were used. This sample was added into a MasterMix containing 0.5 µL of 10 µM of each primer in a total volume of 25 µL 1x OneTaq Master Mix.

4.9 Gel electrophoresis

All gels were run on 1% agarose gels in TAE buffer (made from 50x stock). This gel was created by mixing 1g of Agarose with 100 mL 1x TAE buffer (scaled up depending on the desired volume).

4.10 Sanger sequencing

Sanger sequencing was used for each sequencing of this thesis and was carried out by Macrogen Europe and the results were interpreted through QIAGEN CLC workbench.

4.11 *D. discoideum* strains

For this thesis, 4 different *D. discoideum* strains were used:
WT (AX2 background)
AgnA- (AX4 background)
AgnC- (AX2 background)
AgnE- (AX2 background)

4.12 *D. discoideum* cultures

D. discoideum cultures were initially grown on SM agar with *K. aerogenes* until they had grown sufficiently big (around third ways across the bacterial lawn). A metal loop was then used to scrape the *D. discoideum* cells on the edges (on the border to the bacterial lawn) and dipped in 20 mL HL5-C medium (Formedium), where they were allowed to grow further at 22℃.

At later stages of the project, HL5 was no longer used, and instead, a bacterial suspension containing OD_{600}= 2 *K. aerogenes* in SorMC buffer was used to further grow *D. discoideum* strains.

4.13 SorMC buffer

The recipe for the SorMC buffer (Sorensen + $MgCl_2$ & $CaCl_2$ buffer) was adapted from Paschke *et al.* 2019 and was made the following way:

To make 1 L 1x, a 1L flask was filled with ~900 mL dH_2O, to not exceed 1 L when adding the other reagents. 2.036 g of KH_2PO_4 along with 0.547 g of $Na_2HPO_4•7H_2O$ was dissolved in the water and mixed at room temperature. The volume was then brought up to 1 L with the addition of dH_2O. After the reagents had fully dissolved, 50 µL each of 1M $MgCl_2$ and 1M $CaCl_2$ was added to the solution and mixed. The entire solution was then autoclaved before use.

4.14 SM agar

1 L SM agar was made by dissolving 24.9 SM agar powder (Formedium) in 1 L dH_2O followed by autoclaving.

4.15 HL5-C medium

1L HL5-C medium (including Glucose) was made by dissolving 26.55 g of HL5-C powder, supplied by FORMEDIUM™, in 1 L dH_2O followed by autoclaving.

4.16 Transformation of D. discoideum

D. discoideum was transformed through electroporation. They were electroporated in the following way, adapted from Paschke *et al.* 2018:

Cells were allowed to grow in 20 mL OD_{600}= 2 *K. aerogenes* in SorMC buffer overnight, after which the medium was removed by vacuum, and the cells loosened from the plastic by tapping. They were then washed with 20 mL SorMC buffer a few times (by re-using the same 20 mL SorMC, as to not lose cells) until most cells had loosened and suspended in the buffer. This cell suspension was then moved into a 50 mL falcon tube and spun down at 400 x g for 5 minutes, followed by re-suspension in new 20 mL SorMC buffer. This step was repeated once more, this time resuspending in 400 µL SorMC buffer (only if using a 2 mm gap cuvette. 800 µL if using a 4 mm gap cuvette). 1-2 µg of plasmid DNA was added to the tubes and mixed and moved into a 2 mm gap cuvette chilled on ice. Cells were electroporated twice with the following settings: 500 V (1 kV if using a 4 mm gap cuvette), 25 µF, resistance at 50 Ω (if possible, otherwise 100 Ω should also work), zap twice, with a 2-second gap between each zap. Cells were moved to new plates containing 20 mL OD_{600} = 2 *K. aerogenes* in SorMC buffer and allowed to recover for 5 hours before adding G418 (10 µg/mL) selection and allowed to grow for 2-4 days.

4.17 Counting of cells

D. discoideum cells were counted with the use of a Hemocytometer. 10 µL of cell suspension was used in the cytometer.

4.18 Western Blot

Cells were harvested by spinning them down at 400 x g, 4℃ for 5 minutes followed by washing in cold PBS twice, using 10 mL and 5 mL respectively. To lyse the cells, they were resuspended in 0,8 mL FLAG-IP buffer (see below) with the addition of 1x protease inhibitor (Thermo Scientific) and incubated for 30 minutes on ice (with mixing every ~10 minutes). Cells were spun down at ~20 000 rpm, 4℃ for 10 minutes after which the lysate was decanted into cold Eppendorf tubes, and the volume was brought to ~1 mL by adding more FLAG-IP buffer (with protease inhibitor). From this solution, 20 µL was mixed with 4x Laemmli and denatured by boiling for 5 minutes at 95℃.

The samples were then loaded on a Mini-PROTEAN TGX Stain-Free precast gel (BioRad) with PageRuler Prestained Protein Ladder (Thermo Scientific) as reference. The gel was run for 1h at 100V in running buffer (25 mM Tris, 190 mM glycine, 0.1% SDS).

4.19 FLAG-IP buffer

FLAG-IP buffer was made by mixing 50 mM Tris-HCl (pH 7.5) with 150 mM NaCl, 1% Triton X-100 and 1 mM EDTA. This buffer did not contain any protease inhibitor, as this was added right before use.

Acknowledgements

I would like to thank my supervisor Fredrik Söderbom for giving me the opportunity to work with the Dictys in the lab, and learning so much in the process.

Thanks to Bart Edelbroek for supervising and guiding me all throughout the thesis and teaching me countless laboratory methods.

Thanks to the Micro people for creating a very nice environment in the lab to work in.

References

Annesley SJ, Fisher PR. 2009. Dictyostelium discoideum—a model for many reasons. Molecular and Cellular Biochemistry 329: 73–91.

Asghar A, Groth M, Siol O, Gaube F, Enzensperger C, Glöckner G, Winckler T. 2012. Developmental Gene Regulation by an Ancient Intercellular Communication System in Social Amoebae. Protist 163: 25–37.

Åström M. 2021. Search for the Argonaute protein that governs miRNA regulation in Dictyostelium discoideum.

Boesler B, Meier D, Förstner KU, Friedrich M, Hammann C, Sharma CM, Nellen W. 2014. Argonaute Proteins Affect siRNA Levels and Accumulation of a Novel Extrachromosomal DNA from the Dictyostelium Retrotransposon DIRS-1 *. Journal of Biological Chemistry 289: 35124–35138.

Cerutti H, Casas-Mollano JA. 2006. On the origin and functions of RNA-mediated silencing: from protists to man. Current Genetics 50: 81–99.

Dolinoy DC. 2007. Epigenetic gene regulation: early environmental exposures. Pharmacogenomics 8: 5–10.

Engels BM, Hutvagner G. 2006. Principles and effects of microRNA-mediated post-transcriptional gene regulation. Oncogene 25: 6163–6169.

Glisovic T, Bachorik JL, Yong J, Dreyfuss G. 2008. RNA-binding proteins and post-transcriptional gene regulation. FEBS Letters 582: 1977–1986.

Grishok A. 2005. RNAi mechanisms in Caenorhabditis elegans. FEBS Letters 579: 5932–5939.

Hiraoka H, Nakano T, Kuwana S, Fukuzawa M, Hirano Y, Ueda M, Haraguchi T, Hiraoka Y. 2020. Intracellular ATP levels influence cell fates in Dictyostelium discoideum differentiation. Genes to Cells 25: 312–326.

Hobert O. 2008. Gene Regulation by Transcription Factors and MicroRNAs. Science 319: 1785–1786.

Lam JKW, Chow MYT, Zhang Y, Leung SWS. 2015. siRNA Versus miRNA as Therapeutics for Gene Silencing. Molecular Therapy - Nucleic Acids, doi 10.1038/mtna.2015.23.

Liao Z, Kjellin J, Dahlén H, Söderbom F. 2018a. The Dictyostelium discoideum Argonaute protein AgnE regulates cell growth.

Liao Z, Kjellin J, Fransson Å, Söderbom F. 2018b. Functional analyses of RNA interference effectors in Dictyostelium discoideum during growth and development.

Loomis W. 2012. Dictyostelium Discoideum: A Developmental System. Elsevier

Lynn K, Fernandez A, Aida M, Sedbrook J, Tasaka M, Masson P, Barton MK. 1999. The

PINHEAD/ZWILLE gene acts pleiotropically in Arabidopsis development and has overlapping functions with the ARGONAUTE1 gene. Development 126: 469–481.
Paschke P, Knecht DA, Silale A, Traynor D, Williams TD, Thomason PA, Insall RH, Chubb JR, Kay RR, Veltman DM. 2018. Rapid and efficient genetic engineering of both wild type and axenic strains of Dictyostelium discoideum. PLOS ONE 13: e0196809.
Paschke P, Knecht DA, Williams TD, Thomason PA, Insall RH, Chubb JR, Kay RR, Veltman DM. 2019. Genetic Engineering of Dictyostelium discoideum Cells Based on Selection and Growth on Bacteria. Journal of Visualized Experiments : JoVE 58981.
Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB. 2005. Gene Regulation at the Single-Cell Level. Science 307: 1962–1965.
Sharma UK, Chatterji D. 2010. Transcriptional switching in Escherichia coli during stress and starvation by modulation of σ70 activity. FEMS Microbiology Reviews 34: 646–657.
Swarts DC, Makarova K, Wang Y, Nakanishi K, Ketting RF, Koonin EV, Patel DJ, van der Oost J. 2014. The evolutionary journey of Argonaute proteins. Nature Structural & Molecular Biology 21: 743–753.
Valinezhad Orang A, Safaralizadeh R, Kazemzadeh-Bavili M. 2014. Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation. International Journal of Genomics 2014: e970607.
Veltman DM, Akar G, Bosgraaf L, Van Haastert PJM. 2009. A new set of small, extrachromosomal expression vectors for Dictyostelium discoideum. Plasmid 61: 110–118.
Wilson JE, Connell JE, Macdonald PM. 1996. aubergine enhances oskar translation in the Drosophila ovary. Development 122: 1631–1639.
Wu J, Yang J, Cho WC, Zheng Y. 2020. Argonaute proteins: Structural features, functions and emerging roles. Journal of Advanced Research 24: 317–324.
Zhang X. 2006. On the function of the Dictyostelium Argonaute A protein (AgnA) in epigenetic gene regulation.
1998. AGO1 defines a novel locus of Arabidopsis controlling leaf development. The EMBO Journal 17: 170–180.

Supplementary material

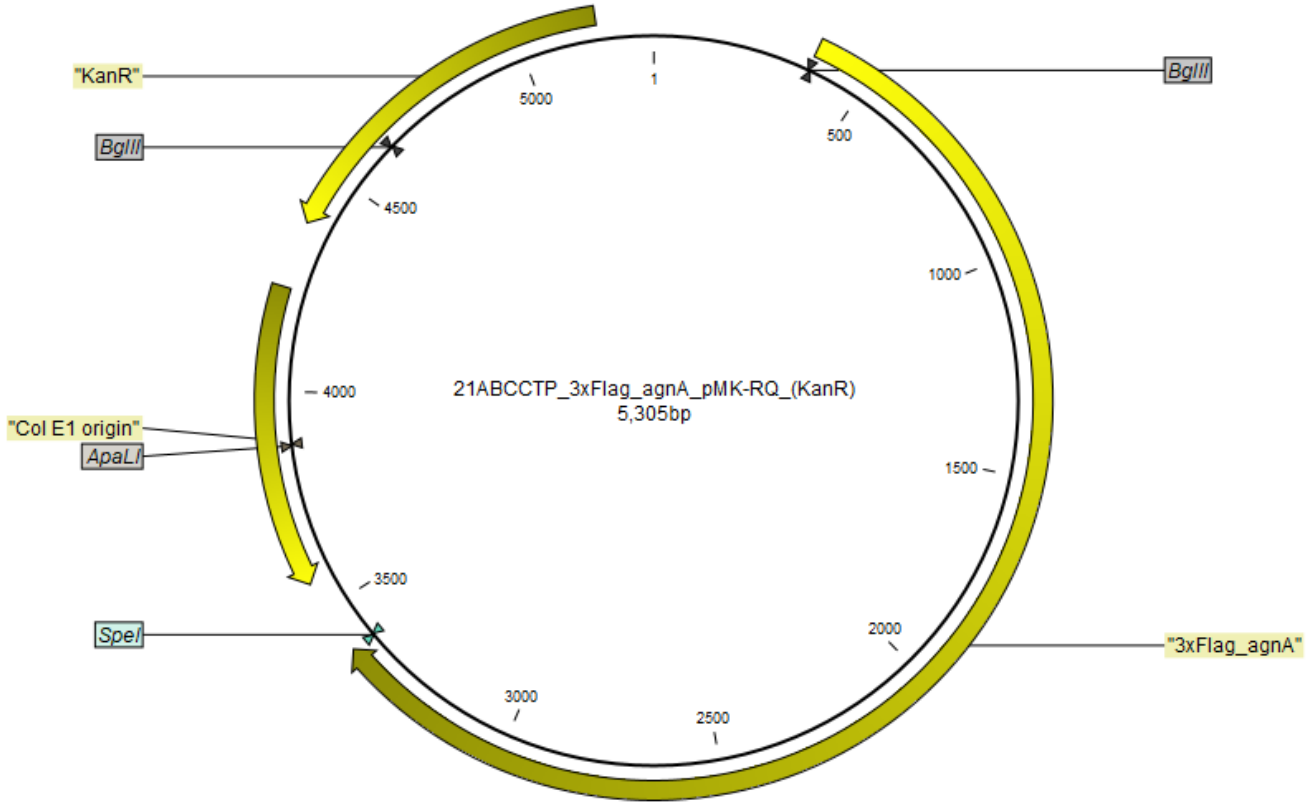
Supplementary Table 1: List of all oligos to produce the results in this study.

Oligo	Type	Sequence	Used for
887	Forward	GCCAATGGTGCATAGTAGAATGG	AgnA insert
826	Reverse	GGATCCAATTGGTAATCCTCCAGC	AgnA insert
1014	Forward	GCAATTACCTCTAGTGGAGATTG	AgnC insert
637	Reverse	CTAACTAGTATATGAGAATCTAGTATTGATG	AgnC insert
922	Forward	GGTTATATATCAAGTGTTCATAC	AgnE insert

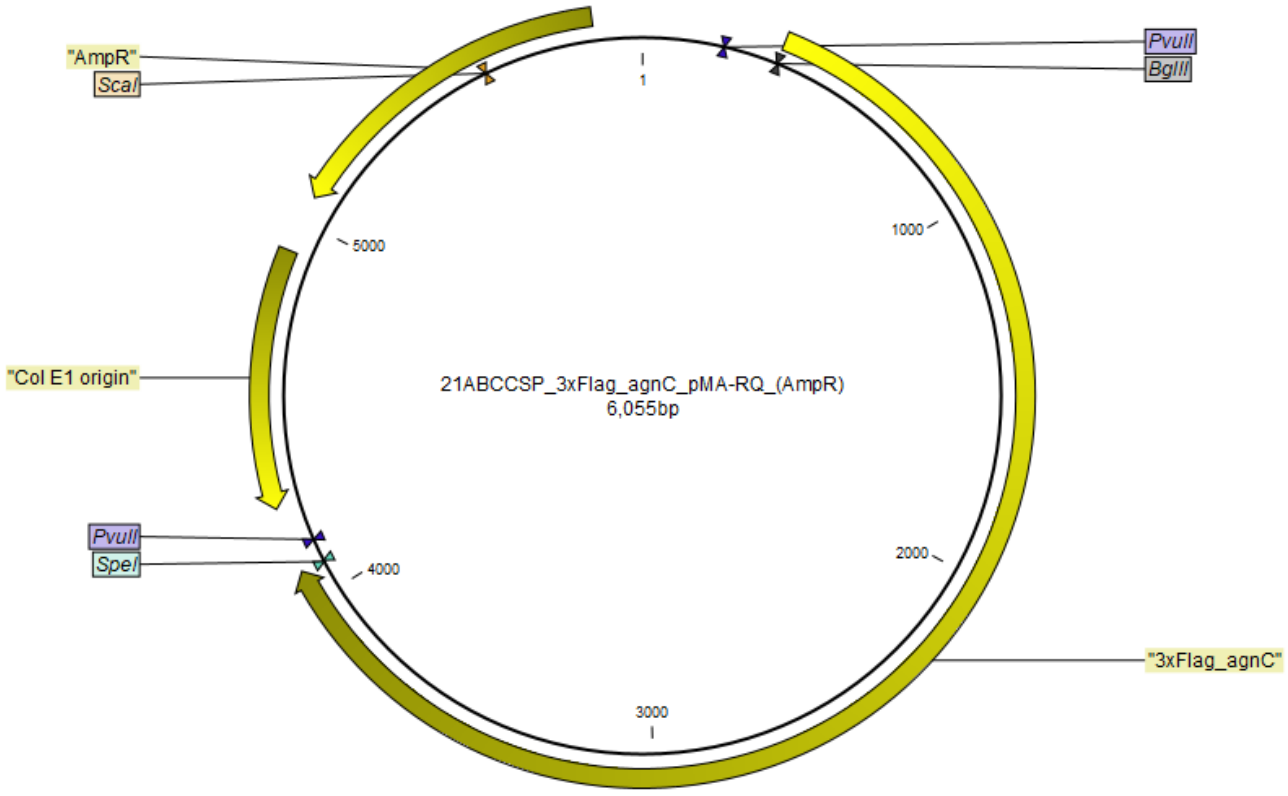
926	Reverse	CCACCACTATCT CTATAACTTCT	AgnE insert
1012	Forward	AATATCCACTTG GAGGAGGTGGA	pDM304 backbone
1013	Reverse	CACGTTCAAAGG GTAAGAATAAGG	pDM304 backbone
373	Reverse	GCTCATGGTGAA TGTATAACAA	pDM304 backbone

Supplementary Table 2: List of all plasmid sizes, along with sizes for the argonaute gene fragments after double digest.

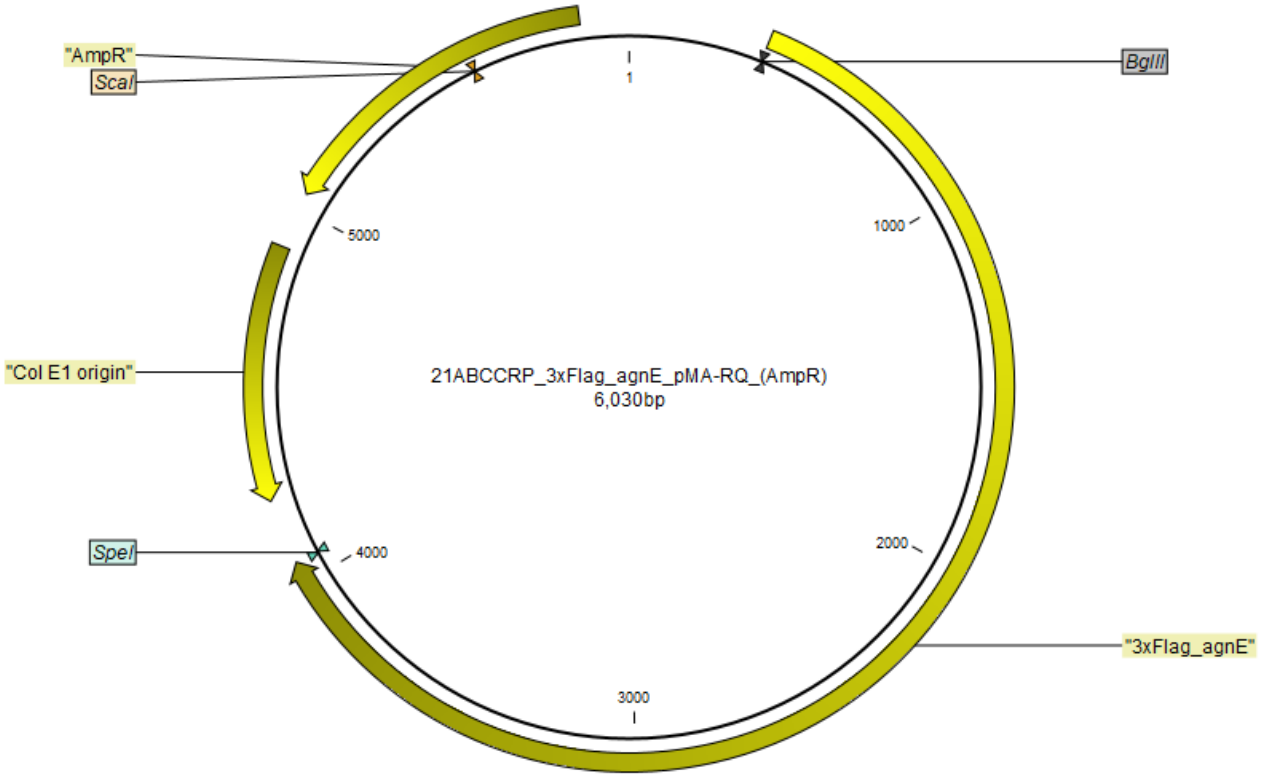
Plasmid/fragment	**Size [basepair]**
AgnA intact plasmid	5 305
AgnA gene fragment	3 027
AgnC intact plasmid	6 055
AgnC gene fragment	3 714
AgnE intact plasmid	6 030
AgnE gene fragment	3 689
pDM304 intact plasmid	6 839



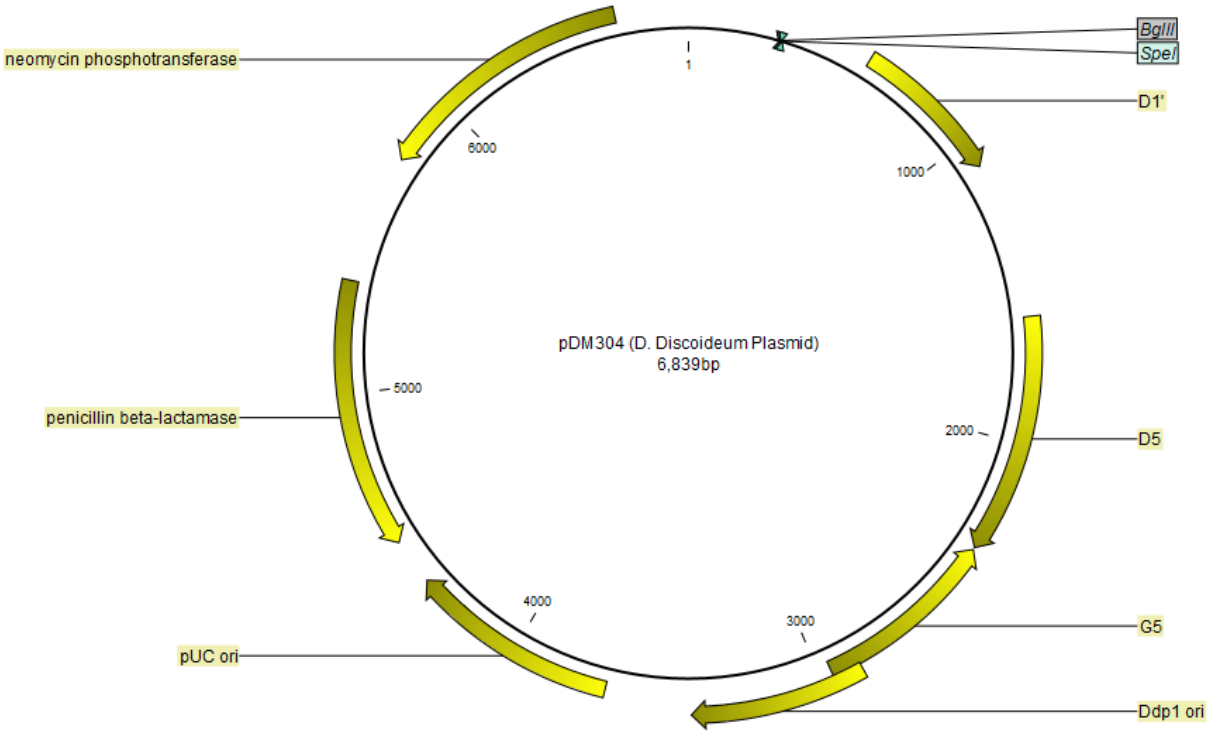
Supplementary figure 1: Showing the plasmid map for AgnA with the original backbone as ordered.



Supplementary figure 2: Showing the plasmid map for AgnC with the original backbone as ordered.



Supplementary figure 3: Showing the plasmid map for AgnE with the original backbone as ordered.



Supplementary figure 4: Showing the plasmid map for pDM304 without any argonaute inserts.