University of Warsaw
Faculty of Biology

Mohamed Abo Baker Elkomy

Functional characterization of the RNA guided DRM2 DNA methyltransferase

Charakterystyka metylotransferazy DNA - DRM2 kierowanej przez RNA

Master thesis

The work done under the direction of
Prof. Dr. Hab. Ewa Bartnik
Department of Biology, University of Warsaw

Warsaw, June 2016

The statement of the thesis supervisor (Oświadczenie kierującego pracą)

I declare that this work was prepared under my direction and certify that it meets the conditions to present it in the procedure for granting a professional title.

Oświadczam, że niniejsza praca została przygotowana pod moim kierunkiem i stwierdzam, że spełnia ona warunki do przedstawienia jej w postępowaniu o nadanie tytułu zawodowego.

Date/ Data Signature/ Podpis kierującego pracą

The statement author work (Oświadczenie autora pracy)

Aware of legal responsibility declare that this dissertation was written by e me alone and does not contain content obtained in a manner inconsistent with applicable regulations.

I also declare that the presented work has not previously been subject to procedures related to obtaining the professional title of university.

I also certify that this version of the work is identical with the attached electronic version.

Świadom odpowiedzialności prawnej oświadczam, że niniejsza praca dyplomowa została napisana przeze mnie samodzielnie i nie zawiera treści uzyskanych w sposób niezgodny z obowiązującymi przepisami.

Oświadczam również, że przedstawiona praca nie była wcześniej przedmiotem procedur związanych z uzyskaniem tytułu zawodowego w wyższej uczelni.

Oświadczam ponadto, że niniejsza wersja pracy jest identyczna z załączoną wersją elektroniczną.

Date/ Data Signature of the author work/ Podpis autora pracy

Summary:

DNA methylation in plants is in part controlled by small RNAs. A key player in RNA directed DNA methylation (RdDM) is Domain Rearranged Methyltransferase 2 (DRM2). In addition to its catalytic domain, which resembles the catalytic domain of animal DNA methyltransferases, the protein has several UBA domains of unknown biochemical significance. Genetic experiments on *Arabidopsis thaliana* indicate that a DRM2 variant lacking all UBA domains is similarly incapacitated as a catalytic mutant. The aim of this work was to elucidate the role of the UBA domains in DRM2. Towards this end, expression plasmids were generated, which should direct expression of DRM2 variants lacking individual UBA domains only after Agrobacterium mediated transformation. Unfortunately, protein expression of the deletion constructs was not observed, even though a reporter control protein was expressed in parallel experiments.

.

Keywords:

DRM2, DNA methyltransferase, RdDM, CHH methylation, UBA domain, Arabidospsis thaliana, Agrobacterium tumefaciens.

Field work

13.4 - Biotechnology

Streszczenie:

Metylacja DNA w roślinach jest częściowo kontrolowana przez małe RNA. Domena metylotransferazy cytozyny (DRM2, ang. domain rearranged methyltransferase 2) pełni kluczową rolę w metylacji DNA zależnej od RNA (RdDM). Ponadto, oprócz jej domeny katalitycznej, która przypomina domenę katalityczną metylotransferaz DNA pochodzenia zwierzęcego, białko posiada kilka domen UBA o nieznanej roli biochemicznej. Eksperymenty genetyczne na Arabidopsis thaliana wskazują, że wariant DRM2 pozbawiony wszystkich domen UBA jest tak samo nieaktywny jak mutant katalityczny. Celem niniejszej pracy było wyjaśnienie roli domen UBA w metylotransferazie DRM2. W tym celu zostały wytworzone plazmidy ekspresyjne posiadające różne warianty DRM2 z brakującymi pojedynczymi domenami UBA ulegające bezpośredniej ekspresji po transformacji z użyciem Agrobacterium. Niestety, nie zaobserwowano ekspresji białek z konstruktów delecyjnych pomimo obecności ekspresji białka kontrolnego w równoległych doświadczeniach.

Słowa kluczowe:

DRM2, metylotransferaza DNA, RdDM, metylacja CHH, domena UBA, Arabidospsis thaliana, Agrobacterium tumefaciens

Dziedzina pracy

13.4 – Biotechnologia

Acknowledgment:

I would like to thank Prof. Matthias Bochtler for hosting me in his laboratory and graciously contributing his scientific advice during the project and with writing this thesis. Also, thanks to Dr. Humberto Fernandes for providing guidance during my master thesis work. Moreover, I credit Dr. Ruslan Yatusevich from Dr. Swiezewski's Laboratory and Dr. Rafal Archacki from Prof Jerzmanowski's Laboratory for sharing plant material with me. I praise Dr.Wayne Dawson for English proofreading.

I am grateful also to Dr. Ayman El-Seedy for his recommendation and help on joining Dr. Humberto Fernandes in this project.

Many thanks to the Foundation of Polish Science (FNP) for support of my thesis under the HOMING PLUS/2013-7/5 grant "Plant-specific methyltransferase DNA structures".

Above all, I am grateful for my parents who supported me along my scientific track and never hesitated to provide guidance and always shared their love.

Thanks to everyone who offered services and resources for me to finish my master degree.

Sincerely,

Mohamed Elkomy

Abbreviation:	Names:
29	truncated version of DRM2 starts from residue 29
29pDONR201	truncated version of DRM2 starts from residue 29 loaded on pDONR201
29pGWB620	truncated version of DRM2 starts from residue 29 loaded on pGWB620
55	truncated version of DRM2 starts from residue 55
55pDONR201	truncated version of DRM2 starts from residue 55 loaded on pDONR201
55pGWB620	truncated version of DRM2 starts from residue 55 loaded on pGWB620
110	truncated version of DRM2 starts from residue 110
110pDONR201	truncated version of DRM2 starts from residue 110 loaded on pDONR201
110pGWB620	truncated version of DRM2 starts from residue 110 loaded on pGWB620
191	truncated version of DRM2 starts from residue 191
191pDONR201	truncated version of DRM2 starts from residue 191 loaded on pDONR201
191pGWB620	truncated version of DRM2 starts from residue 191 loaded on pGWB620
At	*Arabidopsis thaliana*
ATP	Adenosine triphosphate
BAH	Bromo Adjacent Homology
CLSY	Chromatin remodeling protein
CMT3	Chromomethyltransferase 3
DCL3	Dicer like 3
DDT	Dichlorodiphenyltrichloroethane
DMS3	DEFECTIVE IN MERISTEM SILENCING 3
DNA	Deoxyribonucleic acid
DNMT1	DNA Methyltransferase1
DNMTs	DNA methyltransferases
Dpn	*Diplococcus pneumoniae*
DRD1	DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1
DRM	Domain Rearranged Methyltransferase
ECL	Enhanced Chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
F1	generation 1
gm	Gram
H3K9	histone no.3 lysine no.9
HEN 1	HUA Enhancer 1
KTF1	Kow domain-containing Transcription Factor 1
KYP	Kryptonite
M	Molar
MBD	Methyl Binding Domain
MET1	Methyltransferase 1
Mfe	Mycoplasma fermentas
ml	Milliliter

mRNA	Messenger RNA
MTase	methyltransferase
Nt	*Nicotiana tobacum*
PBS	Phosphate-Buffered Saline
PCR	Polymerase Chain Reaction
PHD	Plant Homo Domain
Pme	*Pseudomonas mendocina*
RdDM	RNA Directed DNA methylation
RDM1	RNA-DIRECTED DNA METHYLATION 1
RDR2	RNA Dependant RNA polymerase 2
Sac	*Streptomyces achromogenes*
SAM	S-adenosylmethionine
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SET	Enhancer of zeste
siRNAs	small interfering RNA
Spe	Sphaerotilus species
SRA	RING finger associated
SAWADEE	SAWADEE HOMEODOMAIN HOMOLOGUE 1
Taq	*Thermus aquaticus*
UBA	Ubiquitin Associated domain
UTR	Untranslated Region
Xba	*Xanthomonas badrii*

Table of Contents

1 Introduction

1.1 DNA methyltransferases

DNA methylation occurs in all kingdoms of life. In higher eukaryotes, DNA methylation is limited to methylation of cytosine bases and catalyzed by S-adenosylmethionine (SAM) dependent DNA methyltransferases. In mammals, most DNA outside the regulatory regions (intergenic DNA, coding DNA and repeat elements) is methylated.[1] In plants and fungi, genomic methylation is mainly restricted to the coding sequences of transposons and other repeats. In mammals, DNA methylation is mostly limited to the CpG context (except in embryonic stem cells).[2] In plants, DNA methylation is not only found in the CpG context, but also in CHG and CHH (where H is A, C or T). In the *Arabidopsis thaliana* genome, the CG sites are methylated over 80% of times, CHG sites are 20-100% methylated and CHH sites tend to be below 20%.[3]

1.1.1 *De novo* Vs. Maintenance

DNA methylation is either introduced *de novo,* or after DNA replication to counteract the dilution of methylation by semi-conservative DNA replication. According to their involvement in predominantly one or the other process, DNA methyltransferases are either classified as de novo or maintenance DNA methyltransferases. In mammals, DNMT1 plays the role of the maintenance DNA methyltransferase, whereas the DNMT3A/DNMT3L and DNMT3B/DNMT3L heteroduplexes act as de novo-methyltransferases. In the latter complexes, DNMT3A and DNMT3B are the catalytically active components, but they depend on binding to DNMT3L for activity.

Plants have a larger repertoire of DNA methyltransferases (Figure1.1), which can be classified according to target sequence. Methyltransferase 1 (MET1) is the main maintenance DNA methytransferase. It acts on the symmetric CG context as evidenced by the loss of methylation in this context in *met1* mutants and antisense MET1 transgenic plants. Chromomethyltransferase 3 (CMT3) is a plant specific DNA methyltransferase and its main role is the maintenance of DNA methylation at CHG sites. Mutants of *cmt3* show less methylation at CHG sites. DRM1 and DRM2 methylate primarily CHH sites.

CHH methylation sites are asymmetric-asymmetric means on one strand-, and therefore DNA methylation on the daughter strand cannot be restored based on the methylation of the parental strand. Mutants of *drm1*and *drm2* lack *de novo* methylation at least at the FWA (Flowering Wageningen) and SUP (Superman) loci.[4]

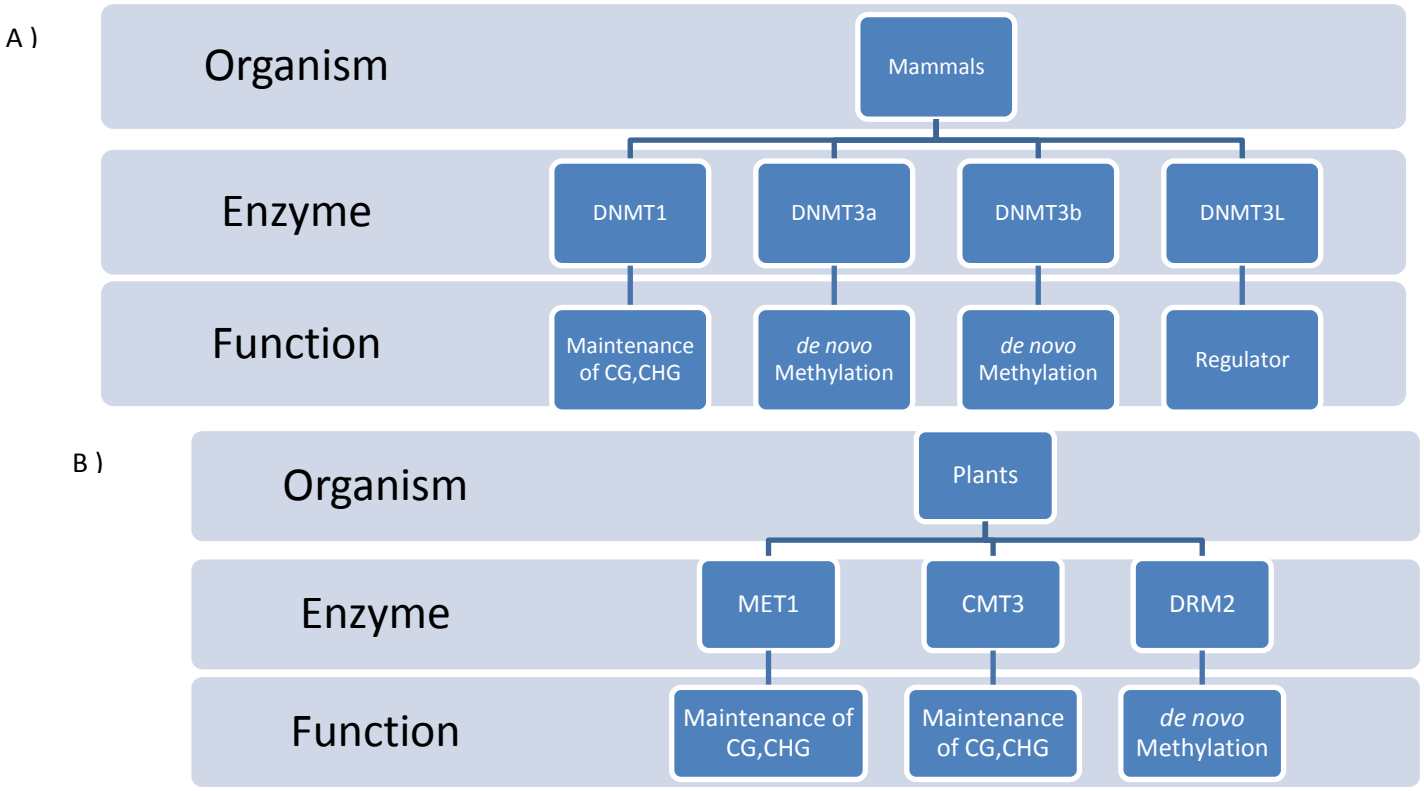


Figure 1.1: Identity and function of the main DNA methyltransferases in animals and plants

1.1.1.1 MET Family

Methyltransferase 1 (MET1) is the plant ortholog of the mammalian DNA methyltransferase1 (DNMT1), and shares 50% amino acid identity at the C-terminal domain, and 24% amino acid identity in the N-terminal domain. The Arabidopsis *met1* mutant shows a severe reduction in its CG methylation pattern. Comparison of transcript levels in the *met1* mutant and wild type plants reveals that expression of both methylated and unmethylated genes increases in the *met1* mutant, suggesting a role of MET1 in transcriptome silencing. In addition, the *met1* mutant lacks expression of other genes like

ibm1, sadnu6-1, and RPS. In addition, MET1 is responsible for methylation of CG during the gametophytic phase through DNA replication. Analysis of the *met1* mutant showed that paternal and maternal genomes have essential roles in the regulation of seed size, and that hypomethylation correlated with reduced seed size. Maternal genome hypomethylation does not affect seed size.[5]

1.1.1.2 CMT Family

The chromomethyltransferase (CMT) family controls non-CG methylation and is composed of CMT1, CMT2, and CMT3. Mutations in CMT3 were found to lead to global loss of CHG and CHH methylation. Using bisulfite sequencing it has been shown that CMT3 has preferentially methylates CHG sites. It has a normal phenotype as its mutation only affects methylation. CMT methyltransferases contain a special chromodomain amino acid motif that is essential for guiding proteins in heterochromatin, by recognition of histone modifications, which gives insights for its role in modifying the DNA heterochromatin area.[5, 6]

The *cmt3* mutation eliminates epigenetic silencing of SUP and PAI genes in Arabidopsis. In addition to methylating DNA at the CHG context at SUP and PAI loci, it also methylates other endogenous sequences at Ta3 transposons. Moreover, the sadhu6-1 gene, which is strongly expressed in the *met1* mutant, is not present in the *cmt3* mutant.[5]

CMT3 contains a BAH domain at the N-terminus, a DNA methyltransferase domain at the C-terminus and a chromodomain embedded inside the DNA methyltransferase domain. The structure shows an aromatic cage within the BAH domain and Chromo domains that allows them to recognize H3K9me2 (Histone 3 lysine 9 dimethyltransferase).[6]

1.1.1.3 DRM family

1.1.1.3.1 DRM structure

Domain Rearranged Methyltransferase (DRM) is found in plants and is the *de novo* DNA methyltransferase homologue of the mammalian DNMT3, but displays a rearranged catalytic motif between I-V and VI-X, with the latter being closer to the N-terminus of

the protein than the former. In contrast to animal DNA methyltransferases and MET1 and CMT3 plant methyltransferases, the domain-rearranged methyltransferase contains three Ubiquitin Associated domains (UBAs) at its N-terminus.[7, 8, 9]

These UBA domains consist of three helices connected by two conserved loops. This arrangement forms a conserved hydrophobic patch that may act as a common protein interacting interface. The available crystal structure of DRM2 is only for the catalytic domain (Figure 1.2), which performs the DNA methylation, but the three-dimensional arrangement of the DRM2 and UBA domains is still unknown. Moreover, this effort only revealed the structure of the *Nicotiana tobacum* homologue, while the vast majority of genetic and biochemical data arose from work on the model organism *Arabidopsis thaliana*. The structure reveals the methyltransferase (Mtase) domain in the complex with Sinefungin, which is a SAM analog. The terminal 30 residues (259-288) at the N-terminus of NtDRM2 Mtase form a long loop that wraps around the surface of the core methyltransferase domain consisting of the remaining residues (289-608). The first residue of the core methyltransferase is adjacent to the C-terminus of the protein. In addition, the same N and C terminus of Dnmt3 are adjacent to each other, showing that despite the domain rearrangements the DRM2 protein retains the traditional DNA methyltransferase fold. Moreover, the structure reveals a homodimer arrangement (Fig 1.2), and further analyses – such as mutating residues involved in dimerization – indicate that this dimerization is important for catalytic activity. [8] DRM2 and DRM3 show the same domain arrangements with the N-Terminal UBA domains and C-terminal cytosine methyltransferase domain. Moreover, DRM1 shows sequence similarity with DRM2.[10, 11]

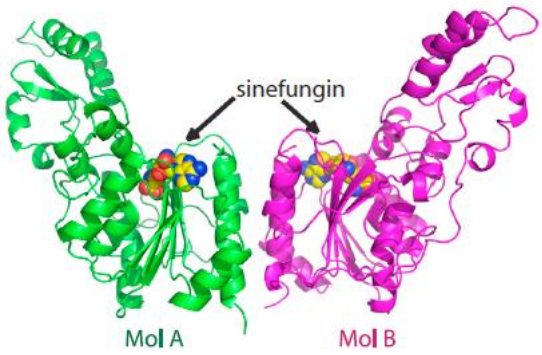


Figure 1.2: Overall structure of NtDRM2 in ribbon representation. The two monomers are displayed with different colors (green and magenta), and the sinefungin cofactor analogue is shown in space-filling representation. [10]

1.1.1.3.2 DRM function

DRM1 and DRM2 are responsible for methylation of cytosine at CHG and CHH sites. DRM1 is expressed in the mature egg cells and DRM2 is expressed in the developing embryo and roots. [12]

The *drm1/drm2* mutant has the phenotypic features of the single *drm2* mutant. [5, 12] One gene used as a test case of the involvement of DRM on DNA methylation is the FWA. Plants show a late flowering phenotype when this promoter is methylated. Transformation of an FWA transgene with a methylated promoter gives rise to late flowering in wild type (wt) *Arabidopsis thaliana* plants, and to an early flowering phenotype in the *drm1/drm2* mutant. [9]

DRM3 cannot substitute DRM2 activity *in vivo* but it is nonetheless required for establishment of RdDM, which suggests that it may play a role analogous to DNMT3L in mammals. However, it was later found that DRM3 does not bind to DRM2. Instead, the protein physically interacts with RNA Polymerase V and promotes its activity. [11, 13] The functional importance of the UBA domains in DRM2 has been tested by single mutations of phenylalanine residues into alanines in three UBA domains, at its conserved MGF motif present between the first two α-helices of the domain. The mutations do not compromise DRM2 stability or localization, but they lead to reduced DNA methylation in non-CpG context. At the CHH sites, the decrease in methylation is comparable with the effect of a mutation that abolishes DRM2. It is not clear why the UBA domains of DRM2 are required for the function of the protein. [8, 11]

1.2 Mechanism of RNA dependent DNA methylation

DRM2 is involved in RNA dependent DNA methylation. Apart from DRM2, the pathway contains two main components, which are plant specific atypical DNA dependent RNA polymerases named Pol IV and Pol V (Fig 1.3). They have their own specific subunits but share subunits with each other and with RNA Pol II. [14, 15]

Pol IV synthesizes RNA transcripts, which are then converted to double stranded RNA duplexes by RNA Dependent RNA polymerase (RDR2). Double stranded RNA is then cleaved into 24nt primary siRNA by Dicer like 3 (DCL3). HUA Enhancer 1 (HEN1) then methylates the 2′OH group of the 3′-terminal nucleotide of siRNA, which is subsequently loaded on AGO4 (Figure 1.3). [14, 16]

Pol V transcribes candidate loci selected by the DDR complex, consisting of Defective in Meristem Silencing 3 (DMS3), Defective in RNA-directed DNA methylation 1 (DRD1) and RNA-directed DNA Methylation 1 (RDM1). [14, 16] If there is a match between PolV transcripts and AGO4 bound siRNAs, DRM2 methylates the PolV transcribed locus. [17]

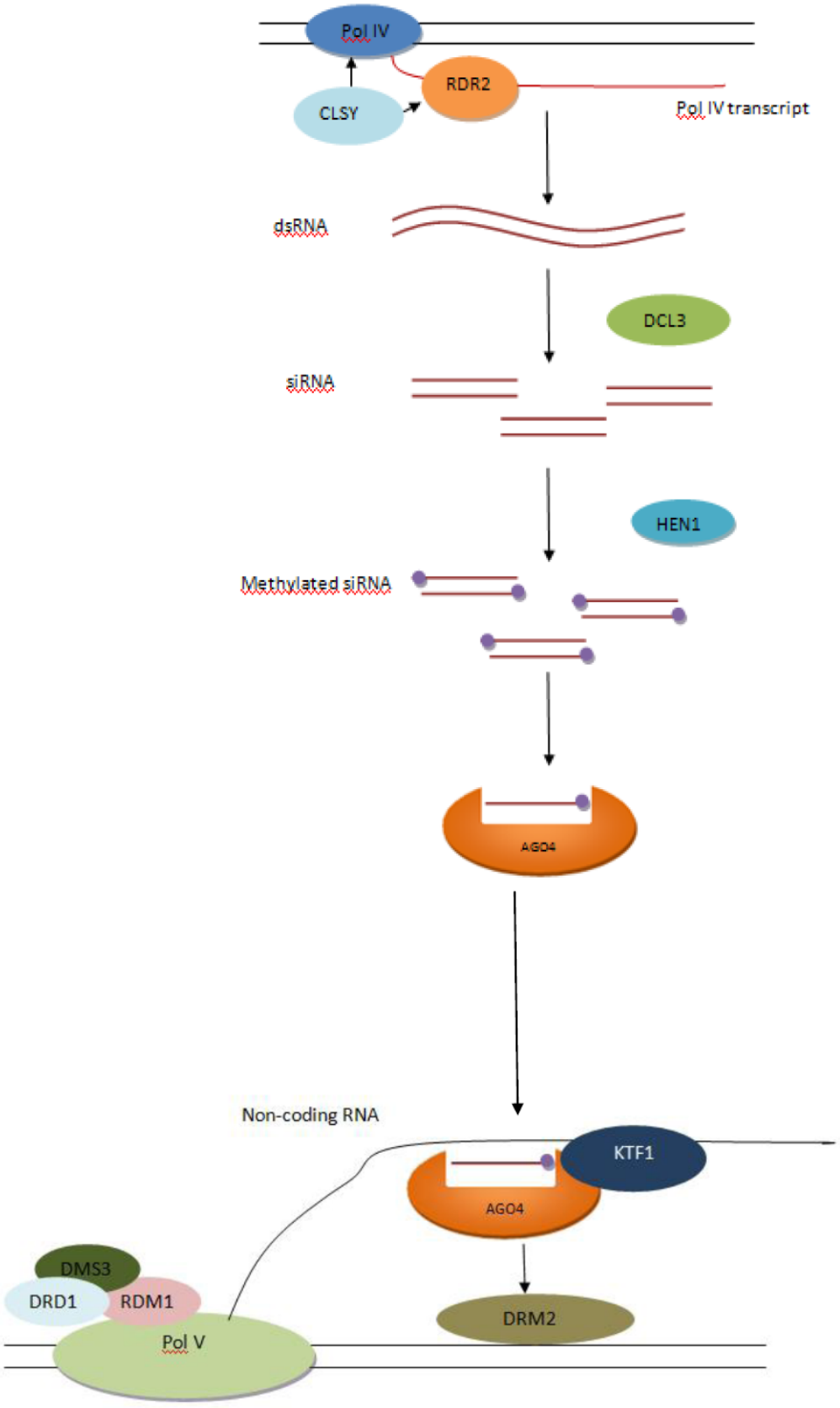


Figure 1.3 Canonical RdDM pathway model. .

1.3 UBA Domains

Ubiquitin is a cellular signaling protein that labels proteins and regulates the cell cycle, DNA repair, and gene silencing. Formally, binding of the Ubiquitin Binding Domain (UBD) to proteins is catalyzed by the actions of several enzymes (E1 for activation, E2 for conjugation, and E3 for ligation) that label such proteins with the Ub mark.[20] Ubiquitin chains are built by isopeptide bonds between the carboxyl group of glycine (residue 76) of ubiquitin and the ε-amino group of another ubiquitin lysine. There are seven possible homotypic linkage types as there are seven lysine residues in ubiquitin (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48 or Lys63). Addition of a single ubiquitin to the targeted protein can change protein activity and localization. [18]

Ubiquitin binding associated (UBA) domains are present in many proteins that play a role in ubiquitin biology. Many UBA domains are involved in protein-protein interactions. They are the largest UBDs as they consist of α-helices. The number of helices in each UBA domain is three helices which contain a conserved MGF motif between helix one and two. There are four classes of UBA domains, which have been concluded after the study of 30 UBA domains. The first one binds to Lys48 linked to polyubiquitin chain preferentially. The second class prefers to bind to Lys63 linked to polyubiquitin chain. The third class does not bind to ubiquitin at all. The fourth class binds to polyubiquitin without any preference to any linkage specificity. [19]

What we try to emphasize in this work is to eliminate each UBA domain one by one (rather than all at once) so we can investigate the role of each UBA domain. In addition, we identify what are the proteins that bind to DRM2 through its UBA domains and we reproduce previous work on mutating the conserved motif (MGF) of UBA domains to disrupt ligation of UBA domain with ubiquitin – to reinforce the conclusions drawn here. The unique contribution of this work is that we use a single mutation of each ubiquitin domain one by one (not all UBA domains in one go).

1.4 Aim of the work

- Identification of proteins that bind to DRM2 via its UBA domains

- Generation of *Arabidopsis thaliana* expression constructs for DRM2 variants with UBA domain deletions

2. Materials

2.1 Kits

Kit	Manufacture
Gateway Cloning	Invitrogen
Plasmid DNA miniprep Purification	A & A Biotechnology
Plant DNA Genomic Extraction	A & A Biotechnology
PCR Clean up	A & A Biotechnology

Table 2.1: List of commercial Kits used

2.2 Enzymes

2.2.1 PCR Enzymes

For the realization of this project three different DNA polymerases (Table 2.2) were used:

A) Phusion PCR polymerase

This enzyme is ideal for amplification of long fragments and it has 5′→ 3′ polymerase activity, and 3′→ 5′ exonuclease activity.

B) Taq DNA polymerase

This enzyme is isolated from an *E. coli* strain that carries the *Taq* DNA Polymerase gene from *Thermus aquaticus* that has 5′→ 3′ polymerase activity and exonuclease activity.

C) PFU polymerase

This is a high-fidelity PCR enzyme, thus with a very low error incorporation. It is used in site directed mutagenesis.

Enzyme	Optimal average amplification size	Manufacture
Phusion PCR polymerase	3.8 kbp	Thermo Fisher Scientific
Taq DNA polymerase	1 kbp	Thermo Fisher Scientific
PFU polymerase	10 kbp	Invitrogen

Table 2.2: List of DNA polymerases used

2.2.2 Restriction Enzymes

Enzyme	Sequence *	Function	Manufacture
Pme1	...GTTT^{V}AAAC... ...CAAA$_{\wedge}$TTTG...	Check insertion of 29 deletion construct in the pGWB620 vector	Thermo Fisher Scientific
Spe1	...A^{V}CTAGT... ...TGATC$_{\wedge}$A...		Thermo Fisher Scientific
Xba1	...T^{V}CTAGA... ...AGATC$_{\wedge}$T...	Check insertion of 55 deletion construct in the pGWB620 vector	Thermo Fisher Scientific
Mfe1	...C^{V}AATTG... ...GTTAA$_{\wedge}$C...	Check insertion of 110 deletion construct in the pGWB620 vector	Thermo Fisher Scientific
Sac1	...GAGCT^{V}C... ...C$_{\wedge}$TCGAG...	Diagnosis for success of site directed mutagenesis on UBA1 domain	Thermo Fisher Scientific
Dpn1	...G^{me}A^{V}TC... ...CT$_{\wedge}$^{me}AG...	Digest methylated DNA	Thermo Fisher Scientific

Table 2.3: List of restriction enzymes used. *: The symbols ^{V} and $_{\wedge}$ represent the sites of restriction cleavage, and ^{me} methylation at the following nucleotide.

2.3 Vectors

Name	Antibiotic resistance	Manufacture	Use
pDONR201	Kanamycin	Invitrogen	Entry clone in Gateway system
pDONR207	Gentamycin	Invitrogen	Entry clone in Gateway system
pGEX-6P-1-DRM2	Ampicillin	Dr Fernandes gift	cDNA template of the DRM2 gene
pGWB620	Spectinomycin	Dr Swiezewski lab gift	Destination vector in Gateway system
pGWB602-Luc	Spectinomycin	Dr Swiezewski lab gift	Destination vector in Gateway system to check expression with luciferase enzyme under a 35S promoter.

Table2.4: List of cloning vectors used

2.4 Oligonucleatide Primers

Name	Function	Product (bp)	Sequence (5'-3')
Drm2fwr_29_pDONR_attB1 Drm2rev_FL_pDONR_attB2	Amplification of deletion construct of DRM2 starting with amino acid no.29 with attB sites	1797bp	5'GGGACAAGTTTGTACAAAAAAGCAGGCTTCATGCAGTGTA GGGTCGAAAATC3' 5'GGGACCACTTTGTACAAGAAAGCTGGGTCAGATCCTCTCA TCCTCGCACGTAC3'
Drm2fwr_55_pDONR_attB1 Drm2rev_FL_pDONR_attB2	Amplification of deletion construct of DRM2 starting with amino acid no.55 with attB sites	1721bp	5'GGGACAAGTTTGTACAAAAAAGCAGGCTTCATGTTAGTTC AGATGGGCTTCTC3' 5'GGGACCACTTTGTACAAGAAAGCTGGGTCAGATCCTCTCA TCCTCGCACGTAC3'
Drm2fwr_110_pDONR_attB1	Amplification of deletion construct of DRM2 starting with amino acid no.110	1516bp	5'GGGACAAGTTTGTACAAAAAAGCAGGCTTCATGTCCAAGT CCAAGGCTATTG3'

Drm2rev_FL_pDONR_attB2	with attB sites		5’GGGACCACTTTGTACAAGAAAGCTGGGTCAGATCCTCTCA TCCTCGCACGTAC3’
Drm2fwr_191_pDONR_attB1 Drm2rev_FL_pDONR_attB2	Amplification of deletion construct of DRM2 starting with amino acid no.191 with attB sites	1311bp	5’GGGACAAGTTTGTACAAAAAAGCAGGCTTCATGTCAAATG AAAATGGCAGC3’ 5’GGGACCACTTTGTACAAGAAAGCTGGGTCAGATCCTCTCA TCCTCGCACGTAC3’
DRM2_mf_f73a DRM2_mr_f73a	Site directed mutagenesis for UBA1	Depends on plasmid	5’GTTTTTGCTACATTGTTTGACATGGAGCTCCTGTTGAGATGATTTCTAGAGCGATC3’ 5’GATCGCTCTAGAAATCATCTCAACAGGAGCTCCCATGTCAAACAATGTAGCAAAAAC3’
Drm2fwr_-30nt_pDONRattB1 Drm2rev_FL_pDONR_attB2	Amplification of drm2 with its UTR region starting with upstream nucleotide -30	3541 bp	5’GGGACAAGTTTGTACAAAAAAGCAGGCTTCCTCAAACTAAAAATTCTCTG3’ 5’GGGACCACTTTGTACAAGAAAGCTGGGTCAGATCCTCTCA TCCTCGCACGTAC3’
Drm2fwr_-126ntpDONRattB1 Drm2rev_FL_pDONR_attB2	Amplification of drm2 with its UTR region starting with upstream nucleotide -126	3637 bp	5’GGGACAAGTTTGTACAAAAAAGCAGGCTTCGTTTCCTCCTCCAGTAAACTG3’ 5’GGGACCACTTTGTACAAGAAAGCTGGGTCAGATCCTCTCA TCCTCGCACGTAC3’
Drm2fwr_-501ntpDONRattB1 Drm2rev_FL_pDONR_attB2	Amplification of drm2 with its UTR region starting with upstream nucleotide -501	4012 bp	5’GGGACAAGTTTGTACAAAAAAGCAGGCTTCGTCATTGCATGCATCCAAAG3’ 5’GGGACCACTTTGTACAAGAAAGCTGGGTCAGATCCTCTCA TCCTCGCACGTAC3’
Drm2rev_FLstop_pDONR_aatB2	Reverse primer for DRM2 with stop codon and att-B site		5’GGGACCACTTTGTACAAGAAAGCTGGGTCTTAAGATCCTC TCATCCTCGCACGTAC3’
Drm2fwr_195_pGEX6p1 Drm2rev_244_pGEX6p1	Amplification of DRM2 insert to detect insertion in pDONR 201	144 bp	5’TAGCTAGGATCCAAAATACGGTCTTTGGTGAAG3’ TAGCTAGCGGCCGCGTAAAACTCACTAAATTCC3’

Figure 2.5: Primers used in this work; attB sites are shaded, start codons underline and mutated codons are red colored.

2.5 Bacterial strains

A) *Agrobacterium tumefaciens*

This strain is a soil borne bacteria that causes growth of gall disease in plants. This occurs by inserting a piece of bacterial DNA (T-DNA) into the plant's genome. T-DNA is located in the tumor inducing plasmid (Ti-Plasmid) and bordered on two sides by special DNA sequences called T-DNA borders. The strain used in this work (GV3101) is disarmed from tumor inducing genes but keeps the capability to transform T-DNA or whatever is engineered within the T-DNA border regions.

B) DH5α

This strain of *E. coli* was developed in the laboratory, for laboratory cloning procedures. This strain is not pathogenic and contains multiple mutations that enable high-efficiency transformations.

C) Top 10

This strain is acquired from Thermo Fisher Scientific. It has enhanced genomic DNA cloning capabilities. It is marked with high efficiency to clone methylated DNA.

2.6 Plant Material

Plant	**Source**
Arabidopsis thaliana (Col-0)	Dr Swiezewski lab gift
Arabidopsis thaliana ΔDRM1/DRM2	NASC (The European Arabidopsis Stock Centre) [Stoke number N16383]
Nicotiana benthamiana WT	Prof Jerzmanowski lab gift
Nicotiana tabacum WT	Dr Swiezewski lab gift

Table 2.6: Plant seeds used

2.7 Culture Medium

A) Bacterial Medium

The following mediums were prepared by the IBB technical team

- Liquid Broth (Tryptone, Yeast Extract, NaCl)
- SOC medium (Tryptone, Yeast Extract, NaCl, KCl)
- YEB medium (Beef extract, Yeast extract, Peptone, Sucrose, $MgSO_4$, Distilled water, Agar)

B) Plant Medium

MS solid for seedling growth

0.25% Murashige and Skoog

1% sucrose

0.6% Agar

pH = 5.7

MS liquid medium for agro-seedling transformation (co-cultivation media)

0.25% Murashige and Skoog

1% sucrose

100 μM acetosyringone, 0.005% ➔ 50 μl added fresh to each 20 ml of medium

Silwet L-77 ➔ 3 μl added fresh to each 20 ml of medium

2.8 Antibiotics

Name	**Stock Concentration (solvent)**	**Final concentration**
Kanamycin	50 mg/ml (H_2O)	50 μg/ml
Gentamycin	10 mg/ml (H_2O)	10 μg/ml

Rifamycin	50 mg/ml (MetOH)	100 μg/ml
Spectinomycin	100 mg/ml (H_2O)	100 μg/ml
Ampicilin	100 mg/ml (H_2O)	100 μg/ml

Table2. 7: Antibiotics used in selection medium

2.9 Antibodies

Name	**Dilution**	**Host**	**Manufacture**
Anti-myc	1/5000	Rabbit	Sigma
Anti rabbit	1/1000	Rabbit	Sigma
Anti-mouse	1/1000	mouse	Abcam

Table 2.8: list of antibodies used on Western blots

2.10 Agarose gel

2.10.1 Agarose gel electrophoresis

Agarose gel electrophoresis is a technique used to separate and analyze DNA. The gel might be used to look at the DNA quality, quantify, and/or to isolate a particular band. The DNA is visualized in the gel by addition of GelRed ™ and exposure to UV light.

1% Agarose gel
0.5 gm of Agarose
50 ml of TBE buffer

2.10.2 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

This is a method of gel electrophoresis to separate denatured proteins based on the the mass of the given protein. The proteins are dissolved in a solution of sodium dodecyl sulfate (SDS), a detergent that dissociates the secondary and tertiary interactions between

proteins, and charges them negatively. The largest molecules move slowly through the gel while smaller molecules move faster and result in bands near the end of the gel.

10 ml of 12% SDS-PAGE gel

4.5 ml of 40 % Acrylamide/Bisacrylamide

4.5 ml of Tris pH 8.8

150 µl of SDS 10%

90 µl of APS 10%

7.5 µl of TEMED

2 ml of Stacking gel

0.28 ml of 40 % Acrylamide/Bisacrylamide

0.26 ml of Tris pH 8.8

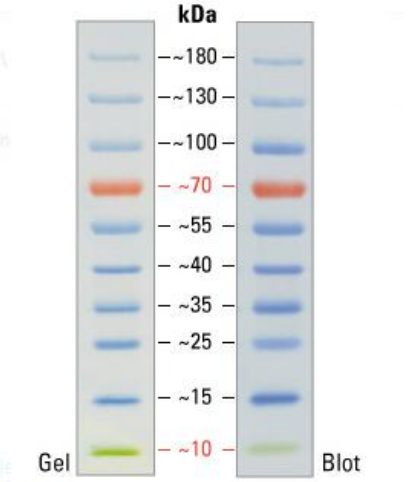
50 µl of SDS 10%

25 µl of APS 10%

7.5 µl of TEMED

1.46 ml H_2O

PageRuler™ Prestained Protein Ladder, 10 to 180 kDa



SDS-PAGE band profile of the PageRuler Prestained Protein Ladder
Images are from a 4-20% Tris-glycine gel (SDS-PAGE) and subsequent transfer to membrane.

2.11 Buffers

Cold H buffer

50 mM HEPES

0.25 M Sucrose

1 mM EDTA

Total Protein Extraction buffer 1

0.1 M EDTA, pH 8

0.12 M Tris-HCl pH 6.8

4% W/V SDS

10 V/V β-mercaptoethanol

5% v/v glycerol

0.005% w/v bromophenol blue

Total Protein Extraction buffer 2

2 mM EDTA, pH 8.0

50 mM Tris-HCl, pH 7.2

3 mM β-mercaptoethanol

0.1% Triton X-100

150 mM $NaCl_2$

SDS-PAGE running buffer 10x

248 mM Trizma Tris base

1.92 M Glycine

1% w/v SDS

Protein sample loading buffer (5X)

0.25% bromophenol blue

0.5M DDT

Glycerol 50%

SDS 10%

0.25 M Tris-HCl

Final pH = 6.8

PBS 10x

100 mm $Na_2HPO_4.7H_2O$

1.37 m NaCl

27 mm KCl

18 mm KH_2PO_4

PBST 1 L

1 L PBS

1 ml Tween 20

Blot transfer buffer 1L

2.5 g Glycine

58 g Tris

4 ml SDS 10%

200 ml Ethanol

Chemiluminescence developing solution

1 ml of 1 M Tris pH 7.4

50 µl Luminol

20 µl Communic acid

Up to 10 ml of MilliQ water

10 µl of H_2O_2 added at the moment of capturing

Autoclaved $CaCl_2$ buffer for *E. coli* competent cell preparation

60 mM $CaCl_2$

10 mM PIPES

15% Glycerol

Final pH = 7.0

Washing solution for *Agrobacterium tumefaciens*

10 mM $MgCl_2$

100 µM Acetosyringon

5X stock solution of TBE in 1 L of H_2O

54 g of Tris base

27.5 g of Boric acid

20 mL of 0.5 M EDTA (pH 8.0)

2.12 Bioinformatics

Name of Program	Usage	Source
NCBI Blast	Sequence Alignment	http://blast.ncbi.nlm.nih.gov/Blast.cgi
APE	Sequence analysis	http://biologylabs.utah.edu/jorgensen/wayned/ape/
NEB Tm Calculator	Tm Calculation	http://tmcalculator.neb.com/#!/
Protein Molecular Weight	Calculating Protein Molecular Weight	http://www.bioinformatics.org/sms/prot_mw.html

Table 2.9: list of programs used

2.13 Sources of UBA domains

Name	Accession number
Chain A, Crystal Structure Of The Mud1 Uba Domain	1Z96_A
Chain A, Solution Structure Of The Uba Domain Of Human Tudor Domain Containing Protein 3	1WJI_A
Chain A, Solution Structure Of Rsgi Ruh-027, A Uba Domain From Mouse Cdna	1WHC_A

3 Methods

3.1 DNA cloning methods

3.1.1 Amplification of DNA fragments via polymerase chain reaction (PCR)

PCR is used to amplify specific fragments of DNA and make them ready to be later cloned, and –in this project- transformed into plants. To perform PCR, the following components are necessary: (1) a DNA template that contains the targeted fragment, (2) nucleoside triphosphates (dNTPs), which are the nucleic acid units necessary for building new copies of the desired fragment, (3) the oligonucleotide primers (the pair of oligonucleotide sequences that work as a start and initiation molecule for the DNA polymerization) and (4) the DNA polymerase enzyme.

To make a full DNA polymerization reaction, there are three phases: First) Denaturation where the temperature varies between 94^{o}C to 98^{o}C (in this project) to split the two strands of DNA. Second) Annealing, which depends on the melting temperature (Tm) value of the oligonucleotide primers and which, in turn, depends on its sequence, where these oligonucleotide sequences attach to the single-stranded DNA (ssDNA). Third) Extension, at this step, the DNA polymerase enzyme binds to primers and starts to recruit dNTPs along ssDNA, where the enzyme that is used will depend on the fragment one intends to amplify. Analysis of PCR reactions can be performed using agarose gel electrophoresis, and by the addition of dye, visualized under UV light.

A) Large fragments by Phusion DNA Polymerase

Component	**Amount**
HF buffer 10x	10 µl
dNTPs 10 mM	1 µl

Forward Primer 10 µM	2.5 µl
Reverse Primer 10 µM	2.5 µl
Template DNA 100 ng/µl	0.5 µl
Phusion DNA polymerase enzyme	0.5 µl
Water	33 µl

Table 3.1: Components of Phusion PCR reaction

98^{o}C	30 seconds	
98^{o}C	10 seconds	5 Cycles
55^{o}C	30 seconds	
72^{o}C	2 minutes	
98^{o}C	10 seconds	30 Cycles
72^{o}C	30 seconds	
72^{o}C	2 minutes	
72^{o}C	10 minutes	
4^{o}C	Until storage	

Table 3.2: Phusion PCR program

B) Small Fragments Taq DNA polymerase

Component	**Amount**
Buffer 10x	10 µl
dNTPs 10 mM	1 µl
Forward Primer 10 µM	2.5 µl
Reverse Primer 10 µM	2.5 µl
Template DNA 100 ng/µl	3.75 µl
Taq DNA polymerase enzyme	0.25 µl
Water	30 µl

Table 3.3: Components of Taq PCR reaction

95°C	3 minutes	
95°C	30 seconds	35 Cycles
69°C	30 seconds	
68°C	20 seconds	
68°C	10 minutes	
4°C	Until storage	

Table 3.4: Taq PCR program

C) Mutagenesis by PFU Polymerase

Component	**Amount**
Buffer 10x	5 µl
dNTPs 10 mM	1 µl
Forward Primer 10 µM	0.5 µl
Reverse primer 10 µM	0.5 µl
Template DNA 100 ng/µl	0.5 µl
PFU turbo DNA polymerase enzyme	0.5 µl
Water	42 µl

Table 3.5 : Components of the mutagenesis reaction

95°C	2 minutes	
95°C	30 seconds	30 Cycles
65 °C	45 seconds	
68 °C	7 minutes	
68 °C	4 minutes	
4 °C	Until storage	

Table 3.6: Mutagenesis cycles program

3.1.2 Chemically Competent *E. coli* DH5α cells preparation and transformation

A) Preparation

Streak cells of *E. coli* DH5α on a plain LB agar plate without any antibiotics and incubate at 37^{o}C overnight. Pick a single colony with an edge of a tip to inoculate 30 ml of liquid LB without antibiotics then incubate at 37^{o}C with shaking 200 RPM overnight. Use 2 ml of overnight culture to inoculate 200 ml of LB without antibiotics then incubate at 37^{o}C with shaking 200 RPM until Optical Density (OD) at 600 nm reaches 0.6 (check every hour). Place four 50 ml tubes on ice. When the OD_{600} reaches 0.6, pour 200 ml into pre-chilled tubes and incubate on ice for 30 minutes. Spin with maximum speed at 4^{o}C for 10 minutes then remove supernatant. Resuspend each cell pellet in a 50 ml tube with 15 ml of sterile cold $CaCl_2$ by smooth pipetting. Incubate cells on ice for 15 minutes. Spin with maximum speed at 4^{o}C for 10 minutes. Resuspend each cell pellet in each 50 ml tube with 4 ml of sterile cold $CaCl_2$ by smooth pipetting. Aliquot cells into 400 μl aliquots in 1.5 ml tubes. Leave aliquots on ice for 1 hour. Freeze in liquid nitrogen. Place tubes in racks and freeze in -80^{o}C

B) Transformation by thermal shock

Add 1-2 μl of DNA (120 ng/μl is needed) to 50 μl chemically competent *E. coli* DH5α. Place on ice for 30 minutes. Heat shock cells at 42^{o}C for 1 minute. Place again on ice for 2 minutes. Add 450 μl of LB to cells. Incubate at 37^{o}C for 1-2 hours. Take 100 μl of cells and plate on LB agar with antibiotic selection. Keep plate(s) at 37^{o}C overnight and the next day successfully transformed cells with resistance in transformed plasmid will grow. Take by an edge of a tip a single colony and propagate in 5 ml LB with antibiotic selection at 37^{o}C with 200 RPM shaking overnight. From this point, transformed cells are ready for plasmid isolation.

3.1.3 Preparation, transformation of electro competent Agrobacterium cells

A) Preparation

Streak *Agrobacterium tumefaciens* GV3101(pMP90RK) on LB plate with antibiotics (Gentamycin and Rifampycin). Grow at 28^{o}C for 2 days. Place several colonies in 100 ml LB media with antibiotics then grow overnight at 28^{o}C. Transfer 50 ml cell culture into fresh 500 ml LB media in a 2.5 L flask. Let cells grow at 28^{o}C with 200 rpm shaking until optical OD_{600} reaches 1 (after 4 hours). Transfer cell culture to 50 ml cold tubes and place them on ice for 20 minutes. Centrifuge cells at maximum speed under 4^{o}C for 15 minutes. Pour off media completely and resuspend cells in each tube completely but gently with 50 ml of cold sterile water. Centrifuge cells at maximum speed under 4^{o}C for 15 minutes. Repeat steps 8-9 four times. Pour off media and resuspend cells in each tube with 5 ml of cold sterile 10% glycerol. Combine suspended cells from each tube into one cold 50 ml tube. Centrifuge cells at maximum speed under 4^{o}C for 10 minutes. Pour off media and finally resuspend cells completely but very gently in 2 ml of cold sterile 10% glycerol. Divide into 100 µl aliquots and freeze cells in liquid nitrogen. Store cells at -80^{o}C. They can be used for at least 6 months.

B) Electro-transformation

Dilute plasmid to 15 ng/µl. Mix 2 µl of diluted plasmid with 50 µl of *Agrobacterium tumefaciens* in a 1.5 ml Eppendorf tube. Transfer mixture to sterile electroporation cuvette that has been chilled for 20 minutes on ice. Set the electroporator to 2.00 V. Snap cuvette into proper orientation, push in and hold both pulse buttons until it beeps (it usually takes 2 seconds). Add 1mL LB or ice cold SOC media to the cuvette and pipette up and down to mix. Then transfer to tube. Incubate at 28^{o}C for 2 hours shaking then streak suspension on LB agar to make single colonies. Incubate plates at 28^{o}C for two days. Take a single colony and spread it with edge of tips along the whole fresh LB agar

plate. Incubate plates at 28^{o}C for two days. *Agrobacterium tumefaciens* is now loaded with plasmid and ready for plant transformation

3.1.4 Gateway Cloning

It is a universal system for cloning DNA sequence into multiple destination plasmids and (in this project) it is used for protein expression in plants. It uses a lambda phage *in vivo* system *in site*-specific recombination to clone DNA *in vitro*. The crucial sequence for DNA recombination is the att sites.

The reaction aim is to insert the DNA fragment of interest with flanking att-B sites into the donor vector with att-P sites to generate an entry clone with att-L sites. This reaction is called a BP reaction.

The next reaction is called a LR reaction where the DNA sequence of interest is transferred from the entry clone to the destination vector to obtain the expression clone. This expression clone is loaded with the DNA sequence of interest flanked with an att-L sites and that allows expression in the desired host (i.e., in this case, plants).

For Gateway cloning, the steps have been conducted according to the manufacturer's kit protocol, but the reaction volumes were adjusted as presented in tables 3.7 and 3.8

Element	**Amount**
Purified PCR product 100 ng/µl	1.5 µl
pDONR201 plasmid 190 ng/µl	0.51 µl
BP Clonase enzyme	1 µl
TE buffer	7 µl

Table 3.7: Gateway BP reaction mix

Element	**Amount**
Entry clone (pDONR201 with DRM2 deletion constructs) 150ng/ µl	1 µl
Destination vector (pGWB620) 150 ng/µl	1 µl

LR Clonase enzyme	1 µl
TE buffer	7 µl

Table 3.8: Gateway LR reaction mix

3.2 Plant transformation

3.2.1 Plant growth conditions

A) *Arabidopsis thaliana* and *Nicotiana benthamnia* whole plant

Seeds were sterilized by 70% ethanol for 10 minutes then washed with 100% ethanol. Then the seeds are left in cold water in the fridge for 4 days to break dormancy. A mixture of clay soil and perlite is prepared in $7cm^2$ pots. Two seeds were added to each pot. Place pots in the green house with conditions of 16 hours of light and 8 hours of dark (Irrigation system worked each 3 days so manual care was required)

B) *Arabidopsis thaliana* seedlings

Seeds were sterilized by 70% ethanol for 10 minutes then washed with 100% ethanol. Seeds plated in MS medium with vitamins and 1% sucrose with enough space between each seed. Plates were left in growth chamber for 7-10 days (16 hours of light and 8 hours of dark)

3.2.2 Agro-mediated plant transformation

Directly scratch Agrobacterium from fresh plate and wash in washing solution. Resuspend it in co-cultivation media at final $OD6_{00}$ of 0.35. In clean plant dish (6 wells X 5cm diameter) place 30 Arabidopsis seedlings in each well. Add 4 ml of co-cultivation medium and 1 ml of freshly prepared electro transformed agrobacterium. Leave the plate in total darkness in growth chamber for 40 hours. Remove media with micropipette and add 5ml of 1% bleach and leave for 10 minutes. Wash with MilliQ water two times. Seedlings are now ready for protein expression analysis

3.2.3 Luciferase assay for Arabidopsis seedlings

After co-cultivation, seedlings are surface sterilized with 0.1% bleach for 10 minutes. Wash twice in sterile ddH_20. Transfer seedlings on MS solid plates and, in the dark, spray with a LUC-substrate for imaging. Subject the plate for LUC imagining using NightSHADE LB 985 *in vivo* Plant Imaging System

3.3 DNA/ Protein Analyses

3.3.1 DNA plasmid purification

To purify the DNA plasmid from the bacterial cells, the steps as described in the manufacturer's protocol (Table 2.1) were followed.

3.3.2 PCR cleaning

A) PCR cleaning kit

Cleaning PCR products was done following the manufacturer's protocol (Table 2.1)

B) PEG/$MgCl_2$ purification method

Add 120 µl of TE buffer to 40 µl of PCR product. To that mixture add 80 µl of 30% PEG 8000/ 30mM $MgCl_{2.}$ Vortex and centrifuge to maximum speed for 5 minutes. Remove supernantant. Resuspend pellet in TE buffer 20 µl. Preserve in Eppendorf tube at -20^{o}C.

3.3.3 Plant Genomic DNA Extraction

To purify genomic DNA from *Arabidopsis thaliana*, the steps as described in the manufacture's protocol of the kit (Table 2.1) were followed.

3.3.4 DNA concentration measurement

Switch on Nanodrop Spectrophotometer device. Open the lid and clean the socket with 2 μl of MilliQ water two times. Add 2 μl of DNA sample buffer or MilliQ water (depending on DNA solvent). Close the lid and press Measure blank. Open the lid again and clean with tissue. Load 2 μl of DNA samples. Close the lid and press measure. Take the read from the screen.

3.1.5 Plant total protein extraction

Weigh the plant sample. Place sample in pre-chilled mortar and liquid nitrogen solution. Smash sample with mortar pestle. Add 3 volumes (v/w) of cold protein extraction buffer and keep smashing. Transfer solution into Eppendorf tube. Add 1/5 of loading sample (5X) to mixture in Eppendorf. Boil mixture at 95^{o}C for 5 minutes and let it cool to be ready for SDS-PAGE analysis.

3.1.6 Gel electrophoresis

A) Agarose gel electrophoresis

After preparing the 1% Agarose gel, the first wells were loaded with 3 μl of DNA molecular ladder 100 bp. The following wells were loaded with DNA samples mixed

with an appropriate amount of the Loading dye and Gel Red. The gel was run at 160 V for 45 minutes. The fragments were visualized using UV light.

B) Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Assemble the tray component of protein unit. Pour off the resolving gel between the two combined glasses. To make the resolving gel even, add water by micropipette. After the polymerization of the polyacrylamide gel, pour off the water outside unit. Add stacking gel very carefully. Place the comb to form wells so the protein samples can be loaded. Remove the comb very carefully and add two glasses to the Electrophoresis unit. Add the running buffer then load the marker and samples. Adjust the power supply to 210 V after connecting the cathode and anode cables. Leave it for 1 hour and 30 minutes. Remove the gel and place it in coomassie blue for staining or place in the transfer blotter device.

3.1.7 Western Blot

Place 3 filter papers in the device and add a transfer buffer to cover it. Put polyacrylamide gel on top of the three filter papers. Gently place nitrocellulose membrane on top. Put three more filter papers above the nitrocellulose membrane. Add the transfer buffer to cover whole sandwich and eliminate air bubbles using a round glass rod. Combine the second metal part of blotter and connect wires on the appropriate direction. Run power supply at 2 amps/$cm^{2.}$ After transfer for 90 minutes, carefully remove the nitrocellulose membrane. Wash the membrane with PBST for 5 minutes 4 times. Add the first antibody diluted in a 5% skim milk solution. Incubate for 2 hours. Remove the milk solution and wash with PBST for 5 minutes four times. Add the second antibody diluted in milk solution and incubate for 2 hours. Remove the milk solution and wash with PBST. Add developing solution. Adjust increment to take chemiluminescent picture each 5 seconds. After taking chemiluminescent pictures, take a normal picture for Ladder visualization.

4. Results and Discussion

Existence of UBA domains in DRM2 has been demonstrated by alignment of Sequences for other UBA domains in other organism (check 2.13). (figure 4.A)

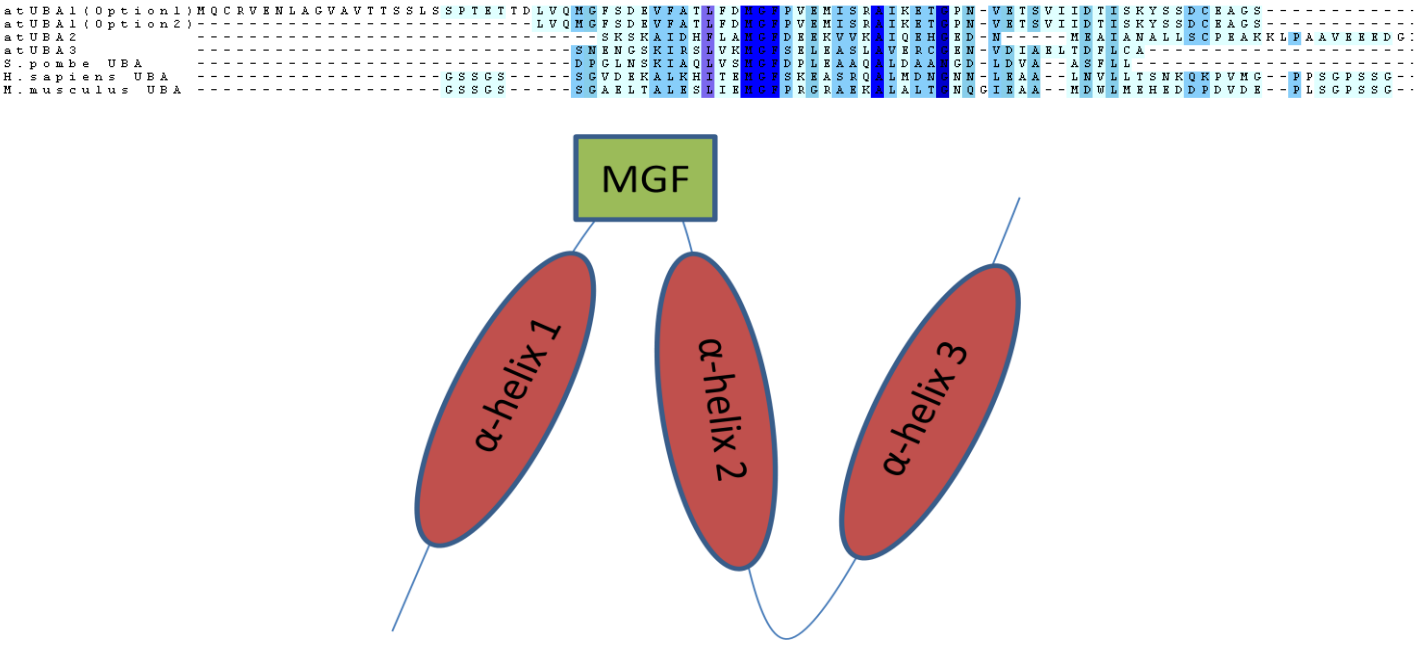


figure 4.A Alignment of atUBA domains and its secondary structure prediction

Also to check the common secondary structure of the UBA domains, I used the Psipred program and it predicted successfully its three alpha helices with the MGF conserved motif between helix 1 and 2 (see appendix for more illustration).

After predicting the secondary structure, it was noticeable that there are two options for the UBA1 domain. The first option is to include residues downstream from residue 29. At this point, there is the MGF motif and the three helices. The second option for the UBA1 domain is the sequence starting from residue 55.

4.1 Characterization of deletion constructs

The cDNA of DRM2 was obtained from Dr. Fernandes. By running att-B primers,

the forward and reverse truncated versions of DRM2 cDNA were synthesized as follows.

.

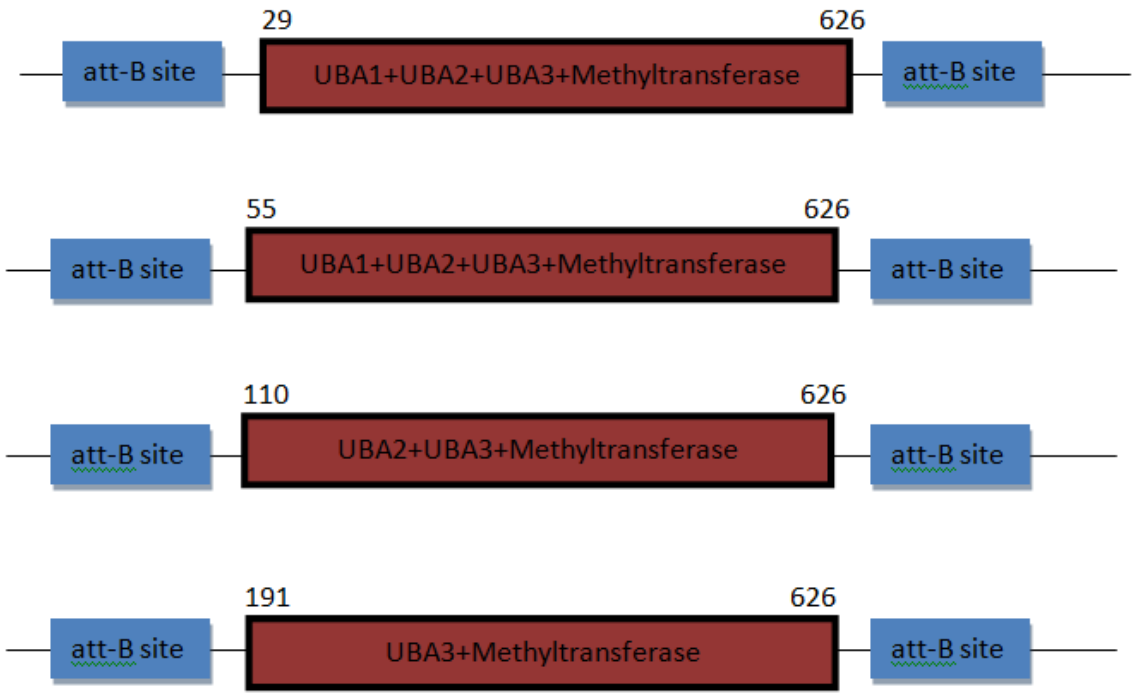


Figure 4.B PCR products of att-B primers

These truncated versions of DRM2 include different lengths of UBA domains to check what the role of each UBA domain is respectively. Numbers present amino acid residues where each sequence starts and ends.

PCR products of att-B primers which include truncated versions of atDRM2 were purified from dNTPs and put into the gateway cloning system to end up with four deletion constructs loaded on pGWB620 plasmid.

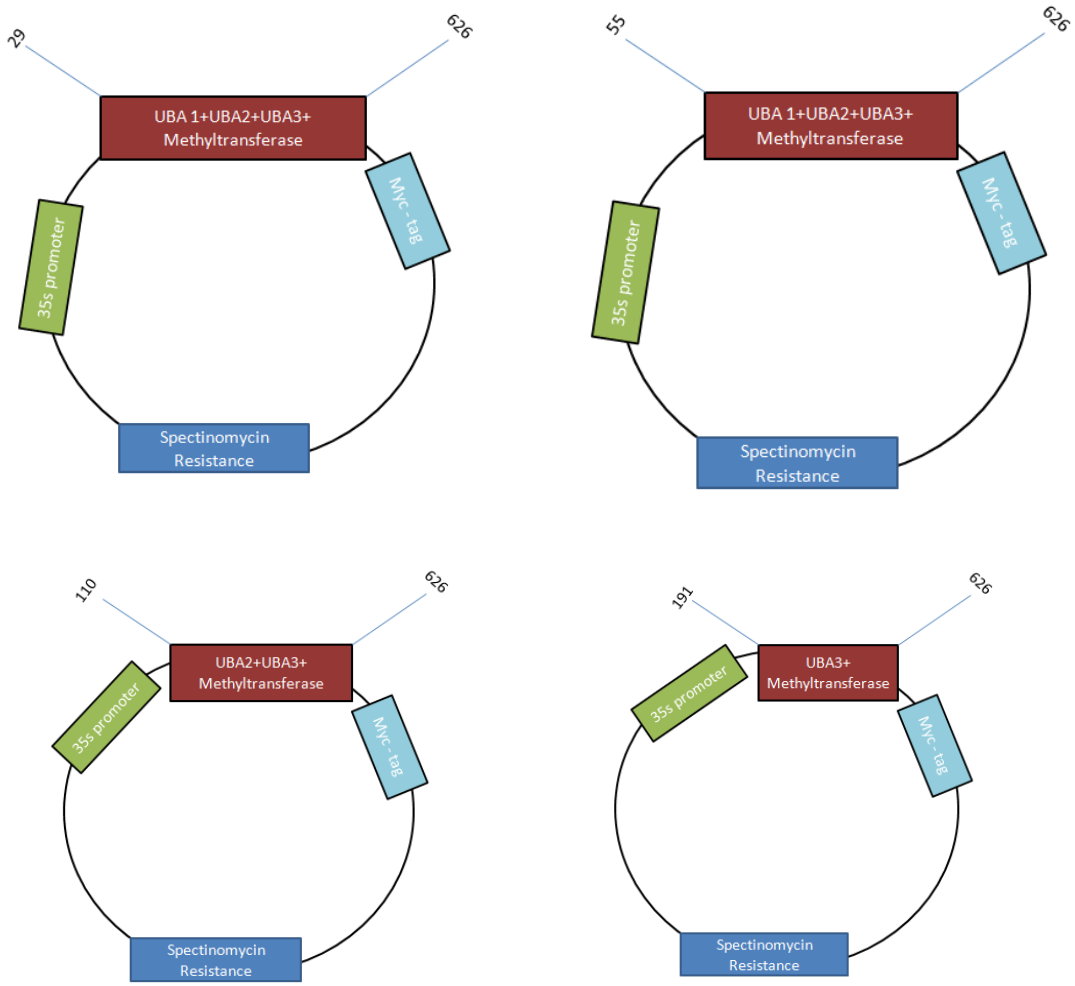


Figure 4.C Final deletion constructs loaded on pGWB620 plasmid

These deletion constructs have the 35s promoter for transcription initiation at the N-terminus. At the C-terminus, there are 10x myc tags for immunoblot detection followed by Tnos Termination to end transcription.

4.2 Screening DRM2 deletion constructs

After obtaining cDNA of DRM2 from Dr. Fernandes, the coding regions for different constructs were amplified by PCR and with the incorporation of the att-B sites (figure 4.1). Later these fragments will be used for gateway cloning.

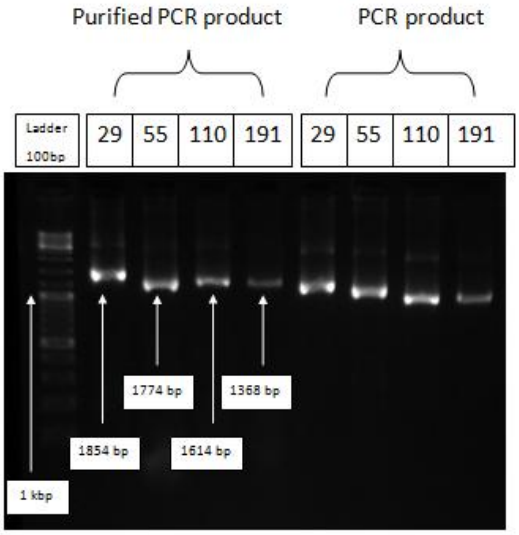


Figure.4.1: 1.5 % agarose gel of the PCR amplification, and purification, of the different DRM2 constructs. First lane 100bp ladder followed by purified PCR products (4 lanes). Last 4 lanes are for PCR products.

To generate different constructs of AtDRM2 with N-terminal deletions of its UBA domains, primers have been designed to amplify sequences starting from the beginning of the different UBA domains (Table 2.5) and with merging of att-B sites, for gateway cloning. All constructs were finished with the genuine stop codon of the myc-tag protein after the catalytic domain. The different deletion constructs start with

1) Residue No. 29, thus harboring the UBA1 (option 1), UBA2, and UBA 3 domains. Length of the product is 1854 bp.
2) Residue No. 55, thus harboring the UBA1 (option 2), UBA2, and UBA3 domains. Length of the product is 1774 bp.
3) Residue No. 110, thus harboring the UBA2 and UBA3 domains. Length of the product is 1614 bp.
4) Residues No. 191, thus harboring the UBA3 domain. Length of the product is 1368pb.

The att-B sites are built-in on the PCR products for site-specific recombination, thus allowing cloning on entry vectors and posterior swapping of the deletion construct from those vectors to different destination vectors. Actually, this is one of advantages of gateway cloning; to be able to switch and swap sequences of interest between different vectors. This opens the possibility of, for example, quickly changing a protein tag, or the promoter type used if necessary.

The amplification of deletion constructs through PCR nonetheless leaves byproducts that inhibit the gateway cloning. Thus, they needed to be cleaned of dNTPs and oligonucleotides. The PEG/$MgCl_2$ method gave clearer products compared to the products obtained directly obtained from the kit. Therefore, PEG/$MgCl_2$was used as the cleaning step on this protocol. After the PCR products of deletion constructs and purification steps were completed, the samples were analyzed using agarose gel electrophoresis. The bands showed the success of generating deletion constructs, and their purity (Figure 4.1).

4.3 Site directed mutagenesis

PCR was conducted to single mutate UBA1 on pDONR 201 at its conserved MGF motif. To investigate the function of the UBA1 domain, a point mutation was introduced to stop the activity of the UBA1 domain. This happened by changing amino acid number 73 from phenylalanine to alanine, which was predicted to render the UBA1 domain unable to bind to ubiquitin. This mutagenesis was performed directly on the pDONR201 vector harboring the AtDRM2 deletion construct no 55.

4.3.1 Restriction digest

Checking the success of the point mutation of the UBA1, the Sac1 enzyme was used to identify the results of the single point mutation (Figure 4.2).

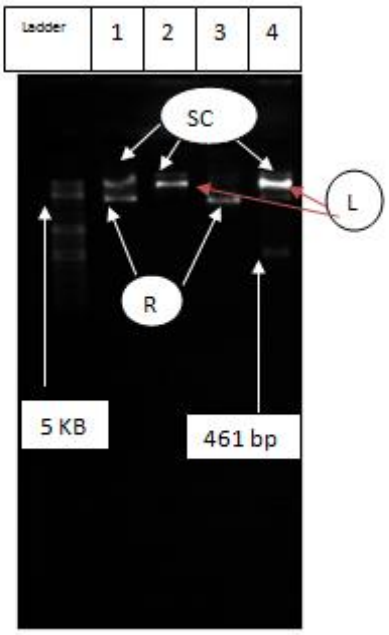


Figure 4.2: Restriction analysis visualized on a 1.5% agarose gel. first lane 100 bp Ladder 1) pDONR201.55wild. 2) pDONR201.55wild with Sac1. 3) pDONR201.55 mutant. 4) pDONR201.55 mutant after Sac1 digestion.

The single point mutation was introduced to pDONR201-55 deletion construct. This point mutation also introduced a Sac1 restriction site, and made the pDONR201-55 plasmid carry two cleavage sites, instead of the unique one in the previous mutagenesis. The pDONR 201-55 deletion was constructed as 5881 bp and showed two forms of plasmid in which the upper one is supercoiled and the lower is relaxed. When Sac1 was applied to the wild construct (pDONR 201- 55 deletion construct) it makes the construct linear, which lies somewhere between the relaxed and supercoiled form. There is also a faded band, which is suggested to be in the supercoiled form of the wild type construct. The pDONR 201-55 deletion construct with the mutation in UBA1 domain migrates

normally and shows only the relaxed form of plasmid. As previously mentioned, the construct was exposed to the Sac1 restriction enzyme, and this was cut into two bands with one fragment weighing 5420 bp and the other 461 bp. It is obvious that there are three forms of the plasmid: supercoiled (SC), linear (L) and relaxed (R). The only spliced fragment that can be seen is the 461 bp fragment.

4.4 Gateway cloning for deletion constructs

4.4.1 BP reaction

After purifying the PCR products, the BP reaction was carried out to load the truncated versions of DRM2 on pDONR201, which is a gateway entry plasmid and has antibiotic resistance for Kanamycin. This resistance allows only bacteria with pDONR201 to grow on the selective media. Later, a single colony was selected and enriched in LB for DNA extraction to obtain the entry clone with four deletion constructs loaded on pDONR201.

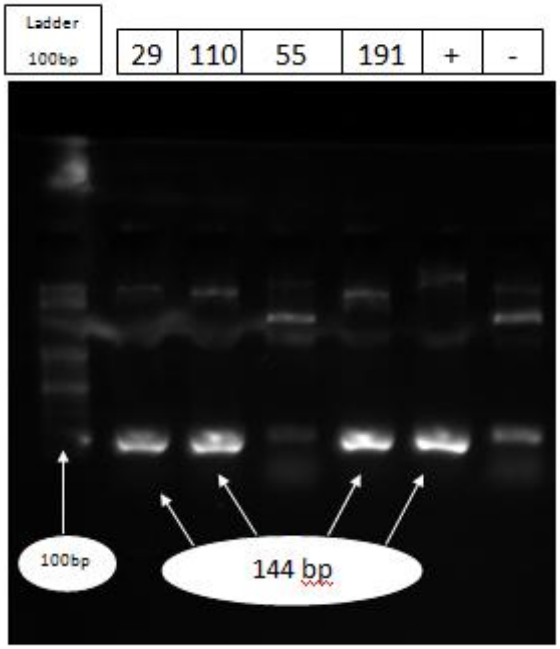


Figure 4.3: Colony PCR and 1.5 % agarose gel analysis of the entry vector harboring the different DRM2 constructs. First lane is 100bp ladder followed by PCR products of inserted truncated version of DRM2 in pDONR201. Last two lanes are positive and negative control

The incorporation of the deletion constructs was confirmed by PCR analysis using as templates the isolated plasmids and DRM2 specific primers. All bands are equal in molecular weight, which is 144 bp (when visualized on agarose gel) due to the usage of the same set of primers for the four plasmids/constructs (Figure 4.3). These primers were annealed at the end of the atDRM2 gene. It is of notice that, construct No. 55 does not show a band with the expected molecular weight. One possibility is the existence of contamination of growing bacteria on the selective media. Therefore, the deletion construct No. 55 was redone. A new PCR product was prepared and used for the BP reaction.

After repeating the BP reaction for construct 55, the Xba1 enzyme was applied to verify the No. 55 deletion construct (Figure 4.4). This enzyme cuts the DRM2 gene in the middle of UBA1 and at the end of the DRM2 gene, generating two bands with expected sizes of 4271 bp and 1610 bp. The first lane shows two forms of plasmid that are supercoiled or relaxed. The second lane also shows two forms of plasmid but there is a fragment which is not visible (1610) because of its low intensity. The successful cut shows that deletion construct No. 55 is ready as an entry clone. To confirm the integrity of the constructs, the samples were sent for sequencing to Genomed Sequencing center (Warsaw, Poland). After sequencing, the BLAST program was used for alignment.

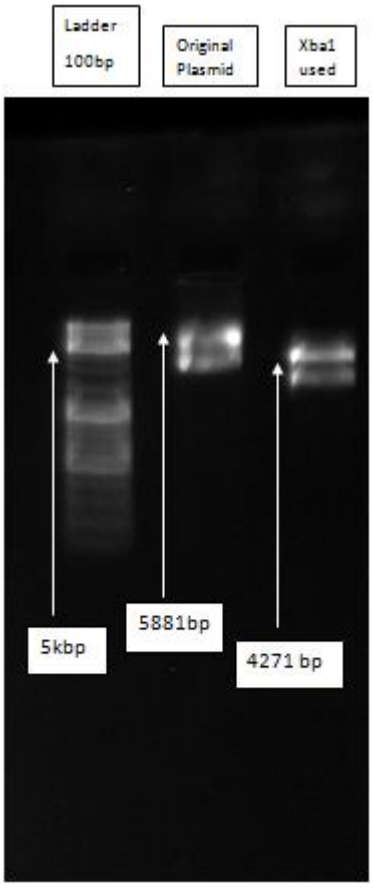


Figure 4.4: XbaI restriction analysis visualized on a 2 % agarose gel. The first lane is a ladder, the second lane is the 55 deletion construct of DRM2 loaded on pDONR201 without using restriction digest, the third lane is 55 deletion construct of DRM2 loaded on pDONR201 after using Xba1 restriction enzyme

4.4.2 LR reaction

After the sequence confirmation, entry clones with deletion constructs are ready for LR reaction. Deletion constructs were recombined from pDONR201 to the destination vector pGWB620 by the help of att-sites and LR clonase. The pGWB620 has a 35s promoter to increase the level of transcription from the inserted sequence. In addition, it has a myc-tag at the end of the att-B site to allow detection of the inserted sequence's protein using immunoblot precipitation. The antibiotic selection for this plasmid is spectinamycin. A single colony was isolated and prepared for DNA extraction yielding expression clones

loaded with four deletion constructs of DRM2. Deletion constructs 191 did not grow on the DH5α strain, so the TOP10 strain was used as alternative.

4.4.3 Restriction digest

To avoid mislabeling, and as a way of fast economical validation, a restriction digest was carried out on the destination vectors harboring the DRM2 constructs. Pme1+Spe1, Xba1, and Mfe1 were used to check the successful insertion of deletion constructs 29, 55, and 110, respectively (Figure 4.5).

1) 29pGWB620 as 12080 bp (A) and when it was cut by both enzymes Pme1+ Spe1, it generated two fragments which had lengths of 10277 bp (B) and 1162 bp (C)

2) 55pGWB620 as 12005 bp (D) and when it was cut by the Xba1 enzyme, it generated three fragments which had lengths of 10277 bp (E), 1610 bp (F) and 118 bp (not visible due to low intensity) .

3) 110pGWB620 as 11840 bp (G) and when it was cut by Mfe1 enzyme, it generated two fragments which had length 11339 bp (H) and 501 bp (not visible due to low intensity).

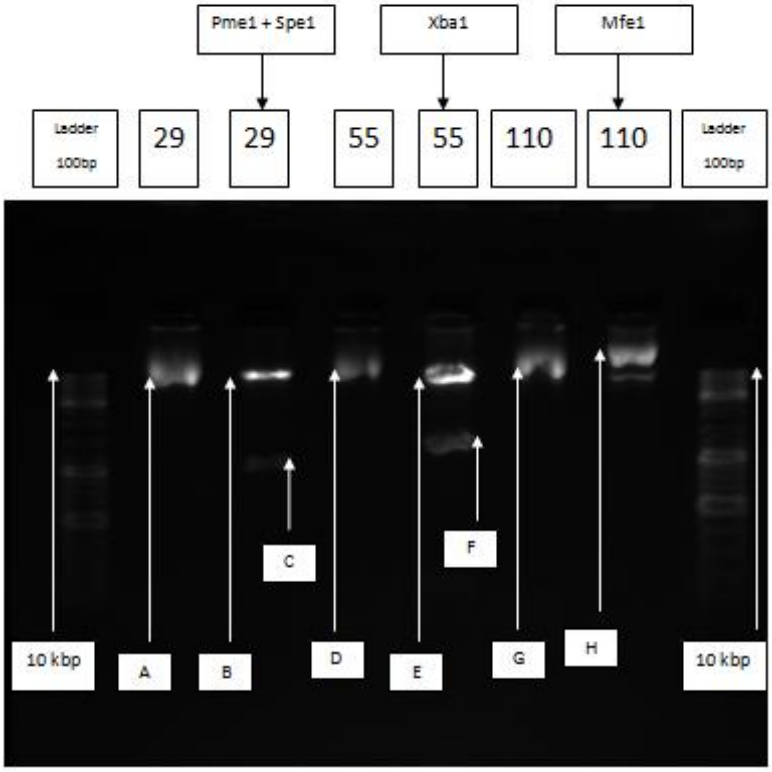


Figure 4.5: Restriction analysis on pGWB620 harboring different DRM2 constructs, visualized on a 1.5 % agarose gel. First and last lane is ladder while in the middle there are 29, 55, 110 deletion constructs of DRM2 loaded on pGWB620 with and without using restriction digest.

4.5 Characterization of Plant transformation

The deletion constructs loaded on pGWB620 were cloned in the DH5α and top 10 strains. They were transformed to ΔDRM1/DRM2 *Arabidopsis thaliana* through *Agrobacterium tumefaciens.* Knock out mutants ΔDRM1/DRM2 were established by insertion of t-DNA inside the gene sequence of both DRM1 and DRM2. This kind of insertion disrupted the function of the DRM2 gene.[9] The 55pGWB620 and 110pGWB620 were also transformed by leaf infiltration to *Nicotiana benthmania*.

The total protein of both plants was extracted by different buffers (Cold H buffer- Total protein extraction buffer1- Total protein extraction buffer 2). The most effective buffer was Total protein extraction buffer 2 due to the presence of NaCl, which maintains the ionic strength of the solution and solubility of the extracted protein. Afterwards, immunoblotting was performed to detect the expression of the truncated versions of atDRM2. By using a myc tagged protein (the source is E-coli) as a positive mark for

immunoblotting, there was no signal for the existence of deletion construct expression at the expected molecular weight which is 160 kDa. On the other hand, the signal for immunoblot at molecular weight 70 kDa for four deletion constructs was detectable.

4.6 DRM2 protein expression

More than one protocol was used to extract total protein from plants. Afterward, the most successful result obtained from SDS-PAGE, which is Total Protein Extraction buffer 2, was used. The first try was using Cold H buffer (Figure 4.6) where different protein samples from 8 days old *Arabidopsis thaliana* seedlings were extracted. This experiment was separated into two trials, one with loading suspension of protein samples before centrifugation and the second with loading supernatant after centrifugation of the protein sample to maximum speed.

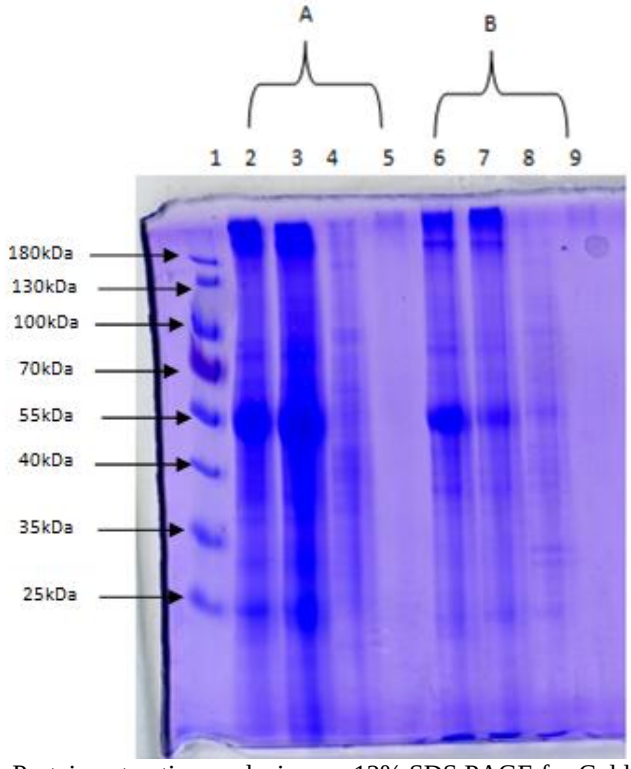


Figure 4.6: Protein extraction analysis on a 12% SDS PAGE for Cold H buffer extraction solution 1- Prestained protein ladder 10 to 180 kDa 2- At wild type 3- At with myc tag protein 4- At with pGWB602 insert 5- At with pGWB620 loaded with 55 deletion construct.6- At wild type 7- At with myc tag protein 8- At with pGWB602 insert 9- At with pGWB620 loaded with 55 deletion construct. (A) refers to usage of whole suspension before centrifugation. (B) refers to usage of supernatant of suspension after centrifugation.

The disadvantage of this buffer is that clear sharp bands for the protein product are missing.

The second try was with the Total Protein Extraction buffer 1 and Total Protein Extraction buffer 2 where the loaded samples were the three weeks old *Nicotiana Benthamnia* and 8 days old *Arabidopsis thaliana* seedlings.

.

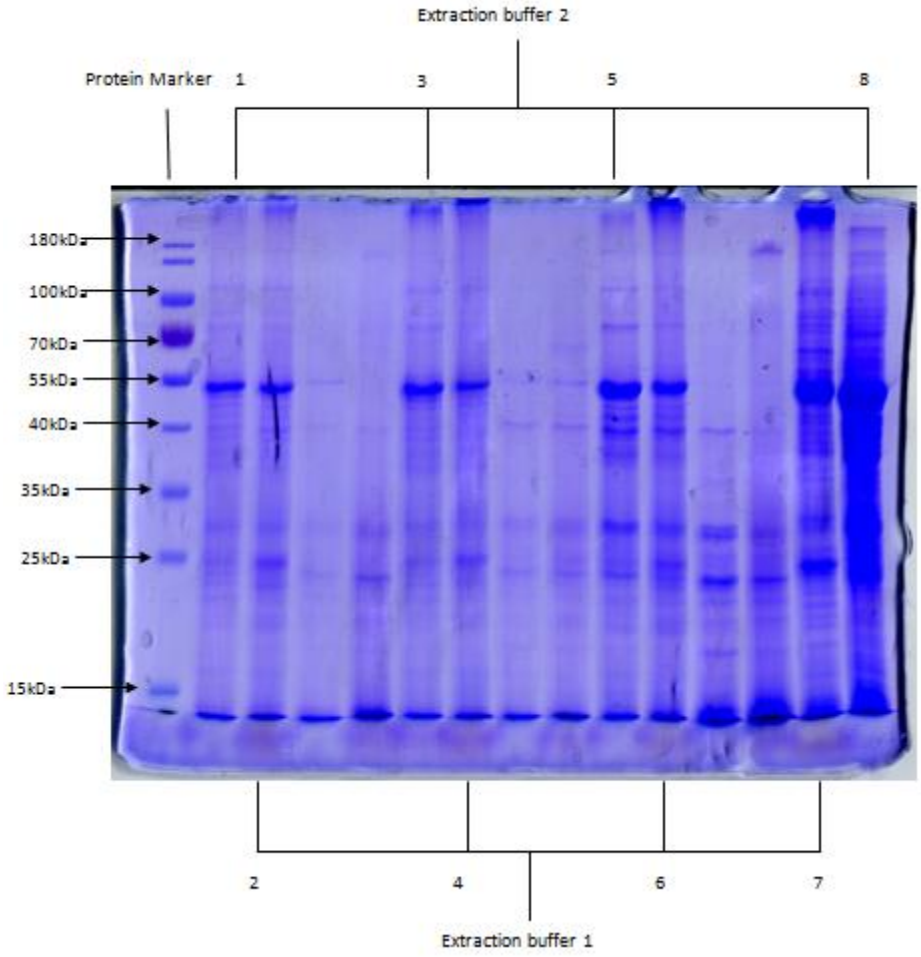


Figure 4.7 : 12% SDS-PAGE with protein marker 10 to 180 kDa in first lane. Extraction buffer 2 has been applied to following (1)deletion construct 55, (3) deletion construct 110 and (5)wild type which are *Nicotiana benthamnia* Leaves. (8) *Arabidopsis thaliana* Seedlings wild type. Extraction buffer 1 has been applied to following (2) deletion construct 55, (4) deletion construct 110 and (6) wild type which are *Nicotiana benthamnia* Leaves. (7) *Arabidopsis thaliana* Seedlings wild type

From the gel, we can see sharp bands and a large amount of the total protein could be extracted using the total protein extraction buffer 2 (Figure 4.7)

4.6.1 *Arabidopsis thaliana*

The inserted vector, which is pGWB 620, has a myc-tag where it is possible to detect a small amount of the protein of inserted sequences with the anti-myc antibody. Detecting the expression of the four deletion constructs was negative. To make sure that no mistakes were made using blotting, a positive control with the myc tag was used so it can be detected by Western blotting.

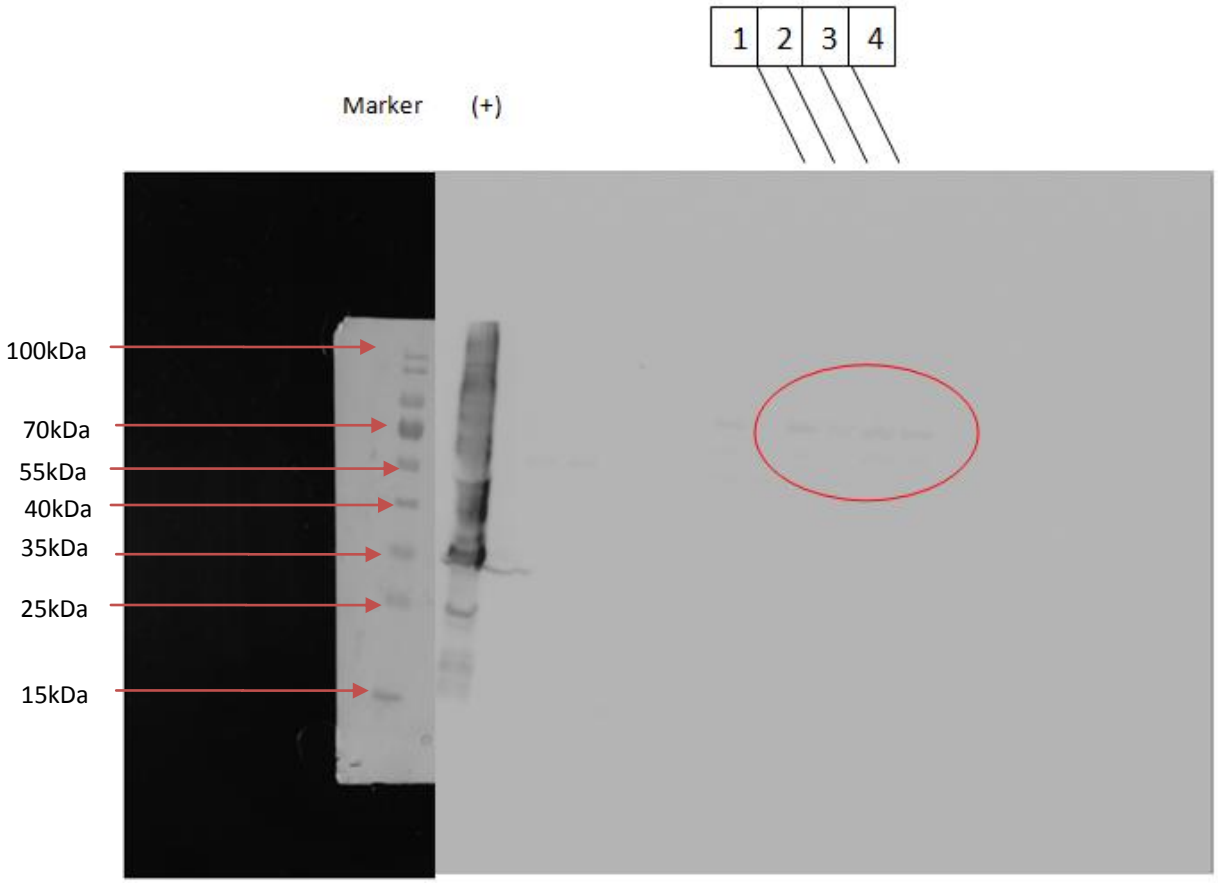


Figure 4.8: two images of ladder and ECL reaction has been put together. Western blot on the rescue AtDRM2 protein C-myc tagged. The first lane is protein marker and the second lane shows the positive control, where rest 1- deletion construct 29, 2- deletion construct 55, 3- deletion construct 110, 4-deletion construct 191. Lanes show no protein detection due to non expression of desired protein, or endogenous cleavage of the tag.

The protein samples of the four deletion constructs were blotted into the nitrocellulose membrane. The results of the first and second antibody were used to obtain fluorescence emission, which is detected by exposing the sample to the CCD imager. This exposure is done after subjecting the membrane to developing solution to make ECL reaction. The positive is nicely detected but test samples show emission of light at a different molecular weight ~70 kDa where it should weigh 160kDa. This suggests that the deletion construct proteins were subjected to proteolysis during the extraction process and kept its myc-tag attached with the DRM2 shortened at its N-terminus or its non-specific binding.

4.7 Explaining the negative result of Immunoblotting

Expression of deletion constructs could not be detected based on the results of the immunoblotting outcome. Steps like cloning and transformation were carried out to obtain the transformed plants with the truncated versions of atDRM2. Probably one of these steps did not work out correctly.

4.7.1 Expression with Luciferase assay

The virulence of the Agrobacterium tumefaciens that was used in these experiments has been questioned with respect to its efficiency. Agrobacterium has the capability of transforming its t-DNA into plant organisms. The pGWB vector series has borders with the t-DNA. To check the virulence and efficiency of the transformation process, a Luc assay was established. The pGWB 602 was loaded with the Luciferase enzyme (Figure 4D). If the transformation process is correct, then an emission of light should be detectable when we spray the transformed plants with the luciferin substrate. The experimental results indicated that emission of light occurred using the NIGHTSHADE LB 985 *in vivo* plant imaging system. It shows successful expression of the luciferase gene in most Arabidopsis seedlings (Figure 4.9).

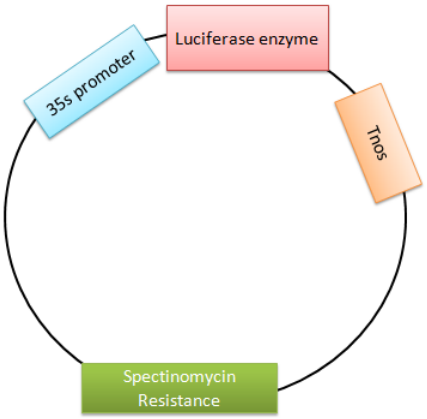


Figure 4D: pGWB 602 with Luciferase insert

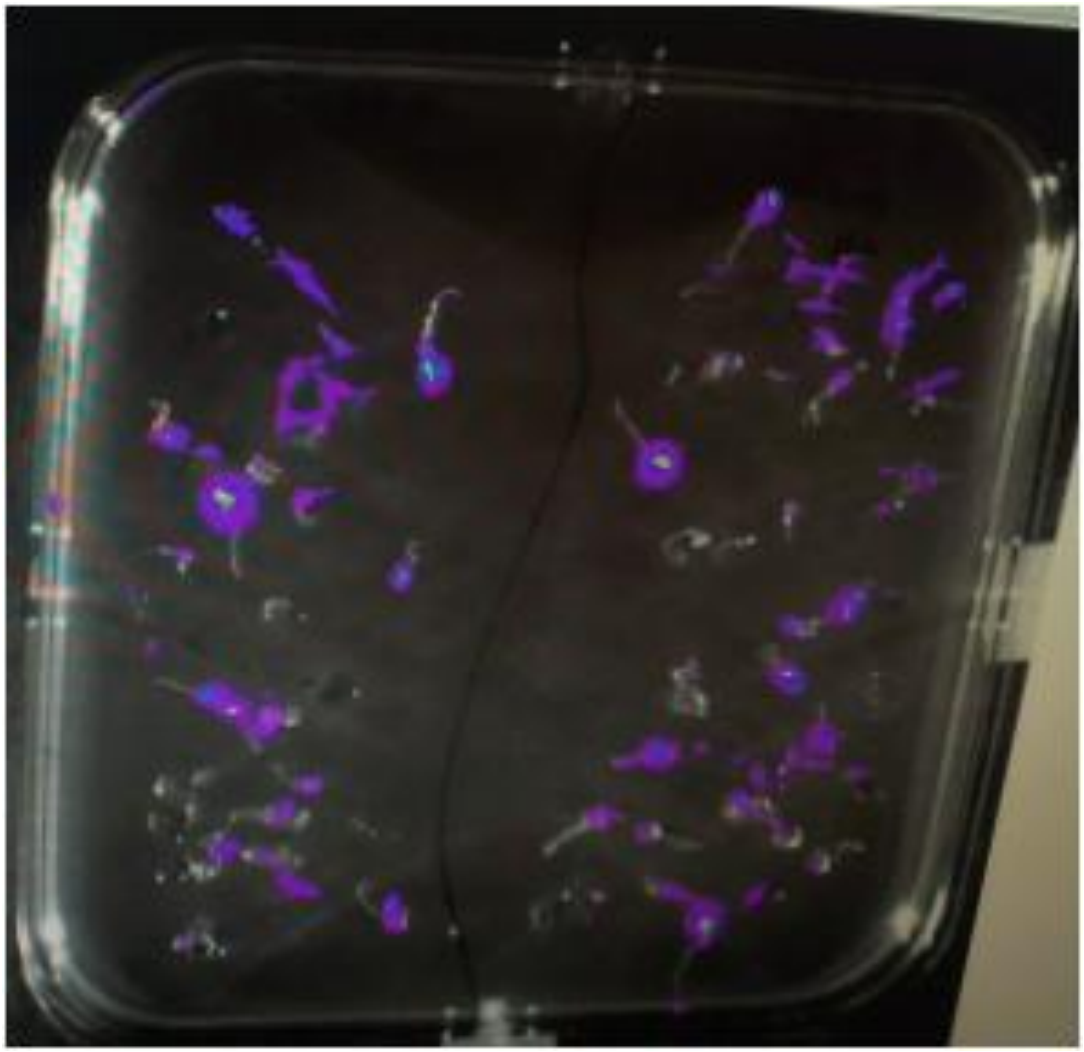


Figure 4.9: Luciferase assay, *Arabidopsis thaliana* seedlings with expressed Luciferase enzyme were sprayed with LUC substrate.

4.7.2 Amplification of 5’ untranslated region of DRM2

The luciferase assay demonstrated that Agrobacterium tumefaciens infection worked and its virulence appeared to be not the cause for the lack of expression of the deletion constructs. This suggested that the failure of the expressed deletion constructs may have been because of the absence of the 5′ UTR region where there is a ribosomal binding site to facilitate the translation process of the mRNA to the protein. In addition, the 3′ UTR should be considered because it contains the polyadenylation site that attenuates the digestion of mRNA. It was noted that pGWB620 contains sites for RBS and polyadenylation. Within the time constraints of the project, I was only able to include the 5′UTR to DRM2 genomic DNA by applying the 5′UTR primers. These PCR products can be obtained by any person coming after me to express DRM2 in other organisms.

The upstream region of DRM2 has been successfully amplified using PCR including the -501, -126 and -30 5′UTR nucleotides (Figure 4.16). Each construct has been prepared twice, one is a 5′ UTR with the full length of DRM2 (not including the stop codon) and the other is a 5′UTR with the full length of DRM2 and a stop codon. This stop codon is to facilitate the addition of tagged proteins at the N-terminus side. While those with the full length sequences can only provide an addition of C-terminus tagged proteins.

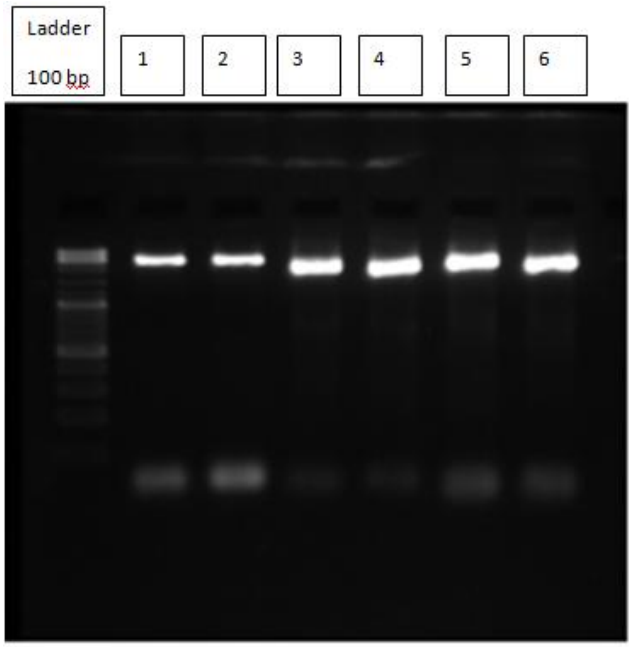


Figure 4.10: 1.5 % agarose gel of the PCR amplification, and purification, of the different DRM2 constructs. First lane is 100bp ladder, 1) -30 5' UTR + full length of DRM2. 2) -30 5' UTR + full length of DRM2 + stop codon. 3) -126 5' UTR + full length of DRM2 4)) -126 5' UTR + full length of DRM2 + stop codon. 5) -501 5' UTR + full length of DRM2. 6) -501 5' UTR + full length of DRM2 + stop codon.

4.8 Negative result from expressing DRM2

4.8.1 Gateway cloning

The possibility of inefficient detection of truncated versions of DRM2 in plants could be weakly related to the gateway cloning procedures. The system could introduce a nonsense mutation that would lead to an emerging stop codon in the middle of the sequence under study where neither full protein nor the myc-tag protein can be expressed. That explanation is not criticizing the gateway cloning system as the three organisms (E.coli, A. tumefaciens and A. thaliana) could introduce this mutation during the process of carrying deletion constructs. This might explain the weak signal as sequencing was done to confirm the precise insert of multiple deletion constructs.

4.8.2 Transformation

One clue on the absence of a DRM2 signal from the expression is the way that deletion constructs have been transformed into *Arabidopsis thaliana* through *Agrobacterium tumefaciens.* The virulence of *Agrobacterium tumefaciens* has been tested by the LUC

assay and it looked efficient. However, this may be a possible angle to address this issue. The product of the deletion constructs might affect the machinery of Ti-plasmid of A. tumefaciens in transferring t-DNA into A. thaliana. The inhibition reaction could be initiated during the incubating of *A. tumefaciens* on agar medium before transformation.

4.8.3 Transcription and translation

Chances could play a part and the transformed t-DNA could have been placed between the following regions:

1) the site where the sequence is methylated so that no mRNA transcription was possible.
2) the larger gene where the t-DNA is bordered with two splicing sites so that the mRNA product of the deletion constructs could be excised completely.

In addition to that, the deletion constructs may have been transcribed and translated properly but the amount of expressed protein was not enough to detect by the anti-myc antibody.

This suggests that in future experiments; RNA detection techniques should be added to help quantify the amount of mRNA produced by the deletion constructs. In addition, RNA detection techniques can be used to determine if either of the deletion constructs has been successfully transcribed or no expression at all has occurred.

Summary of Achievements

1- Deletion constructs loaded on expression vector pGWB620
2- Site directed mutagenesis of UBA 1 loaded on pDNOR201
3- Arabidospsis Thaliana seedling Transformants with deletion constructs
4- PCR products of upstream regions of DRM2 attached to genomic DRM2

References

1- Weber, M., & Schübeler, D. (2007). Genomic patterns of DNA methylation: targets and function of an epigenetic mark. Current Opinion in Cell Biology, 19(3), 273–280. http://doi.org/10.1016/j.ceb.2007.04.011

2- Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., Ecker, J. R. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature, 462(7271), 315–322. http://doi.org/10.1038/nature08514

3- Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., Jacobsen, S. E. (2008). Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature, 452(7184), 215–219. http://doi.org/10.1038/nature06745

4- Cao, X., & Jacobsen, S. E. (2002). Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. Proceedings of the National Academy of Sciences of the United States of America, 99 Suppl 4, 16491–16498. http://doi.org/10.1073/pnas.162371599

5- Huang, J., Wang, H., Xie, X., Zhang, D., Liu, Y., & Guo, G. (2010). Roles of DNA methyltransferases in Arabidopsis development. Journal of Biotechnology, 9(50), 8506–8514. http://doi.org/10.5897/AJB09.1941

6- Du, J., Zhong, X., Bernatavichute, Y. V, Stroud, H., Feng, S., Caro, E., Jacobsen, S. E. (2012). Dual Binding of Chromomethylase Domains to H3K9me2-Containing Nucleosomes Directs DNA Methylation in Plants. CELL, 151(1), 167–180. http://doi.org/10.1016/j.cell.2012.07.034

7- Du, J., Johnson, L. M., Jacobsen, S. E., & Patel, D. J. (2015). DNA methylation pathways and their crosstalk with histone methylation. Nature Reviews Molecular Cell Biology, 16(9), 519–532. http://doi.org/10.1038/nrm4043

8- Zhong, X., Du, J., Hale, C. J., Gallego-bartolome, J., Feng, S., Vashisht, A. A., Jacobsen, S. E. (2014). Molecular Mechanism of Action of Plant DRM De Novo DNA Methyltransferases. Cell, 157(5), 1050–1060. http://doi.org/10.1016/j.cell.2014.03.056

9- Cao, X., & Jacobsen, S. E. (2002). Role of the Arabidopsis DRM Methyltransferases in De Novo DNA Methylation and Gene Silencing, 12(02), 1138–1144

10- Cao, X., Springer, N. M., Muszynski, M. G., Phillips, R. L., Kaeppler, S., & Jacobsen, S. E. (2000). Conserved plant genes with similarity to mammalian de novo DNA methyltransferases, 97(9), 4979–4984.

11- Henderson, I. R., Deleris, A., Wong, W., Zhong, X., Chin, H. G., Gregory, A., Jacobsen, S. E. (2010). The De Novo Cytosine Methyltransferase DRM2 Requires Intact UBA Domains and a Catalytically Mutated Paralog DRM3 during RNA – Directed DNA Methylation in Arabidopsis thaliana, 6(10). http://doi.org/10.1371/journal.pgen.1001182

12- Calarco, J. P., & Martienssen, R. A. (2012). Imprinting : DNA Methyltransferases Illuminate Reprogramming. Current Biology, 22(21), R929–R931. http://doi.org/10.1016/j.cub.2012.09.030

13- Zhong, X., Hale, C. J., Nguyen, M., Ausin, I., Groth, M., Hetzel, J., Jacobsen, S. E. (2015).

DOMAINS REARRANGED METHYLTRANSFERASE3 controls DNA methylation and regulates RNA polymerase V transcript abundance in Arabidopsis. Proceedings of the National Academy of Sciences, 112(3), 911–916. http://doi.org/10.1073/pnas.1423603112

14- He, X.-J., Chen, T., & Zhu, J.-K. (2011). Regulation and function of DNA methylation in plants and animals. Cell Research, 21(3), 442–465. http://doi.org/10.1038/cr.2011.23

15- Huang, L., Jones, A. M. E., Searle, I., Patel, K., Vogler, H., Hubner, N. C., & Baulcombe, D. C. (2009). An atypical RNA polymerase involved in RNA silencing shares small subunits with RNA polymerase II. Nature Structural & Molecular Biology, 16(1), 91–93. http://doi.org/10.1038/nsmb.1539

16- Zhang, H., & Zhu, J. K. (2011). RNA-directed DNA methylation. Current Opinion in Plant Biology, 14(2), 142–147. http://doi.org/10.1016/j.pbi.2011.02.003a

17- Böhmdorfer, G. Rowley, M. J., Kuci, J., Zhu, Y., Amies, I., & Wierzbicki, A. T. (2014). RNA-directed DNA methylation requires stepwise binding of silencing factors to long non-coding RNA, 181–191. http://doi.org/10.1111/tpj.12563

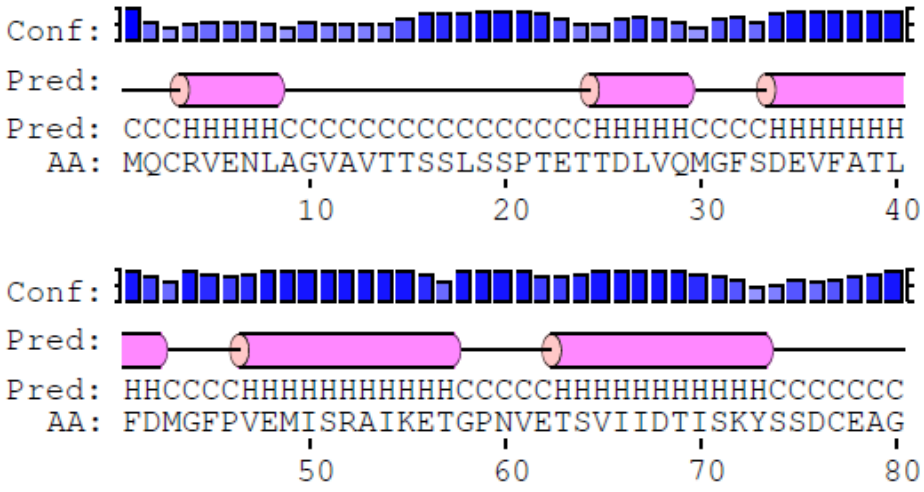
18- Dikic, I., Wakatsuki, S., & Walters, K. J. (2009). Ubiquitin-binding domains - from structures to functions. Nature Reviews. Molecular Cell Biology, 10(10), 659–671. http://doi.org/10.1038/nrm2767

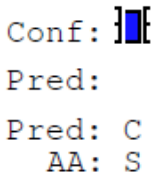
19- Hurley, J. H., Lee, S., & Prag, G. (2006). Ubiquitin-binding domains. The Biochemical Journal, 399(3), 361–372. http://doi.org/10.1042/BJ20061138

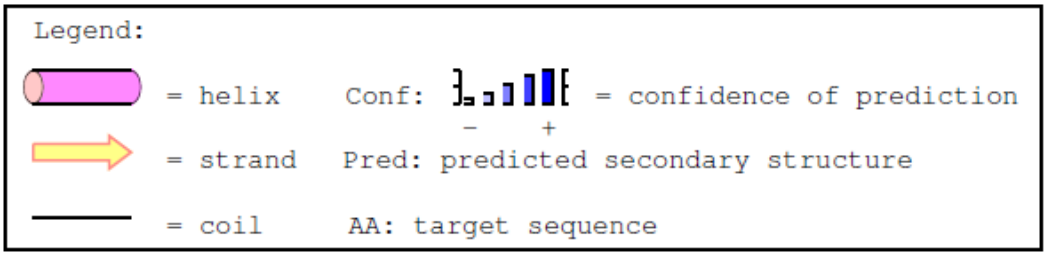
Appendix

atUBA1(Option1)

MQCRVENLAGVAVTTSSLSSPTETTDLVQMGFSDEVFATLFDMGFPVEMISRAIKETGPNVETSVIIDTISKYS
SDCEAGS

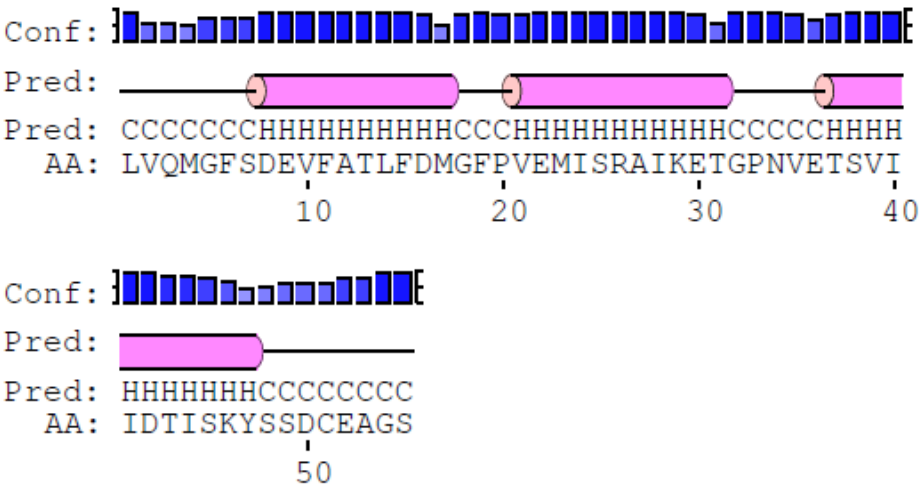


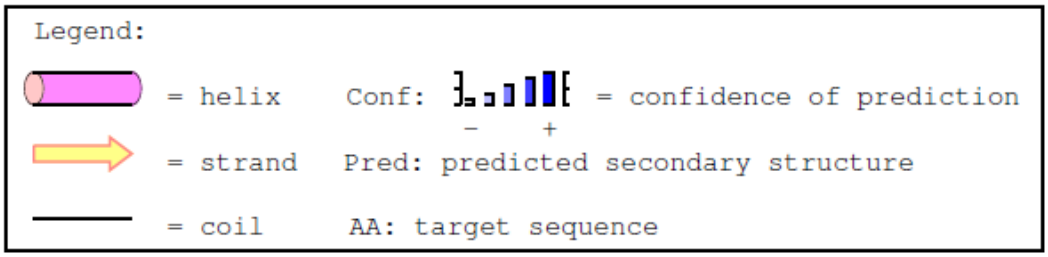




atUBA1(Option2)

LVQMGFSDEVFATLFDMGFPVEMISRAIKETGPNVETSVIIDTISKYSSDCEAGS





atUBA2

SKSKAIDHFLAMGFDEEKVVKAIQEHGEDNMEAIANALLSCPEAKKLPAAVEEEDGIDWSSSDDDTNYTDML
NSDDEKDPN

atUBA3

SNENGSKIRSLVKMGFSELEASLAVERCGENVDIAELTDFLCA