UNIVERSITY OF TARTU

Faculty of Science and Technology

Institute of Bioengineering

Oleksandr Syritka

Describing the non-canonical Restriction-Modification type II system (RMII) in *Pseudomonas putida* PaW85

Bachelor's Thesis (12 ECTS)

Curriculum Science & Technology

Supervisor(s):

Research Fellow, PhD Andres Ainelo

Associate Professor, PhD Hedvig Tamman

Tartu 2026

Describing the non-canonical Restriction-Modification type II system (RMII) in *Pseudomonas putida* PaW85

Abstract

Bacteriophages are the most abundant biological entities on Earth and exert a major selective pressure on bacterial populations. To counteract phage infection, bacteria have evolved a diverse arsenal of anti-phage defense mechanisms, among which Restriction-Modification (RM) systems are the most prevalent. RM systems protect the host by methylating self-DNA at specific recognition sequences while cleaving incoming foreign DNA that lacks the corresponding modification. This thesis focuses on a non-canonical type II RM system (RMII) encoded in *Pseudomonas putida* PaW85, a biotechnologically important soil bacterium. The RMII locus comprises a putative restrictase (*R, PP_3988*) and two distinct methylases (*M1, PP_3989* and *M2, PP_5651*). Using phage susceptibility assays with an expanded CEPEST phage collection, the system was confirmed to reduce infection efficiency of phages from genus clusters G16 and G17 by up to 200-fold. To determine the functional contribution of each methylase, deletion strains lacking either or both methylases were constructed and used for restrictase expression experiments. Both methylases were shown to be important, which refers to redundancy. Bioinformatic and structural analyses show that the M1 and M2 are unrelated, probably coupled together by horizontal gene transfer. M1 is a cytosine methyltransferase with structural similarity to methylases of known type IIA(S) systems, while M2 is evolutionarily unrelated to M1 and shares structural similarity with adenine methyltransferases. These findings characterize a non-canonical RMII system in *P. putida* PaW85 and provide a foundation for future biochemical studies of the restrictase and for informed genomic engineering of this industrially relevant bacterium for possible biotechnological applications.

Keywords

Bacteriophages, restriction-modification system, phage defense

CERCS: B230 Microbiology, Bacteriology, Virology, Mycology

Institute name: Institute of Molecular and Cell Biology

Research group: Bacteriophage group

Pseudomonas putida ebatüüpilise tüüp II restriktsiooni-modifikatsiooni süsteemi kirjeldamine

Lühikokkuvõte

Bakteriofaagid on kõige arvukamalt esinevad bioloogilised üksused Maal ja avaldavad bakteripopulatsioonidele suurt valikulist survet. Faagiinfektsioonide vastu võitlemiseks on bakteritel väljakujunenud mitmekesine faagivastaste kaitsemehhanismide arsenal, mille hulgas on kõige levinumad restriktsioon-modifitseerimise (RM) süsteemid. RM-süsteemid kaitsevad peremeesorganismi, metüleerides oma DNAd spetsiifilistest äratundmisjärjestustest. Samal ajal lagundavad nad sissetulevat võõr-DNAd, millel puudub vastav modifikatsioon. Käesolev bakalaurusetöö keskendub mittekanoonilisele II tüüpi RM-süsteemile (RMII) biotehnoloogiliselt olulises mullabakteris *Pseudomonas putida* PaW85. RMII lookus sisaldab oletatavat restriktaasi (*R, PP_3988*) ja kahte erinevat metülaasi (*M1, PP_3989* ja *M2, PP_5651*). Kasutades faagitundlikkuse teste laiendatud CEPEST faagikollektsiooniga, kinnitati, et süsteem vähendab G16 ja G17 faagiperekonda kuuluvate faagide infektsiooni efektiivsust kuni 200 korda. Metülaaside funktsionaalse panuse määramiseks konstrueeriti deletsioonitüved, millel puudus restriktaas ning üks või mõlemad metülaasid, ja neid kasutati restriktaasi ekspressioonikatsetes. Tulemused näitasid, et restriktaas muutub toksiliseks ainult siis, kui mõlemad metülaasid puuduvad, kinnitades, et M1 ja M2 kaitsevad kumbki iseseisvalt peremeesorganismi genoomi restriktaasi eest. Bioinformaatilised järjestuse- ja struktuurianalüüsid näitasid, et metülaasid ei ole omavahel sarnased ja on ilmselt kokku sattunud horisontaalse geeniülekandega. M1 on tsütosiinmetüültransferaas, millel on struktuuriline sarnasus teadaolevate IIA(S) tüüpi süsteemide metülaasidega, samas kui M2 jagab struktuurilist sarnasust adeniin-metüültransferaasidega. Saadud tulemused iseloomustavad *P. putida* ebatüüpilist RMII süsteemi ja loovad aluse restriktaasi edasisteks biokeemilisteks uuringuteks ja selle tööstuslikult olulise bakteri teadlikuks genoomseks muutmiseks biotehnoloogia hüvanguks.

Võtmesõnad:

Bakteriofaagid, restriktsiooni-modifikatsiooni süsteemid, faagikaitse süsteemid

CERCS: B230 Mikrobioloogia, Bacterioloogia, Viroloogia, Mükoloogia

TABLE OF CONTENTS

TERMS, ABBREVIATIONS AND NOTATIONS

ATC – Anhydrotetracycline

ATP – Adenosine triphosphate

bp – Base pair(s)

CEPEST – Collection of Environmental Phages for *Pseudomonas putida* Strains in Tartu

CFU – Colony-forming unit(s)

Cip – Ciprofloxacin

CRISPR – Clustered Regularly Interspaced Short Palindromic Repeats

crRNA – CRISPR RNA

DNA – Deoxyribonucleic acid

Gm – Gentamycin

HGT – Horizontal gene transfer

Km – Kanamycin

LB – Lysogeny broth

M – Methylation subunit (in type I and III RM complexes)

M1 – Methylase 1 (PP_3989 in the RMII locus of *P. putida* PaW85)

M2 – Methylase 2 (PP_5651 in the RMII locus of *P. putida* PaW85)

mRNA – Messenger RNA

MTase – Methyltransferase

NAD+ – Nicotinamide adenine dinucleotide (oxidised form)

NAD(P)H – Nicotinamide adenine dinucleotide (phosphate), reduced form

NTP – Nucleoside triphosphate

OD580 – Optical density measured at 580 nm wavelength

PCR – Polymerase chain reaction

PDB – Protein Data Bank

PPCQ – Proline-Proline-Cysteine-Glutamine (conserved catalytic motif in cytosine methyltransferases)

R – Restriction subunit (in type I and III RM complexes); also used as a label for the restrictase gene (PP_3988) in the RMII locus

RecBCD – Exonuclease V complex (helicase-nuclease involved in DNA repair and recombination)

REase – Restriction endonuclease

RM – Restriction-Modification

RMI – Restriction-Modification type I system (as used in strain designations in this work)

RMII – Restriction-Modification type II system (the non-canonical system studied in this work)

RNA – Ribonucleic acid

S – Specificity subunit (in type I RM complexes)

SAM – S-adenosyl-L-methionine

Sm – Streptomycin

SM buffer – Sodium-Magnesium buffer

SUMO – Small Ubiquitin-like Modifier (affinity tag used for protein purification)

TM-score – Template Modelling score (measure of structural similarity between protein structures)

VBR – Virus-to-Bacteria Ratio

WT – Wild type

INTRODUCTION

Bacteriophages (phages) are the most abundant biological entities on Earth outnumbering their bacterial hosts by at least an order of magnitude in most environments (Güemes et al., 2016). They are obligate intracellular parasites that rely entirely on their bacterial hosts for replication (Sharma et al., 2017). Phages strongly shape bacterial populations primarily by exerting strong selective pressure through infection and killing. In response to this predatory pressure, bacteria have evolved a wide range of defense mechanisms (Chevallereau et al., 2022).

These defense systems include alterations in surface receptors that block phage adsorption, nuclease-based mechanisms that degrade injected viral DNA, and abortive infection systems in which an infected cell sacrifices itself before the phages mature. Among the nuclease-based systems, Restriction-Modification (RM) systems are the most abundant and are encoded in the majority of bacterial genomes (Georjon & Bernheim, 2023). These systems consist of a methyltransferase that modifies specific short DNA sequences in the host genome to prevent degradation by a restriction endonuclease, which recognizes the same sequence and cleaves any DNA that lacks the protective modification. RM systems are classified into four major types (I–IV) based on their subunit composition, recognition sequence properties, cleavage mechanism, and cofactor requirements (Kommireddy & Nagaraja, 2013). Type II RM systems are the most well-studied due to their predictable cleavage properties, which have made them irreplaceable tools in recombinant DNA technology. However, many non-canonical type II systems have also been identified and further subclassified. One of the subtypes includes systems that encode two distinct methylase genes to provide full modification of an asymmetric recognition sequence (Loenen et al., 2014).

Pseudomonas putida has been successfully employed in industrial biotechnology for its metabolic versatility and tolerance to oxidative stress (Nikel & de Lorenzo, 2018). However, phage contamination during industrial fermentations may lead to lower product yields or complete batch failures. Therefore, characterizing the specific arsenal of defense systems in biotechnologically relevant bacteria is needed. To enable knowledgeable construct of strains that would be less prone to infection and more consistent in production (Xu et al., 2025).

With the development of CEPEST, the largest collection of environmental phages for *P. putida* to date, many predicted antiphage defense systems can now be validated (Brauer et al., 2024). This work aims to describe a non-canonical type II restriction-modification system (RMII) in *P. putida* PaW85 harbors two distinct methylases and a putative restrictase, raising questions about the functional roles of each component.

1 LITERATURE REVIEW

1.1 Bacteria and phage interactions

Bacteria are extremely diverse and occupy a wide range of marine and soil habitats (Lozupone & Knight, 2007). There, they are accompanied by bacteriophages, viruses that infect bacterial hosts and shape bacterial populations by using two main replicative strategies: lytic and lysogenic cycles (Sharma et al., 2017).

At the beginning of the lytic cycle, the phage attaches to the bacteria through the complementarity of the host's surface receptors and the tail proteins of the virus. Then, the bacteriophage's genetic material is inserted into the host cell. Afterward, the phage's genetic material is transcribed into viral mRNA. This mRNA hijacks the host's ribosomes to translate viral proteins needed for complete host takeover and the assembly of new phage particles. The phage genome is replicated, and later its copies are packaged into viral capsids. The bacterial cell is eventually lysed, and the viral progeny "bursts out" from the cell and proceeds to infect other cells (Sharma et al., 2017). Sometimes parts of the host's chromosome get mistakenly incorporated into capsids; therefore, when a new host is infected, the captured fragment may recombine and become part of the bacterial genome. This is one of the ways transduction can occur, which is part of the group of mechanisms of horizontal gene transfer (HGT) within bacterial populations. Thus, phages serve not only as selective agents but also facilitate the evolution of their hosts (Borodovich et al., 2022).

The lysogenic cycle starts to differ from the lytic cycle when the phage's injected genetic material is incorporated into the host's genome. Lytic replication is suppressed, and the inserted phage genome propagates with the bacterial chromosome as a prophage (Sharma et al., 2017). The latter can carry genes that provide a fitness advantage or drastically affect the bacteria's phenotype. For example, when *Vibrio cholerae* is infected with the lysogenic CTX phage, it becomes toxigenic because the prophage encodes the cholera toxin (Waldor & Mekalanos, 1996).

The possibility of establishing such a relationship with the host and, eventually, of deciding between killing and surviving within, combined with the extreme abundance of phages, makes them one of the main forces shaping bacterial populations in colonized environments (Chevallereau et al., 2022). Bacteriophages are a major source of mortality, killing approximately 20% of the ocean's microbial biomass daily. In doing so, phages not only exert selective pressure on bacterial populations but also influence the niches they inhabit. In the end of the infection cycle, the nutrients stored in cells are commonly released back into the

environment; therefore, more carbon and energy are kept in the microbial loop and are not passed further in the food chain to larger organisms. Conversely, this nutrient pulse could also benefit survivors within the population. Furthermore, viruses prevent the most competitive and rapidly growing bacterial taxa from taking over, allowing rarer and more diverse species to persist (Suttle, 2007).

The nature of the habitat promotes diversity as well. For instance, soil environments, compared to marine ones, host greater bacterial and viral diversity since migration is restricted and populations are better isolated, resulting in a higher number of locally unique communities. This also leads to a larger variation in the Virus-to-Bacteria Ratio (VBR) across diverse soil types. In soils with low organic matter and poor water content, the VBR can reach values over 1000, as bacterial populations decline while viral populations persist. In such conditions, when the host is scarce, lysogeny is the preferred replication strategy (Srinivasiah et al., 2008). However, if the conditions improve, phages can switch to the lytic cycle and exploit actively growing bacteria to rapidly increase their numbers (Nabergoj et al., 2017). To counteract viral activity, bacteria have evolved a diverse arsenal of defense strategies to prevent or halt infection (Georjon & Bernheim, 2023).

1.2 General bacterial defense strategies

The constant arms race between bacteria and phages has led to the evolution of highly diverse anti-phage defense mechanisms (Georjon & Bernheim, 2023). A significant driver of this diversity is the fitness cost of maintaining these systems. It is impractical for individual cells to carry every defense system, since it would be energy expensive and increase the risk of autoimmunity. Therefore, bacteria split the burden by sharing defense-encoding genes within the population through horizontal gene transfer (Bernheim & Sorek, 2020). These systems often cluster within the genome and form defense islands, genomic regions enriched in defense genes and DNA mobility-related genes such as recombinases, transposases, and integrases. Thus, defense islands serve as hotspots for the integration of incoming genetic elements. This further facilitates the spread of defense systems across bacterial communities by allowing rapid gain and loss of these systems in response to changing selective pressures. The discovery of defense islands has benefited the research as well. Many novel bacterial defense systems have been identified using a guilt-by-association principle, which involves the analysis of genes consistently associated with defense islands (Makarova et al., 2011).

To provide a comprehensive overview of defense systems, their classification is based on the mechanisms by which they hinder phage infection. The strategies mirror the key stages of the

lytic cycle: preventing phage DNA entry, interfering with an injected viral genome, and sacrificing the infected cell to limit phage production and protect the bacterial population (Georjon & Bernheim, 2023).

1.2.1 Preventing the phage genome entry

For the phage genome to enter the cell, the virus must attach to the cell surface and inject its genetic material. The most straightforward strategy for bacteria to avoid phage adsorption is to alter or block the surface receptors required for the process (Chevallereau et al., 2022). This can be achieved through mutations in genes encoding the receptors or through biochemical modifications of the receptors. For example, some *Pseudomonas aeruginosa* strains can protect themselves against pili-specific bacteriophages by masking their binding sites through glycosylation of type IV pili proteins (Harvey et al., 2018). However, biochemical modifications and especially alterations in receptor gene sequences may result in disadvantageous phenotypes. Therefore, bacterial populations also protect themselves through phase variation, which allows cells to alternate between different surface structures by dynamically regulating gene expression, usually via methylation (Chevallereau et al., 2022). For example, phase variation in the *lic2A* gene in *Haemophilus influenzae* changes lipooligosaccharide structure on the surface, which is required for phage adsorption. This creates phage-resistant subpopulations of the single clone, ensuring the population's survival (Zaleski et al., 2005). Certain bacteria utilize sophisticated strategies to prevent phage attachment. For example, some bacteria release vesicles that act as decoys because they contain the same receptors that are found on the cell membrane. As a result, phages bind to the vesicles instead of the bacterium and waste their infection potential (Manning & Kuehn, 2011).

Even if the bacteriophage has successfully attached to the surface, the infection can still be prevented by blocking genome injection. Some protein structures within the bacterial envelope have the capacity to do this. For example, the prophage-encoded SieA protein blocks the injection of DNA from the P22-like phages into *Salmonella enterica* by interfering with the formation or function of the channel required for genome delivery (Leavitt et al., 2024).

1.2.2 Disrupting an injected viral genome

After the injection, the next layer of defense involves impairing the phage genome, which is usually achieved by degrading the viral nucleic acid. The most common nuclease-based defense systems are Restriction-Modification (RM) systems, which distinguish between host and foreign DNA using the methylation status (Georjon & Bernheim, 2023). The CRISPR-Cas system, in turn, behaves as an adaptive immune system. During a phage infection, a specialized

complex of Cas proteins identifies and integrates short fragments of viral DNA into the CRISPR array on the bacterial chromosome. This array consists of short stretches of direct repeats that are interrupted by previously acquired unique viral sequences termed spacers. These spacers are then transcribed into the CRISPR RNAs (crRNAs), which guide the nucleases to the matching sequences of the invading viral DNA. In this manner, the CRISPR-Cas systems store information from past infections and use it to halt future phage infections (Hille et al., 2018).

1.2.3 Abortive infection systems

Even if the phage successfully establishes an infection, its replication can still be interrupted if the host induces its own death or metabolic arrest. The size of viral progeny is minimized, and the bacterial population is protected. The systems responsible for this last layer of defense are termed abortive infection systems, which consist of the sensing and effector modules. Phage infection is detected either by recognizing viral components, such as tail tube proteins, capsids, and nucleic acids, or by monitoring the host's essential systems for malfunctions caused by viral activity. Once the sensing components are activated, cell death is commonly achieved through the following mechanisms: depletion of essential nutrients, such as ATP, NAD+, and nucleotides; non-specific degradation of both host and viral genomes; and disruption of the cellular membrane. (Georjon & Bernheim, 2023)

1.2.4 Prophage-encoded defense systems

Prophages survive by persisting and replicating inside the host; therefore, they would benefit from keeping it alive and protecting it against competing phages. Consequently, many defense systems are carried by prophages. These systems can belong to any class of defense strategies described before (Chevallereau et al., 2022; Georjon & Bernheim, 2023). Furthermore, prophages can provide resistance through a specific mechanism known as superinfection immunity. This defense mechanism is an indirect consequence of the prophage maintaining its lysogeny. To persist in the cell, the prophage inhibits its own lytic cycle by continuously producing the repressor, which binds the operator sequences of the lytic genes. As a result, when another phage infects the bacterium, its lytic genes are immediately silenced by the repressor already present in the cytoplasm. This mechanism neutralizes only very closely related phages because repressors must be highly compatible with the operator sequences of the invading phages (Chevallereau et al., 2022).

1.3 Restriction-Modification systems

Restriction-Modification (RM) systems defend the host against phages by cleaving incoming viral DNA. They consist of the sequence-specific restriction endonuclease (REase), which,

upon recognition, cleaves DNA by catalyzing the hydrolysis of phosphodiester bonds, and the methyltransferase (MTase) that prevents degradation of the host's own DNA by methylating adenine or cytosine bases in the same recognition sequence (Figure 3) (Kommireddy & Nagaraja, 2013). Cleavage of phosphodiester bonds by REases can be viewed as a transfer of the phosphoryl group to water. For that, they typically require a divalent metal ion, such as Mg^{2+}, as a cofactor to stabilize the negative charge of the leaving group and activate a water molecule for hydrolysis (Pingoud & Jeltsch, 2001). As for MTases, their activity is largely dependent on S-adenosyl-L-methionine (SAM), which is a universal methyl group donor (Figure 1). The sulfur atom of the methionine is covalently linked to an adenosine moiety. The resulting positive charge on a sulfonium cation strongly attracts electrons from the attached methyl group. Consequently, the electrophilicity of the methyl group increases, making it a good leaving group (Jędrzejewski et al., 2025).

RM systems are classified into four types primarily based on the following properties of their restriction components: structure, energy requirements, and cleavage mechanism (Kommireddy & Nagaraja, 2013).

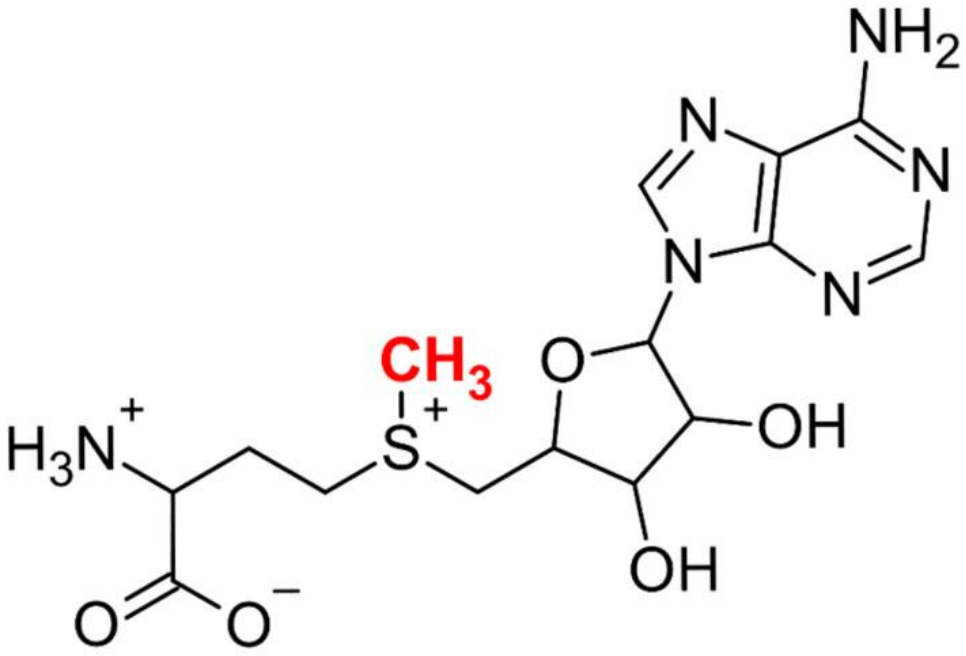

Figure *1*. Structure of S-adenosylmethionine. Adapted from (Jędrzejewski et al., 2025). The molecule consists of a methionine residue covalently linked to an adenosyl moiety, with the reactive methyl group highlighted in red.

1.3.1 Type I RM systems

The key feature of type I RM systems is that they are multisubunit proteins that function as a single protein complex and typically contain two R (restriction) subunits, two M (methylation) and one S (specificity) subunit for sequence recognition. When these enzymes encounter

unmethylated DNA, they behave mainly as restrictases and cleave unpredictably at variable sites away from the recognition sequence. The sequence they recognize is usually asymmetric and bipartite, meaning it consists of two specific motifs separated by a spacer, a gap of non-specific nucleotides (Loenen et al., 2014; Roberts et al., 2003). After binding the recognition site, they use ATP to drive DNA translocation, creating DNA loops in the process (Figure 2). Once the translocation is stopped, cleavage is triggered with a metal ion as a cofactor. Since translocation is usually blocked at random by DNA-bound proteins or atypical DNA structures, cleavage can occur at different sites, even if the same site is used for recognition (Murray, 2000; Roberts et al., 2003). On the other hand, these complexes function as methylases when they act on hemimethylated DNA. M subunits use SAM for the transfer of the methyl group (Roberts et al., 2003).

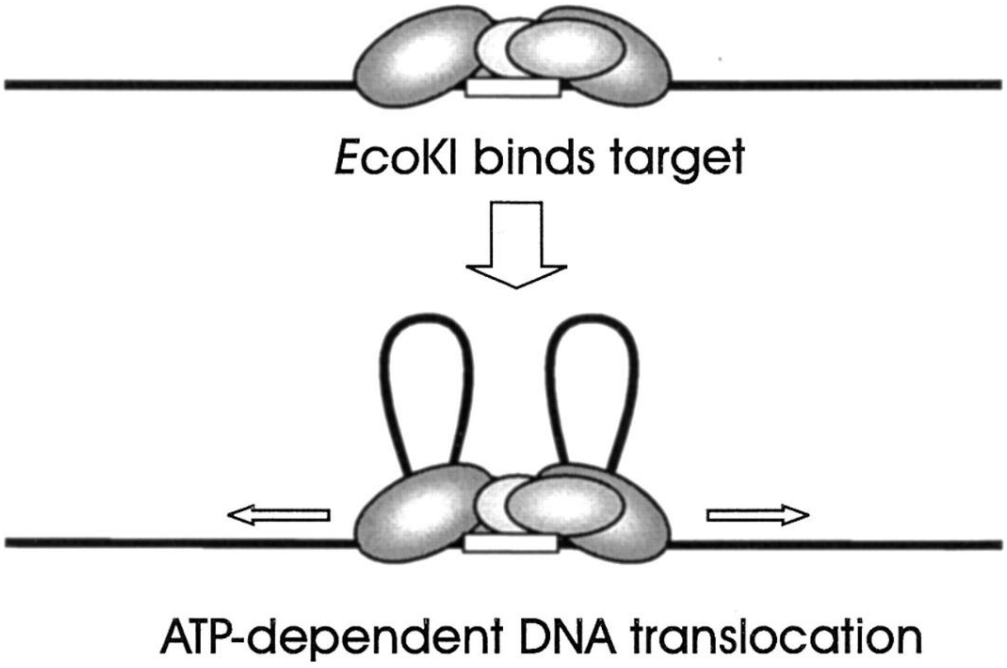

Figure *2*. Schematic representation of DNA translocation by EcoKI (type I RM complex). Adapted from (Murray, 2000). The figure depicts the EcoKI enzyme complex translocating DNA (bottom part of the figure) upon binding the target site (top part of the figure). DNA loops form as the complex continuously "pulls" DNA strands from both directions while remaining bound to the recognition sequence (bottom part of the figure).

1.3.2 Type II RM systems

Unlike type I RM systems, type II systems consist of two separate enzymes, REase and MTase, which are often encoded in the same operon but act independently (Figure 3). REases recognize short (4-8 bp), palindromic sequences and cut the DNA consistently at the same positions within or immediately adjacent to the recognition site. They usually dimerize and use divalent metal ions as cofactors for cleavage. MTases, in turn, typically function as monomers and methylate

the same recognition sequence, using SAM as the methyl group donor. Methylation of both strands of the recognition site is required for efficient protection from restriction. During host genome replication, the newly synthesized strand is unmethylated, while the parental strand remains methylated. This hemimethylation is enough to prevent cleavage. As a result, MTase has time to find the unmodified strand and produce fully methylated DNA. (Pingoud & Jeltsch, 2001)

The ability of type II REases to produce predictable DNA ends after cleavage has made them irreplaceable tools in recombinant DNA technology. Therefore, these REases have been extensively studied, and numerous subtypes that don't follow the canonical description have been identified. The enzymes with the canonical organization belong to the type IIP subtype, where "P" stands for "palindrome". (Loenen et al., 2014)

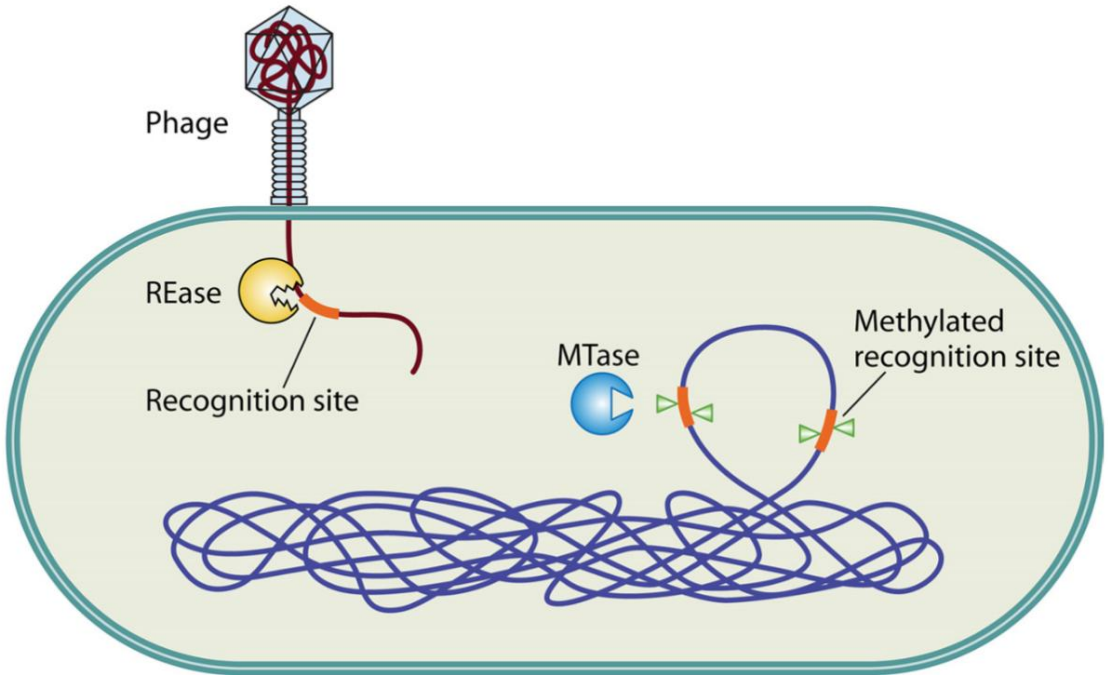

Figure *3*. Basic description of RM type II systems' function in phage defense. Adapted from (Kommireddy & Nagaraja, 2013). The figure presents the basic logic behind RM type II systems as an antiphage defense mechanism. Two separate enzymes, where the methylase (in blue) is responsible for the methylation of the bacterium's genome, which prevents its degradation by the restrictase (in yellow), which cleaves incoming unmodified DNA at a specific recognition sequence (in orange).

1.3.3 Type III RM systems

These systems are typically heterotrimeric complexes consisting of two M subunits and one R subunit. The M subunits are responsible for sequence recognition and transfer of the methyl group from SAM, while the R subunit is responsible for the restriction activity. Compared with other RM system types, type III enzymes employ an intricate ATP-dependent mechanism for

cleavage. They require two copies of an unmethylated asymmetric recognition sequence within the same DNA molecule. Separate M_2R_1 complexes bind each of these sequences with their M subunits (Loenen et al., 2014; Pingoud & Jeltsch, 2001). Then, ATP hydrolysis is used to drive complex rearrangements that result in the dissociation of M subunits and the specific binding of the R subunit to DNA. Once bound, the R subunit is active and able to slide along the DNA molecule in an ATP-independent manner for a short period of time. The cleavage of both strands is triggered when two sliding R subunits collide. Therefore, the recognition sites must also be arranged in an inverse orientation (Figure 4) (Göse et al., 2024). As for the methylation, type III complexes modify only one strand of the recognition sequence, because the sequence is asymmetric. However, such hemimethylation is sufficient to prevent self-digestion of the host chromosome, since the recognition sites are methylated on opposite strands because of their inverse orientation in the genome. When the daughter strand is synthesized during DNA replication, one of the sites remains methylated, and cleavage does not take place (Kommireddy & Nagaraja, 2013; Roberts et al., 2003).

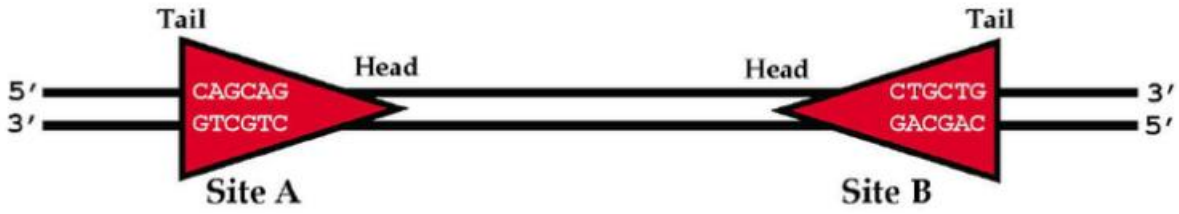

Figure *4*. Inversely oriented recognition sequences of EcoP15I. Adapted from (Nidhanapati K et al., 2012). The recognition site of the Type III RM enzymes is represented as a triangle. The base of the triangle (5`) is the tail, and the tip (3`) of the triangle is the head of the recognition sequence. Recognition sites on a DNA molecule are in a head-to-head orientation, which is required for cleavage by Type III RM enzymes (EcoP15I in this case).

1.3.4 Type IV RM systems

One of the mechanisms that phages utilize to bypass bacterial RM systems is the modification of their own DNA. Type IV systems have evolved to counter this phage adaptation (Kommireddy & Nagaraja, 2013). They are not technically RM systems because they lack the modification component and encode only the restrictase. However, these REases require specific modified DNA bases for cleavage. They generally have low sequence specificity because the host usually lacks the recognized modifications. Some enzymes have even evolved to target phage-specific modifications such as 5-hydroxymethylcytosine. As a result, they can target a wide variety of invading genetic elements. Finally, these enzymes have diverse cofactor requirements. Some use specific metal ions, while others also need a certain NTP for restriction activity (Loenen & Raleigh, 2014).

1.3.5 Other functions of RM systems

The prevalence of RM systems in bacteria underscores their effectiveness as a defense mechanism. However, the defense function alone does not explain their uneven distribution among bacteria or their high recognition sequence specificity. These systems can interact with any foreign DNA that contains an unmethylated recognition site or, in the case of type IV enzymes, a specific modification (Kommireddy & Nagaraja, 2013). Therefore, they also help maintain genome stability by restricting horizontally acquired DNA that might otherwise disrupt the host's genome through recombination (Messling & Williams, 2016). This is further supported by the fact that naturally competent bacteria possess an unusually high number of RM systems (Kommireddy & Nagaraja, 2013). For example, there are typically 13-16 distinct RM systems in the genomes of *Neisseria gonorrhoeae* isolates. Many of these systems are phase-variable and can be randomly switched between "on" and "off" states. This allows the population to regulate the restriction of incoming DNA, and thereby retain beneficial genetic material and exclude the harmful one (Elsener et al., 2026). Furthermore, some phase-variable MTases have been co-opted for epigenetic regulation of certain bacterial genes, leading to phenotypic heterogeneity within clonal populations and greater adaptability to environmental changes. In addition, RM systems prevent the exchange of genetic material between different strains by restricting foreign DNA that lacks native methylation patterns. This isolation promotes the divergent evolution, which can lead to speciation. However, if the incoming DNA is from a closely related bacterium and contains Chi sites recognized by the host's RecBCD end-processing complex, cleavage of this DNA may lead to its integration into the host chromosome via recombination (Figure 5). Therefore, RM systems might not only limit the effects of foreign genetic material on the bacterial genome but also enhance them (Kommireddy & Nagaraja, 2013). Regarding sequence specificity, if it were lowered, the target range of restriction enzymes would broaden. This would be beneficial for bacteria, as it would provide protection against a wider variety of phages and enable tighter regulation of horizontal gene transfer outcomes (Vasu et al., 2012). In practice, however, RM systems exhibit high sequence specificity. Therefore, these systems have been described as selfish genetic elements that compete for specific sequences. Meaning they are genomic parasites that ensure their own persistence without necessarily benefiting the host. This behavior is best understood in type II systems, where REase and MTase are separate enzymes. If RM genes are lost, the remaining MTase is diluted after each cell division cycle. Eventually, a sufficient number of recognition sites won't be modified, and the chromosome will be vulnerable to cleavage by the remaining REase. This forces cells to retain RM genes even if they don't confer a fitness advantage. It has

been observed that two RM systems with the same sequence specificity cannot be maintained in the genome simultaneously. Since the MTase of one system protects the chromosome from the REase of the other, there is no pressure to maintain each system. This competitive exclusion drives the RM systems to specialize in unique sequences (Kusano et al., 1995).

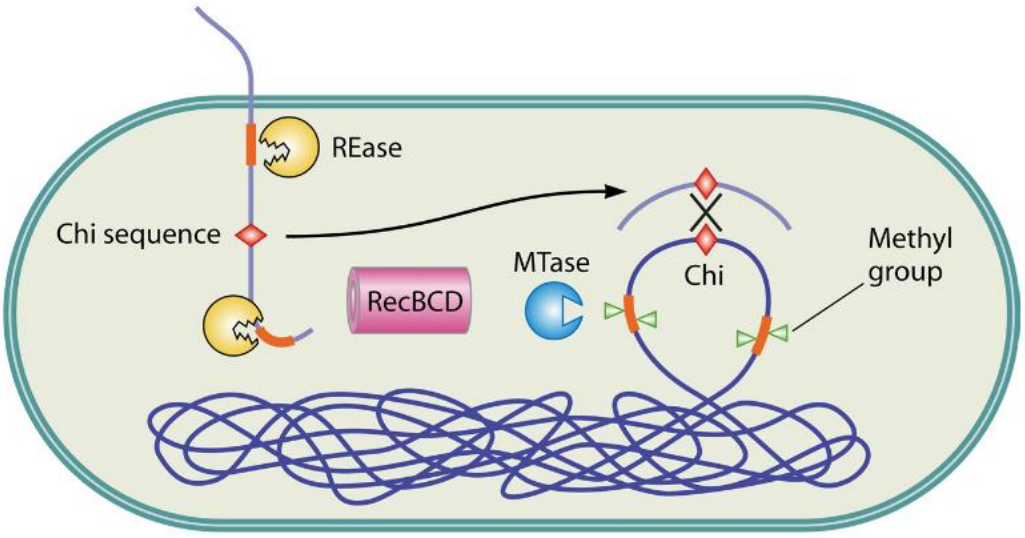

Figure *5*. Facilitation of the integration of incoming genetic elements by RM systems. Adapted from (Kommireddy & Nagaraja, 2013). Restriction of incoming DNA from a closely related bacterium (containing similar Chi sequences) generates DNA fragments that can serve as substrates for homologous recombination by the RecBCD pathway.

1.4 Unusual architectures of Restriction-Modification type II systems

Many type II enzymes deviate from the standard definition, and non-canonical groups have been defined (Loenen et al., 2014). These subtypes are briefly described in Table 1.

Table 1. All RM II subtypes defined so far and their corresponding descriptions.

Subtype	Description	Sources
Type IIP	"P" – palindromic. Canonical RMII systems. Recognize palindromic sequences.	(Pingoud & Jeltsch, 2001)
Type IIA	"A" – asymmetric. Asymmetric recognition site. Thus, they usually require a separate methylase for each DNA strand for complete methylation.	(Roberts et al., 2003)
Type IIS	"S" – shifted cleavage. Cleavage occurs at a defined distance from the recognition site.	(Pingoud & Jeltsch, 2001)
Type IIF	"F" – four. Enzyme is a homotetramer. Two sites are required for cleavage. Cleaves both recognition sites simultaneously.	(Pingoud & Jeltsch, 2001)
Type IIE	"E" – effector. Two sites are required for cleavage. One site is cut, while the other is used as an allosteric effector to stimulate cleavage.	(Pingoud & Jeltsch, 2001)

Type IIC	"C" – combined. R and M genes are fused. Restriction and modification domains within the same polypeptide.	(Roberts et al., 2003)
Type IIB	"B" – both sides. Cleavage occurs on both sides of the recognition sequence.	(Pingoud & Jeltsch, 2001)
Type IIL	"L" – long distance. Cleavage occurs 20 bp away from the recognition site, a larger shift than in type IIS. Multiple sites are required for cleavage. The genome is protected by hemimethylation.	(Callahan et al., 2016)
Type IIG	SAM is used not only as a methyl group donor but also as an allosteric activator to stimulate cleavage.	(Loenen et al., 2014)
Type IIH	"H" – hybrid. Resemble the composition of type I enzymes but cleave DNA like canonical type II REases.	(Loenen et al., 2014)
Type IIT	"T" – two catalytic sites; heterodimers. Use two different catalytic sites for restriction, each of which is specific for one particular strand.	(Roberts et al., 2003)
Type IIM	"M" – methylated. Recognize specific methylated recognition sites, but cleave the DNA like the canonical type II REases.	(Pingoud & Jeltsch, 2001)

These subtypes are based on specific functional traits of the enzymes and are not mutually exclusive. Therefore, a single enzyme often belongs to multiple groups (Loenen et al., 2014). For example, the type IIA subtype represents all type II enzymes that recognize asymmetric sequences. Although type IIS enzymes are defined by their shifted cleavage properties, many also have asymmetric recognition sites. As a result, the majority of these enzymes also belong to the type IIA subtype (Pingoud et al., 2014). Type IIC is another broad subtype that includes all type II systems with restriction and modification domains encoded within the same protein. All type IIB, IIL, and IIG enzymes are also type IIC enzymes (Roberts et al., 2003).

To understand why these subtypes emerged, it is necessary to consider the functional limitations of canonical RM type II systems in host protection. To begin with, MTase is the sole barrier between protection and self-digestion at every recognition site. Processes such as DNA replication or DNA repair require MTases to constantly update the modification status of every recognition site. Any failure to do so may result in restriction of the host's DNA (Loenen et al., 2014). Type IIF and IIE appear to have evolved as solutions to this problem. The two-site requirement for cleavage effectively reduces the risk of self-digestion because a single unmethylated site is not sufficient to trigger restriction. Type IIG systems that require SAM for cleavage also help prevent the restriction of host DNA. SAM requirement ensures that the

cofactor necessary for methylation is present before the restriction can happen (Pingoud et al., 2014).

Another major downside of the canonical type II systems for the host is their limited evolutionary flexibility. It would be beneficial for the host to evolve new target sequence specificities in response to adapting phages. However, in canonical type II systems, REase and MTase are separate proteins with their own DNA-binding domains. Therefore, a mutation that alters the specificity of one enzyme but does not change the specificity of another is toxic to the host (Callahan et al., 2016). Type IIC enzymes solved this problem by fusing the R and M domains, allowing them to share a single target recognition domain. As a result, any change in specificity translates to both restriction and modification activities. This makes specificity changes far more accessible for type II RM systems (Callahan et al., 2016; Loenen et al., 2014).

1.5 *Pseudomonas putida* and its biotechnological relevance

Pseudomonas putida is a Gram-negative, saprotrophic, mesophilic soil bacterium that has been extensively studied for its huge potential as a leading chassis for industrial biotechnology. This potential is connected to its several important qualities. First and foremost, having evolved in harsh soil environments, this bacterium can endure many physicochemical stressors such as heavy metals, organic solvents, and temperature changes. Another advantage is that its metabolism is tuned to generate more NAD(P)H than ATP. Having a higher amount of reducing cofactors makes the bacterium more resilient to high levels of oxidative stress. Therefore, *P. putida* is a highly suitable host for redox-intensive reactions, which are a common limiting factor for product yield in many industrial fermentations (de Lorenzo et al., 2024). Additionally, this bacterium is highly compatible with many standardized synthetic biology tools. This allows the insertion of heterologous pathways for specific industrial purposes through genetic engineering (Nikel & de Lorenzo, 2018).

One of the most relevant features of *P. putida* is its ability to degrade numerous aromatic compounds (Belda et al., 2016). This property has been used in both industrial and environmental applications. For example, modified strains of *P. putida* are used for the production of certain bioplastics from lignin. Several pharmaceuticals are also produced by industrial strains of this bacterium (de Lorenzo et al., 2024). As for environmental applications, *P. putida* strains are used for the bioremediation of lands polluted with toxic aromatic compounds. Another property of this bacterium, which is the ability to colonize the plant rhizosphere and protect against pathogens, has been leveraged in agriculture (Belda et al., 2016).

The PaW85 strain was derived from the soil isolate mt-2 by removing the large catabolic plasmid, pWW0, which is responsible for degrading toluene and xylenes (de Lorenzo et al., 2024). It is therefore isogenic to the widely used industrial KT2440 strain and is the strain used in this work.

1.5.1 Relevance of antiphage defense systems in biotechnology

One of the key challenges in industrial fermentations is phage contamination, which can either hinder production or cause complete batch failure. Therefore, identification and systematic comparison of antiphage defense systems in biotechnologically relevant microorganisms enable the selection of strains with greater immunity against phages (Xu et al., 2025). For example, three different plasmid-encoded systems have been recently validated in certain *Lactococcus* strains. These systems can be transferred to sensitive industrial strains through conjugation. This is particularly beneficial for the food industry, since these strains will not be classified as genetically modified organisms. Consequently, consumer opposition and regulatory hurdles can be avoided (Grafakou et al., 2024).

Phages, in turn, rapidly evolve to bypass bacterial defense systems. Thus, it is necessary to understand the mechanisms behind these systems to engineer the strains with enhanced defensive capabilities without compromising product yields. For example, CRISPR/Cas9 systems can be programmed in *E. coli* strains to include a broad range of phage-specific spacers and provide long-lasting defense. Furthermore, multiple systems with diverse mechanisms can be combined to create multi-layered resistance and significantly boost bacterial defense (Xu et al., 2025).

1.5.2 CEPEST

To effectively study a broad repertoire of antiphage defense systems in *P. putida*, a rich phage library is needed. CEPEST is currently the largest collection of environmental phages for *P. putida*. Initially, a deletion mutant of the PaW85 strain, lacking 4 prophages and 13 toxin-antitoxin systems, was used to isolate the phages from muddy water and soil samples (Brauer et al., 2024). Later, PaW85 mutants harboring more deletions of predicted phage defense loci (Ainelo, unpublished) were used to expand the library, which now consists of over 150 phage isolates. These phages were grouped into 76 species clusters (up to 95% genome identity) and 28 genus clusters (70% genome identity) using sequence identity scores and proteome analysis (Brauer, unpublished).

2 THE AIMS OF THE THESIS

The new phages in the CEPEST collection were isolated with a strain lacking predicted phage defense loci, the functions of which have not been experimentally proven. The main aims of this thesis were:

- Firstly, to identify a functional defense system in *P. putida* PaW85.
- Secondly, to describe the identified RM-II system and its functionality, which involved:
    - Analyzing the RMII locus;
    - Determining the roles of each methylase present in the operon.

3 EXPERIMENTAL PART

3.1 MATERIALS AND METHODS

3.1.1 Bacterial strains, plasmids, primers and media

Growth media: LB (Lysogeny broth) medium (1% tryptone, 0.5% yeast extract, 0.5% NaCl) was used for bacterial liquid cultures. LB agar (1.5% agar) and LB soft agar (0.3% agar) were used for solid media and for the top surface of double agar overlay plates, accordingly. 10 mM of $CaCl_2$ was added to the LB soft agar to increase phage infection efficiency and ciprofloxacin (Cip) at a final concentration of 0.03 μg/mL was added to LB agar for phage susceptibility assays. For cloning procedures, antibiotics kanamycin (Km, final concentration 50 μg/mL), gentamycin (Gm, final concentration 10 μg/mL), and streptomycin (Sm, final concentration 200 μg/mL) were used for selection. Anhydrotetracycline (ATC) was used to induce the expression of the restrictase. SM buffer (sodium-magnesium buffer; 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 8 mM NaCl, 8mM $MgSO_4$, 0.01% gelatin) was used for diluting and storing bacteriophages. Bacterial strains used in the study are listed in Table 2.

Table 2.

Strain	Description	Source
Escherichia coli DH5α λpir	*endA1 hsdR17 glnV44 (= supE44) thi-1 recA1 gyrA96 relA1 φ80dlacΔ(lacZ)M15 Δ(lacZYA-argF)U169 zdg232::Tn10 uidA::pir+*	(Platt et al., 2000)
Escherichia coli BL21(DE3)	*hsd gal (λcIts857 ind1 Sam7 nin5 lacUV5-T7 geen1*)	(Fuerst et al., 1986)
Pseudomonas putida PaW85	Wild-type, isogenic to KT2440	(Bayley et al., 1977)
PaW85 ΔDef2	*P. putida* PaW85 with deletions of 13 toxin antitoxin systems and 4 cryptic prophages	(Brauer et al 2024)
PaW85 ΔDef8	PaW85 ΔDef2 with deletions of predicted phage defense loci Gabija, Wadjet, RMI, HerA-DUF and RMII and PD-T7-1	S. Rosendahl
PaW85 ΔDef12	PaW85 ΔDef8 with deletions of predicted phage defense loci RloC, PDC-S12, PDC-S21 and the M2 of RMII	A. Ainelo
PaW85 ΔPP_3988-3989	*P. putida* with a deletion of RM1 from RMII	I. Polekauskaite
PaW85 ΔPP_5651-3988	*P. putida* with a deletion of RM2 from RMII	This project
PaW85 ΔPP_5651-3988-3989	*P. putida* with a deletion of RM1M2 from RMII	This project

Plasmids used for bacterial cloning are listed in Table 3 and primers in Table 4.

Table 3.

Plasmid	Description	Source
pSNW2	Used for genomic integration of constructed loci into *P. putida*	(Volke et al., 2020)
pSNW2-ΔPP_5651	Contains *M2* deletion locus (from RMII)	This project
pSNW2-ΔPP_5651-3988	Contains *RM2* deletion locus (from RMII)	This project
pQURE-6	Contains the I-SceI nuclease to induce homologous recombination through the double-strand break	(Volke et al., 2021)
pET24d-His10SUMO	Used for recombinant protein expression (*R* gene in this work) in *Escherichia coli* BL21(DE3). Contains His10SUMO Tag.	A. Ainelo
pET24d-His10SUMO+PP_3988	pET24d-His10SUMO plasmid for the *R* gene expression.	This project
babyTOL-PtetO-ISc3	Used for creating the *R* gene expression plasmid for *P. putida* and as a control in the study of R expression effects	M. Hinnu
babyTOL-PtetO-His10SUMO-PP_3988	Contains the *R* gene (from RMII) and His10-SUMO tag for purifying the protein	This project
babyTOL-PtetO-PP_3988	Contains the *R* gene (from RMII)	This project

Table 4.

Primer name	Sequence	Application
pr36	GATCTAGATGGACGGCGTAATCCTGTTC	Creation of ΔPP_5651
pr139	ATCAGTTGTGAATTGTGCTC	Creation of ΔPP_5651
pr140	TAGAATTCCAGACTGCGACGATAAATAC	Creation of ΔPP_5651 and ΔPP_5651- 3988
pr141	GAGCACAATTCACAACTGATGATGGTTTTTGCCTGTAAAAAAATGC	Creation of ΔPP_5651
pr57	TAAGCTTGGCTGTTTTGGC	Amplification of pET24dHis10SUMO
pr58	ACCACCAATCTGTTCGCG	Amplification of pET24dHis10SUMO or His10SUMO tag
pr142	ATCGCGAACAGATTGGTGGTGACGTCTCATTAGATGCAG	Amplification of PP_3988

pr143	CGCCAAAACAGCCAAGCTTACTTTC GAATCTCCAGAAC	Amplification of PP_3988
pr144	TAAAAGCCGAATTCCAGC	Amplification of babyTOL-PtetO
pr145	GCTATGGATCCAATCGCC	Amplification of babyTOL-PtetO
pr146	CTGGCGATTGGATCCATAGCTTTGTT TAACTTTAAGAAGGAGATATACCATG	Amplification of His10SUMO tag, verifying *R* expression construct
pr147	GTGCTGGAATTCGGCTTTTACTTTCG AATCTCCAGAAC	Amplification of PP_3988
pr150	GATCTAGAGTGTTGCTCAGCATCGAAG A	Creation of ΔPP_5651- 3988
pr151	GAGCACAATTCACAACTGATGTATCT CCCCTTTACAAAAGGAA	Creation of ΔPP_5651- 3988
pr40	AAGTTTCGTGCCTTGGTGGA	verifying *R* expression construct
P-ISc3_frw	GTCCTTTTCCGCTGCATAACCTC	verifying *R* expression construct
V1-bfp	TGCCATGGTTGGTCGACTATACTAGC GGATTTGAACGTTGCGAAG	verifying *R* expression construct
fd-term_R	GGACTCCAGTATCGCGAGCCTGATTA AACATTG	verifying *R* expression construct
Fw	GTAAAACGACGGCCAGT	Verifying deletion locus constructs
Rev	CAGGAAACAGCTATGAC	Verifying deletion locus constructs
pr152	CATGGTATATCTCCTTCTTAAAG	Deletion of His10SUMO tag
pr153	GACGTCTCATTAGATGCAG	Deletion of His10SUMO tag

3.1.2 Phages

Bacteriophages used in phage susceptibility assays originate from the widened CEPEST collection of the Institute of Molecular and Cell Biology (unpublished). The used phages shown in the thesis are listed in Table 5.

Table 5.

Phage species	Phage name	The isolator of the phage
16A	Abava-3	S. Rosendahl
16C	Peipsi-1	S. Rosendahl
17A	Kaapa-2	S. Rosendahl
17C	Kaapa-6	S. Rosendahl
17F	Elva-6	K. Männik

3.1.3 Phage susceptibility assays

Bacterial cultures were grown overnight in LB medium. Then, 50 µL of bacterial culture and 50 µL of 1 M $CaCl_2$ were mixed with 5 mL of LB soft agar, which was poured onto solid LB agar plates containing Ciprofloxacin. Afterward, 10-fold serial dilutions of phage stocks were prepared in SM buffer, with dilution factors ranging from 10^0 to 10^8. Finally, 1.5 µL drops of each dilution were spotted on the double-layer agar plates and incubated overnight at 20 °C. Phage susceptibility assays were repeated three times.

3.1.4 Cloning

3.1.4.1 Construction of *P. putida* mutant strains

The specific genomic loci in *P. putida* were deleted using the method of Martínez-García & de Lorenzo, (2011). At the start, the deletion locus is cloned into the pSNW2 plasmid. This plasmid cannot be replicated in *P. putida*. Therefore, under antibiotic pressure, it is kept in the cell via integration driven by site-specific homologous recombination between the introduced deletion locus and homologous genomic DNA. In this work, to generate the deletion loci for later insertion into pSNW2, flanking regions (~500 bp) of the deletion site were amplified using the following primer pairs:

1. pr36 + pr141 and pr139 + pr140 for ΔPP_5651
2. pr139 + pr140 and pr150 + pr151 for ΔPP_5651-3988

The two PCR fragments for each deletion were joined by overlap extension PCR using the two outer primers, which also contain restriction enzyme sites in their 5` ends. The obtained fragments were cloned into pSNW2 using XbaI and EcoRI. Then, the plasmid was propagated in *Escherichia coli* DH5α λpir on LB plates containing Km, checked with PCR using oligos Fw and Rev and positive constructs were sent to sequencing (Core Facility of Genomics, University of Tartu). Next, the correct purified plasmids were transformed into *P. putida* (pSNW2-ΔPP_5651 into PaW85 ΔPP_3988-3989 and pSNW2-ΔPP_5651-3988 into PaW85), and co-integrates were formed. Thereafter, the pQURE-6 plasmid was introduced, from which the I-SceI nuclease expression was induced with 2 mM metatoluate to generate double-stranded breaks at the integrated plasmid in order to excise it from the genome. Then, the cells were plated on LB plates containing Gm and metatoluate (2 mM). Finally, the PCR with outer primers was performed to verify the deletion. The clones where the deletion was successful were selected and stored in glycerol stocks at -80 °C.

3.1.4.2 Restrictase expression in both *E. coli* and *P. putida*

The Gibson assembly was used for cloning of:

- the restrictase gene (PP_3988) into pET24d-His10SUMO for expression in *Escherichia coli* BL21(DE3). The following primer pairs were used: pr57+pr58 (amplification of the vector) and pr142+pr143 (amplification of the PP_3988).
- both the restrictase gene (PP_3988) and His10SUMO Tag into babyTOL-PtetO-ISc3 instead of the *ISc3* gene for expression in *P. putida* PaW85 and derived mutant strains. (The following primer pairs were used: pr144+pr145 (amplification of the vector), pr142+pr147 (amplification of the PP_3988), and pr146+pr58 (amplification of the His10SUMO Tag).

Inverse PCR with primers pr152 and pr 153 was used to delete the His10SUMO Tag from babyTOL-PtetO- His10+SUMO+PP_3988, and the plasmid, which contains the restrictase gene without the His10SUMO Tag (babyTOL-PtetO-PP_3988), was created by ligating the obtained PCR product together. The ligation mixture was electroporated into P. putida PaW85, and the obtained colonies were screened for the deletion of the His10SUMO tag. The primers used for screening and sequencing the correct constructs were P-ISc3_frw or fd-term_R and V1-bfp, and for sequencing the middle region of the product, primers pr40 and pr146 were used (Table 4).

All cloning reactions were performed with Thermo Fisher FastDigest restriction enzymes and T4 DNA ligase according to the manufacturer's protocols.

The plasmids were transformed by electroporation and purified using the FavorPrep Plasmid DNA Extraction Mini kit, (Favorgen) according to the manufacturer's instructions. All obtained plasmids were verified by sequencing at the University of Tartu Core Facility of Genomics.

3.1.5 Measuring bacterial growth

Bacterial cultures were grown overnight. Then, the cells were diluted in LB media in the wells of a 96-well plate to an OD_{580} = 0.1. Afterwards, the plate reader (BioTek Synergy H1) was used to measure bacterial growth overnight at 30 °C without shaking. The obtained OD values have been used to construct the growth curves. The measurements of bacterial growth were taken every 10 minutes.

3.1.6 Bioinformatic analysis of the RMII genes

The sequences of the methylase genes were aligned in UniProt, and the phylogenetic tree, rooted by midpoint, was obtained (The UniProt Consortium, 2025). While Foldseek was used to perform a homology search by structure alignment, the pairwise structure alignment was performed using the tool available in PDB (Bittrich et al., 2024; van Kempen et al., 2024). The

analysis used predicted Alphafold 3D protein structures. The Alphafold server itself was used to model the homodimer of the restrictase gene (*PP_3988*) (Abramson et al., 2024).

3.2 RESULTS

3.2.1 Predicted Restriction-Modification type II in *P. putida* reduces phage infection efficiency

A variety of antiphage defense systems were detected using defense prediction tools. To test whether these potential defense systems confer resistance against phages, a broad screening with all the phages from the published CEPEST collection (Brauer et al., 2024) was conducted. The degrees of phage infection were compared between the following two *P. putida* strains: PaW85 ΔDef2 and PaW85 ΔDef8, which lacks six additional defense systems. The changes in infection efficiencies of different *P. putida* deletion strains were determined by increases in phage plaque numbers within higher phage dilutions with respect to the WT PaW85 or PaW85 ΔDef2 strains. The PaW85 ΔDef8 mutant started to show higher susceptibility to most phages from genus clusters G5 to G13 (data not shown). As the phage collection expanded, multiple phages from genus clusters G16 and G17 were also tested, and these phages also demonstrated significant titre increases on the PaW85 ΔDef8 strain (data not shown). The same phages were then used to infect the single-deletion strains to determine which of the missing systems in the PaW85 ΔDef8 mutant provided protection. One of the systems was validated as a defense system. This system is the predicted RM type II system (RMII) encoded by genes *PP_3988* and *PP_3989* (A. Ainelo, unpublished). The RMII deletion mutant consistently demonstrated elevated susceptibility to phages 16C Peipsi-1, 17A Kaapa-2, 17C Kaapa-6 and 17F Elva-6, with titres reaching a 200-fold difference (Figure 6).

Given the substantial number of RMII system subtypes, a more detailed genomic analysis was needed to elucidate the potential mechanisms underlying the system's function.

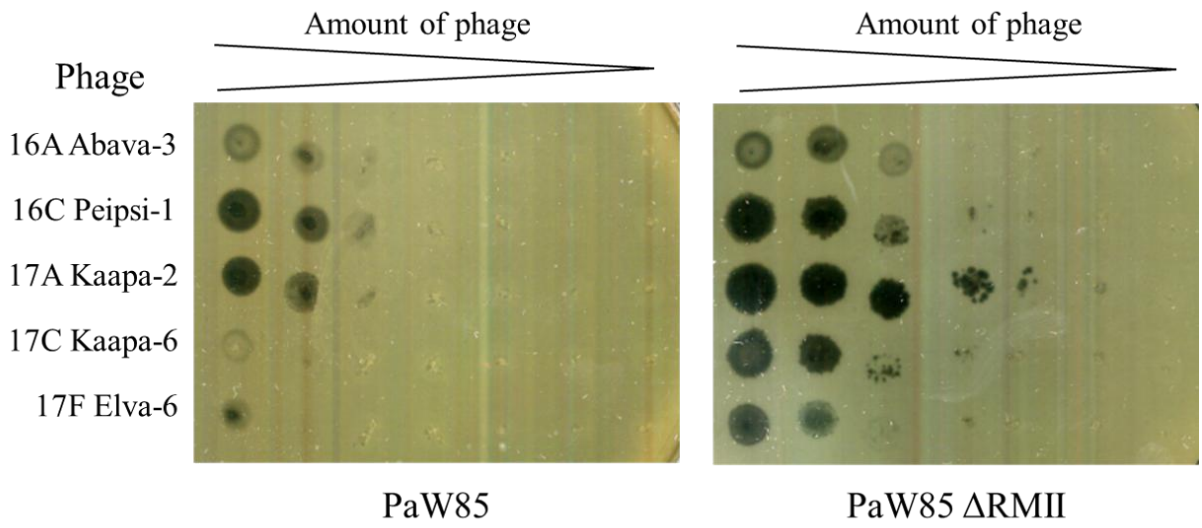

Figure *6*. Phage susceptibility of WT(PaW85) vs ΔRMII. The WT(on the left) and the ΔRMII mutant (on the right) were infected with the phages listed on the left side of the figure.

3.2.2 Analysis of the RMII locus in *P. putida* PaW85

During the inspection of the RMII locus, in addition to the previously predicted genes termed *PP_3988* and *PP_3989*, a third gene, *PP_5651*, was identified in the same operon. However, only *PP_3989* was annotated as a type II methylase in the Pseudomonas database (Winsor et al., 2016), while *PP_3988* was predicted to be the corresponding type II restrictase by DefenseFinder, one of the defense system prediction tools (Tesson et al., 2024). Since *PP_5651* is located with the PP_3988 and PP_3989 genes in the same operon, this suggests a functional link between the genes. The results of structural homology searches based on its AlphaFold model in Foldseek indicate that it is also a methylase, as the top hits across various databases are mostly DNA methyltransferases. Therefore, in this work, the *PP_3988*, *PP_3989*, and *PP_5651* genes were labeled as *R*, *M1*, and *M2*, respectively (Figure 7).

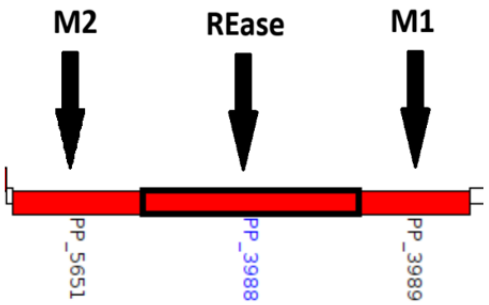

Figure *7*. RMII operon in *P. putida*. The genes are arranged in the genome in the following way: *PP_5651/M2, PP_3988/R* and *PP_3989/M1*. Gene lengths are drawn approximately to scale.

To further investigate the genomic association of these genes, webFlaGs (Saha et al., 2021) was used to determine which of them is more likely to be found together in diverse bacterial genomes (Figure 8). The results show that *R*-like genes are more commonly found with *M2*-like genes than with *M1*-like genes. In fact, there are no instances of the *R* gene appearing without the *M2*, while the *M1* gene is missing in some cases. The presence of a second methylase in the locus sets the RMII system in *P. putida* apart from the canonical RM type II systems. It also raises the questions of whether both methylases are functional and whether one is sufficient to protect the *P. putida* genome from self-digestion.

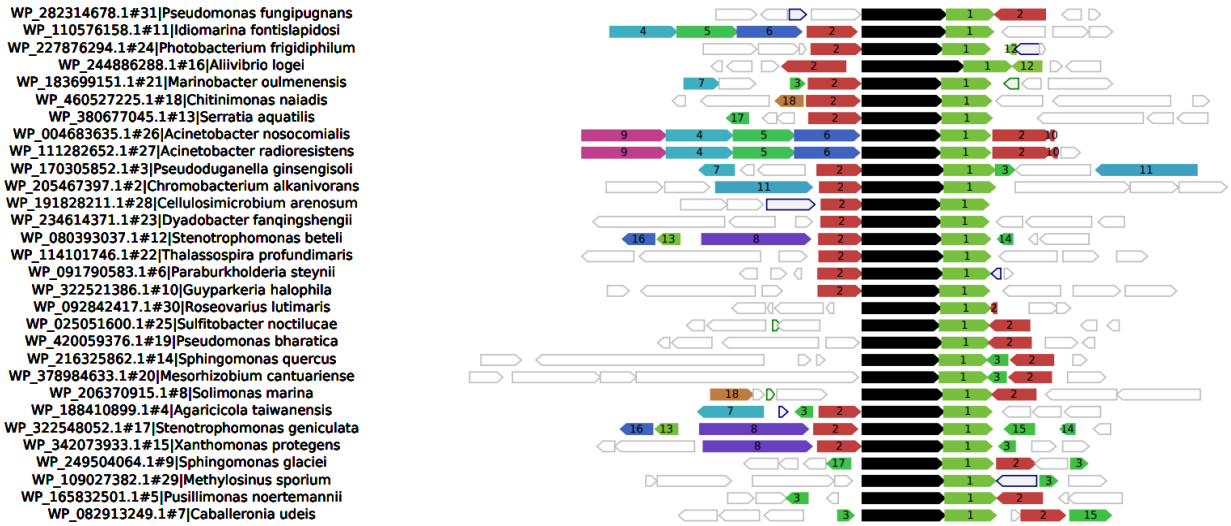

Figure *8*. webFlaGs output for the genes of the RMII locus. The *R*-like genes (black) always co-occur with *M2*-like genes (light green with #1) and often with *M1*-like genes (red with #2).

3.2.3 Attempt to express the restrictase in *E. coli*

After analysis of the RM locus, I attempted to express and purify the putative restriction enzyme (PP_3988) in order to verify its nuclease activity and characterize it further. For that, I cloned the *R* gene into the pET24d-His10SUMO plasmid, which contains a T7 promoter for controlled expression of the inserted gene in T7 RNA polymerase-equipped *Escherichia coli* BL21(DE3). This plasmid also contains the His10SUMO Tag, needed for purification of the expressed protein via affinity chromatography. The constructed plasmid (pET24d-His10SUMO+PP_3988) was transformed into *Escherichia coli* BL21(DE3) via electroporation. Afterwards, the obtained colonies were checked with the colony PCR, and the amplified fragments were put on 1% agarose gel (Figure 9). All these DNA fragments are of incorrect size, suggesting that the restriction enzyme is toxic to the cell and that this *E. coli* strain lacks the methylation pattern needed to prevent cleavage by the enzyme. Therefore, only the colonies that have lost or inactivated the inserted *R* gene survived. These results demonstrate

that the restrictase gene must be expressed in the WT *P. putida* strain that has methylated recognition sites. Thus, the *R* gene was cloned into the babyTOL-PtetO-ISc3 (inducible) plasmid for expression in *P. putida*, which is controlled by anhydrotetracycline (ATC) induction. Since the goal was to purify the restrictase, the His10 SUMO Tag has been added to the N-terminus of the *R* gene. The resulting babyTOL-PtetO- His10+SUMO+PP_3988 plasmid was constructed via the Gibson assembly.

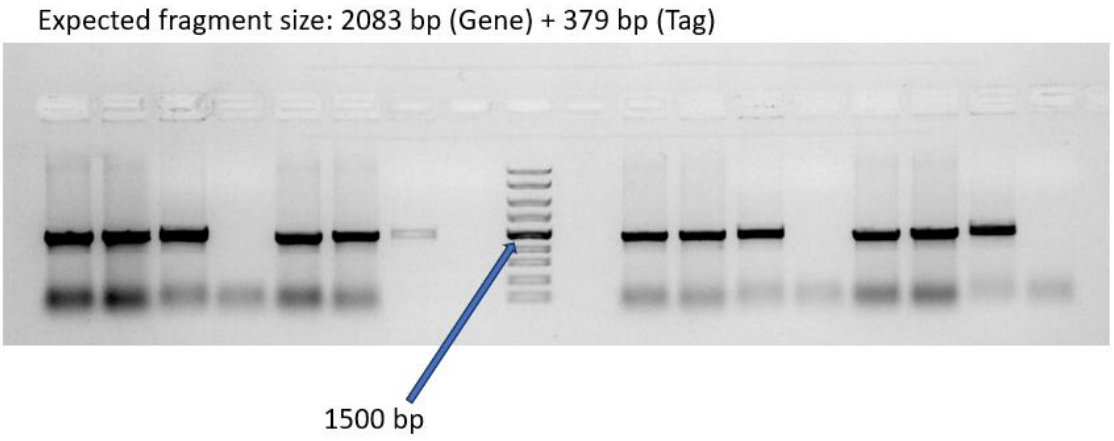

Figure *9*. Amplified insert of pET24d-His10SUMO+PP_3988 plasmid. Agarose gel (1%) showing colony PCR products from E. coli BL21(DE3) colonies transformed with the pET24d-His10SUMO+PP_3988 plasmid. All amplified fragments are of incorrect size relative to the expected product (2462 bp), indicating that clones retaining an intact R gene did not survive.

3.2.4 Verifying the roles of each methylase present in the RMII locus

When the RMII system was predicted, only the *M1* was identified as a methylase. Therefore, during the construction of the first deletion strain, the *M2* gene was not taken into account, and only the *R* and *M1* genes were deleted from the WT strain. Thus, the resulting mutant is essentially ΔRΔM1 (PaW85 ΔPP_3988-3989). This mutant was used for phage susceptibility assays. Although not all of the genes were deleted from the system, the ΔRΔM1 mutant was sufficient to test whether the loss of the system increases phage susceptibility, because the putative restriction enzyme is the one that interferes with the phage infection. However, to verify the functional roles of each methylase, two more deletion strains were constructed: ΔRΔM2 and ΔRM1M2. The ΔRΔM2 (PaW85 ΔPP_5651-3988) strain was obtained by deleting the *R* and *M2* genes from the WT PaW85 strain. Meanwhile, the ΔRM1M2 (PaW85 ΔPP_5651-3988-3989) strain was generated by deleting the *M2* gene from the ΔRΔM1 strain. As a result, all mutants lack the restrictase gene and either of the methylases (ΔRΔM1, ΔRΔM2) or both (ΔRΔM1ΔM2).

To determine which methylase is capable of protecting the genome via methylation, the restrictase was reintroduced into all mutant strains and the WT strain as a control. This experiment was based on the assumption that expression of the restriction enzyme in a strain that lacks methylase genes needed for protection would result in self-restriction and growth inhibition. On this basis, the ΔRΔM1ΔM2 strain was expected to grow poorly, since it lacks both methylases. In contrast, the growth phenotypes of the ΔRΔM1 and ΔRΔM2 would indicate whether M1 or M2 alone is sufficient to prevent restriction. For example, if M1 is functional while M2 is not, growth disruption would be expected in ΔRΔM1 and ΔRΔM1ΔM2 strains, whereas ΔRΔM2 would be expected to grow similarly to the WT strain. Additionally, the PaW85 ΔDef12 strain, which fully lacks RMII (ΔRΔM1ΔM2) and RMI systems, was used in the experiment because it was suspected that the recognition sites could still be protected by the methylation activity of the RMI system in *P. putida*. The cells were electroporated with both the initial babyTOL-PtetO-ISc3 plasmid for the negative control and the babyTOL-PtetO-His10+SUMO+PP_3988 plasmid, which, in addition to the R gene, also contains the His10SUMO Tag for purification. Thereafter, the cells were plated as spots of serial 10-fold bacterial dilutions in LB (10^0-10^5) on 3 different sets of Sm LB plates:

1. LB plate with Sm – no induction of the R gene expression.
2. LB plate with Sm and ATC (final concentration 1 μg/mL) – induction of the *R* gene expression.
3. LB plate with Sm and ATC (final concentration 10 μg/mL) – induction of the *R* gene expression

The induction of gene expression in babyTOL-PtetO-ISc3 by ATC was previously tested only in liquid culture. Therefore, two largely different inducer concentrations were used in the experiments: 1 and 10 μg/mL.

The results do not align with the previously made assumptions (Figure 10). There are no significant differences in growth among the strains or CFUs depending on the expression of the restrictase. Most strains transformed with the babyTOL-PtetO-ISc3 plasmid have fewer colonies than those transformed with the restrictase-containing plasmid. Also, the overall growth on plates with ATC (final concentration 10 μg/mL) is poorer. Thus, ATC (final concentration 1 μg/mL) was used in the subsequent experiments. However, it is important to note that the number of colonies also depends on the transformation efficiency. Overall, no meaningful conclusions could be drawn regarding the functionality of methylases from growing the bacteria on solid media straight after electroporation.

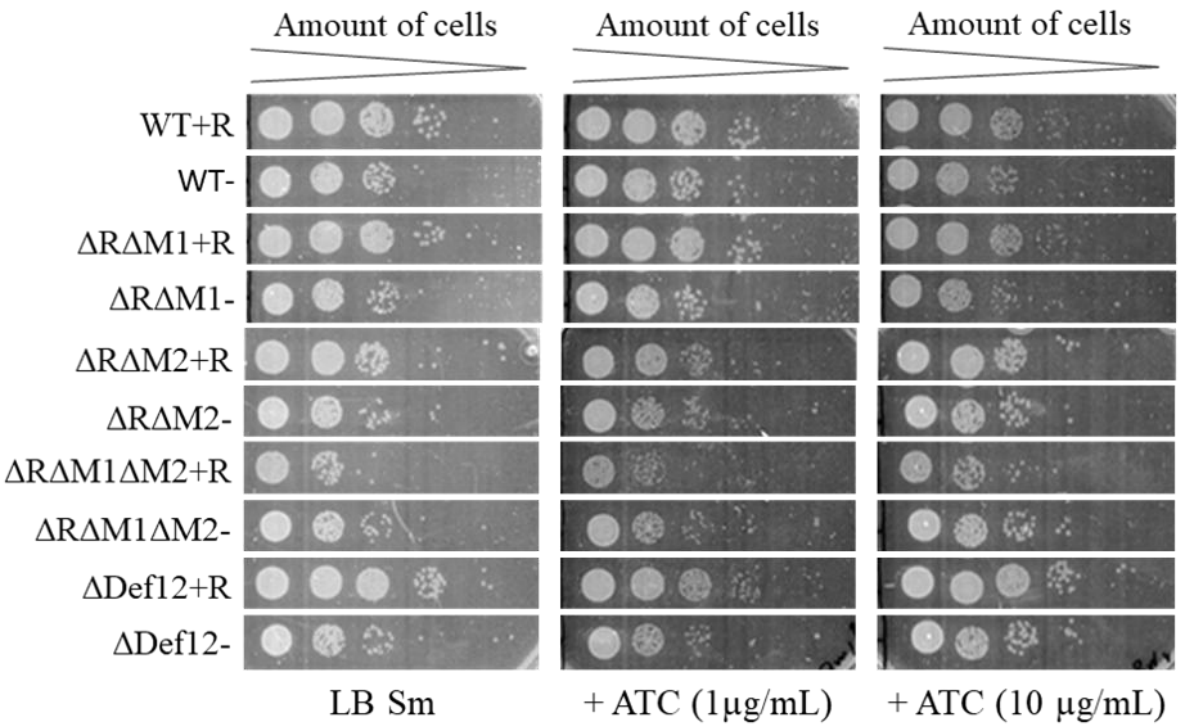

Figure *10*. The strains that were transformed with the initial babyTOL-PtetO-ISc3 plasmid (WT-, ΔRΔM1-, ΔRΔM2-, ΔRΔM1ΔM2- and ΔDef12-) and the babyTOL-PtetO-His10+SUMO+PP_3988 plasmid (WT, ΔRΔM1, ΔRΔM2, ΔRΔM1ΔM2 and ΔDef12). The first column: cells on LB Sm plates. The second column: cells on LB plates with Sm and ATC (1 μg/mL). The third column: LB plates with Sm and ATC (10 μg/mL). The strain names corresponding to bacterial spots are listed on the left side of the figure.

As all constructs were viable and formed colonies after electroporation, liquid LB cultures containing Sm were started from transformed colonies to avoid bias arising from transformation efficiency differences. Then, bacteria were diluted and plated in a similar manner as before. The 10-fold dilutions of the overnight cultures were made and spotted on 2 different sets of Sm LB plates:

1. LB plate with Sm – no induction of the *R* gene expression.
2. LB plate with Sm and ATC (1 μg/mL) – induction of the *R* gene expression.

Once again, no drastic differences in growth were observed between the strains (Figure 11). However, the colonies of both ΔRΔM1ΔM2 and ΔDef12, the strains that lack both of the two methylases, seem slightly more transparent compared to the other strains, but the effect is only barely visible. The absence of the expected reduction of CFUs may be due to the fact that ATC-inducible expression from the babyTOL-PtetO-ISc3 plasmid had previously only been demonstrated in liquid culture and might remain low enough to allow cell growth. However, even without induction, the effects of restrictase toxicity might have been stronger. Since the

His10SUMO tag was fused to the R protein in the construct made, it is possible that the tag interfered with proper enzyme activity.

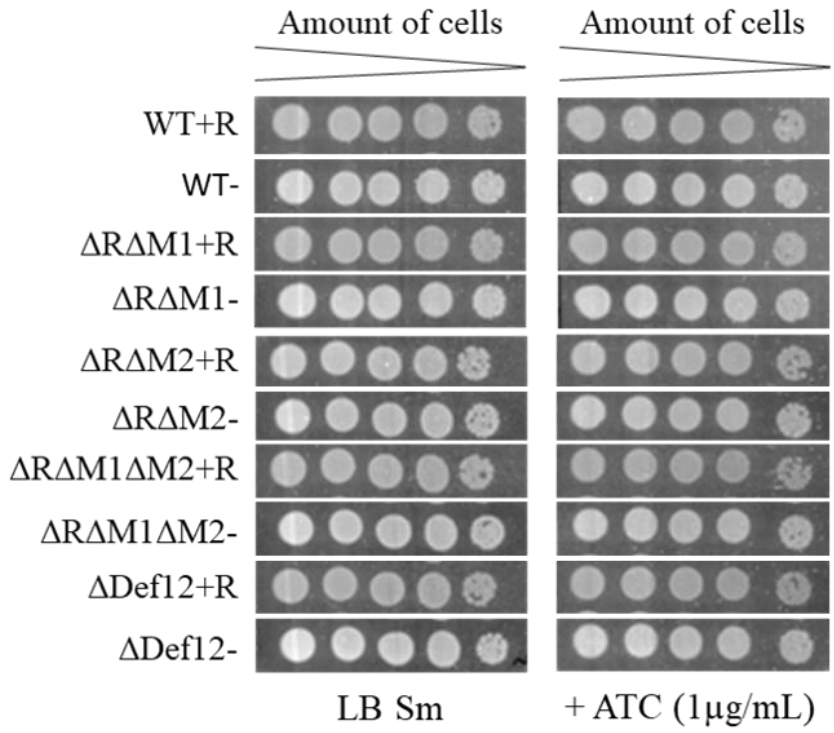

Figure *11*. Transformed strains grown in liquid LB with an antibiotic overnight. The first column: LB plates with Sm. The second column: LB plates with Sm and ATC (1 µg/mL). The strain names corresponding to bacterial spots are listed on the left side of the figure.

To further investigate the effects of restrictase expression, the growth of WT, ΔRΔM1, ΔRΔM2, and ΔRΔM1ΔM2 *P. putida* strains was monitored in liquid LB medium with and without ATC using a plate reader. Each strain was transformed with the following plasmids:

1. babyTOL-PtetO-His10+SUMO+PP_3988 – contains both the *R* gene and His10SUMO Tag. Further referred to as "Tag+".

2. babyTOL-PtetO-PP_3988 – contains only the *R* gene. Further referred to as "Tag-".

3. babyTOL-PtetO-ISc3 – the plasmid without the restrictase. Further referred to as "Control".

The transformed cells were cultured overnight in LB media supplemented with Sm. Then, each strain was diluted to OD580 = 0.1 in LB+Sm media and in LB+Sm+ATC media. The dilutions were made in a 96-well plate, which was sealed with parafilm and placed in the plate reader to monitor bacterial growth overnight via OD580 measurements. The obtained growth data were used to construct the growth curves for each strain (Figure 12).

The growth curves demonstrated that when all strains are transformed with a Tag- plasmid and the restrictase expression is induced, only the growth of the ΔRΔM1ΔM2 strain is disrupted, while the strains that have only one of the methylases deleted (ΔRΔM1 and ΔRΔM2) grow similarly to the WT strain. This means that the restrictase becomes toxic only when both methylases are deleted. Therefore, both M1 and M2 can protect the genome via methylation. This effect was even more pronounced when a tagless restrictase was produced. Also, even when the expression of the tagless restrictase was not induced, the ΔRΔM1ΔM2 strain had an unusual growth pattern, probably due to background expression of the restriction enzyme. However, the inhibition was not as consistent as with the induction. Furthermore, the results indicate that His10SUMO Tag interferes with restriction enzyme activity, as only a slight growth disruption was observed when the expression of tagged restrictase was induced in the ΔRΔM1ΔM2 strain, and no growth defect was observed without induction.

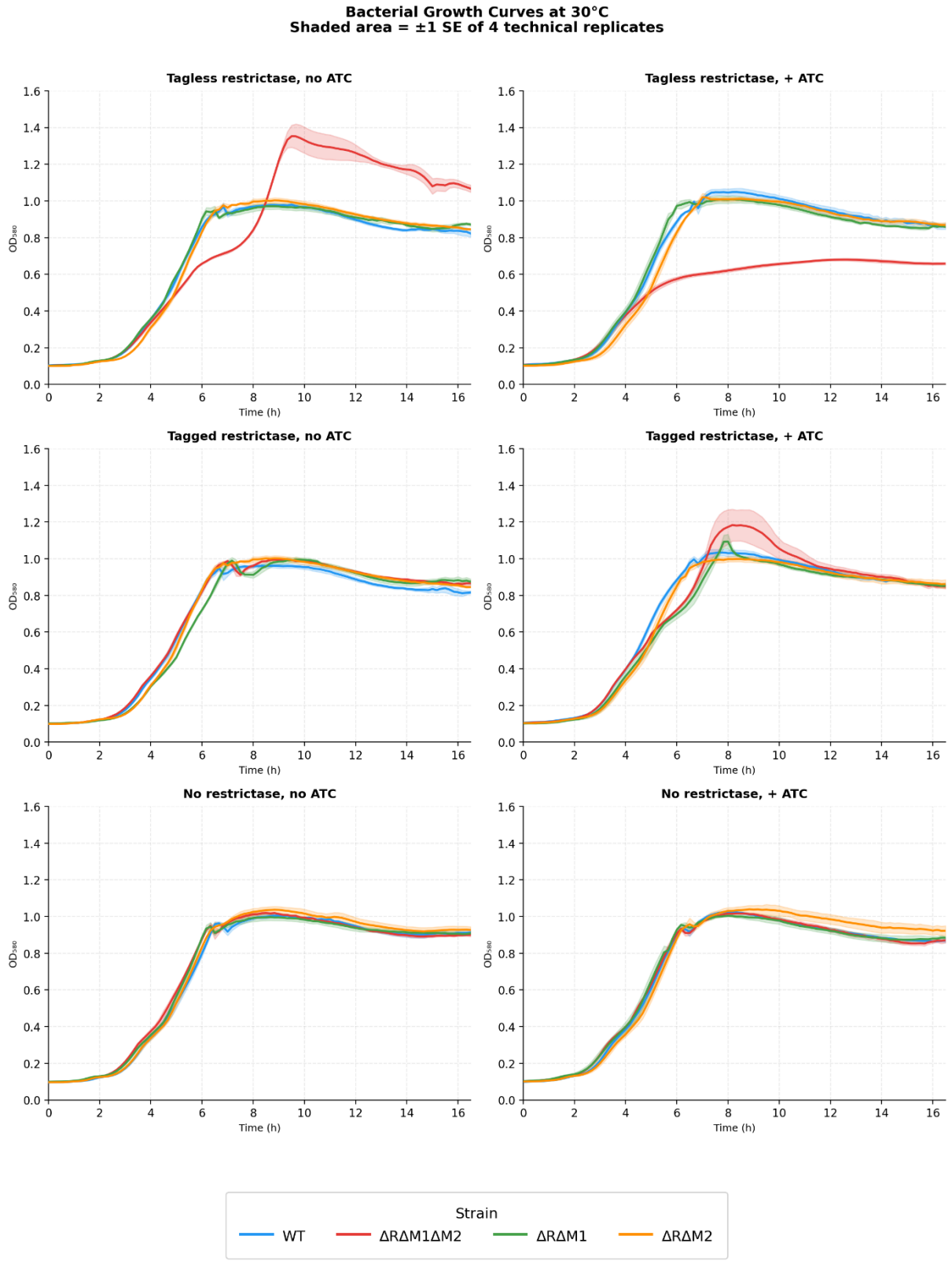

Figure *12*. Growth curves of WT, ΔRΔM1, ΔRΔM2, and ΔRΔM1ΔM2 P. putida strains. The strains express either a tagged restrictase (Tag+), tagless restrictase (Tag-), or a control plasmid without the R gene (Control), with and without ATC induction (1 µg/mL). OD580 was measured every 10 minutes overnight in a 96-well plate reader at 30°C. Growth inhibition in the ΔRΔM1ΔM2 strain upon induction of tagless restrictase expression demonstrates that both *M1* and *M2* are individually sufficient to protect the host genome from self-restriction.

The presence of two functional methylase genes suggests redundancy, yet both genes have been retained in *P. putida*. Therefore, a more detailed analysis of the methylase proteins is needed. Furthermore, RM type II systems with two methylase genes have been described in other bacteria.

3.2.5 Bioinformatic analysis of the genes present in the RMII locus of *P. putida*

Two separate methylase genes are commonly encoded in the RM type II systems when the recognition sequence is asymmetric and requires each enzyme to methylate one strand of the DNA. For example, the *HgaI* RM type II system targets the asymmetric sequence: 5′-GACGC-3′ / 3′-CTGCG-5′. Two methylases modify the internal cytosine on different strands, each using a separate target-binding domain to recognize the strand-specific sequence (Sugisaki et al., 1991).

For comparative analysis, methylase pairs from several well-characterized type II RM systems containing two methylases were selected. These included methylases from the *MboII* system of *Moraxella bovis* [*M1.MboII* (UniProt: P23192) and *M2.MboII* (UniProt: A0A097J9V0)], the *HphI* system of *Haemophilus parahaemolyticus* [*M1.HphI* (UniProt: P50192) and *M2.HphI* (UniProt: P50193)], and the *HgaI* system of *Haemophilus gallinarum* [*M1.HgaI* (UniProt: P25282) and *M2.HgaI* (UniProt: P25283)]. These were compared with the *P. putida* methylases analyzed in this study, *M1* (UniProt: Q88FU3) and *M2* (UniProt: A0A140FWK1).

The results are summarized in Tables 6 and 7: Table 6 presents the structural alignment of all proteins against the *P. putida M1* protein, and Table 7 presents the corresponding alignment results against the *P. putida M2* protein. Sequence identity is used as a measure of amino acid similarity, while TM-score is used to evaluate structural similarity. Proteins with a TM-score above 0.5 are generally considered structurally similar.

Table 6. Comparison of the *M1* protein of *P. putida* to other M proteins

Entry	TM-score	Sequence identity
M1_*P. putida* (compared with all the following entries)	-	-
M2_*P.putida*	0.29	5%
M1.MboII	0.32	7%
M2.MboII	0.3	9%
M1.HphI	0.77	29%
M2.HphI	0.28	8%
M1.HgaI	0.77	20%
M2.HgaI	0.74	19%

Table 7. Comparison of the *M2* protein of *P. putida* to other M proteins

Entry	TM-score	Sequence identity
M2_*P.putida* (compared with all the following entries)	-	-
M1.MboII	0.26	9%
M2.MboII	0.28	7%
M1.HphI	0.37	5%
M2.HphI	0.37	5%
M1.HgaI	0.37	5%
M2.HgaI	0.37	3%

M1 and *M2* showed low sequence and structural similarity to each other, suggesting they are evolutionarily unrelated and do not originate from a recent gene duplication event. Considering the TM-scores over 0.7 for the *M1* protein of *P. putida* with some other methylases, it can be considered structurally similar to methylase 1 from the *HphI* system of *Haemophilus parahaemolyticus* and also to both of the methylases from the *HgaI* system of *Haemophilus gallinarum*. Notably, *M1.HphI*, *M1.HgaI*, and *M2.HgaI* are all cytosine methyltransferases (Lubys et al., 1996; Sugisaki et al., 1991). Sequence alignment identified a conserved PPCQ motif in these three enzymes and the *M1* protein of *P. putida*, which is essential for catalysis (Pinarbaşi et al., 2001) (Figure 13). The cysteine residue within the motif forms a covalent enzyme-substrate intermediate, a necessary step for the subsequent methyl group transfer (Jurkowski et al., 2008).

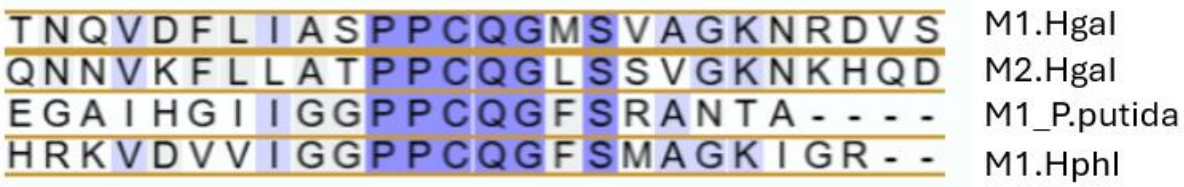

Figure *13*. **Sequence alignment of the conserved PPCQ catalytic motif in selected cytosine methyltransferases.** The alignment includes *M1* of *P. putida* alongside *M1.HphI* (Haemophilus parahaemolyticus), *M1.HgaI* and *M2.HgaI* (Haemophilus gallinarum). The conserved PPCQ motif, which forms the covalent enzyme-substrate intermediate required for methyl group transfer, is highlighted.

The *M2* gene, on the other hand, exhibited poor homology with all the methylases selected for comparison (Table 7). This agrees with the Foldseek results, which show that the top hits do not include any type II RM system methylases. The top hit in the PDB database is an adenine-specific DNA methyltransferase from *Geobacillus stearothermophilus* (Figure 14), with a TM-score of 0.52, only slightly above the limit of the general limit of 0.5.

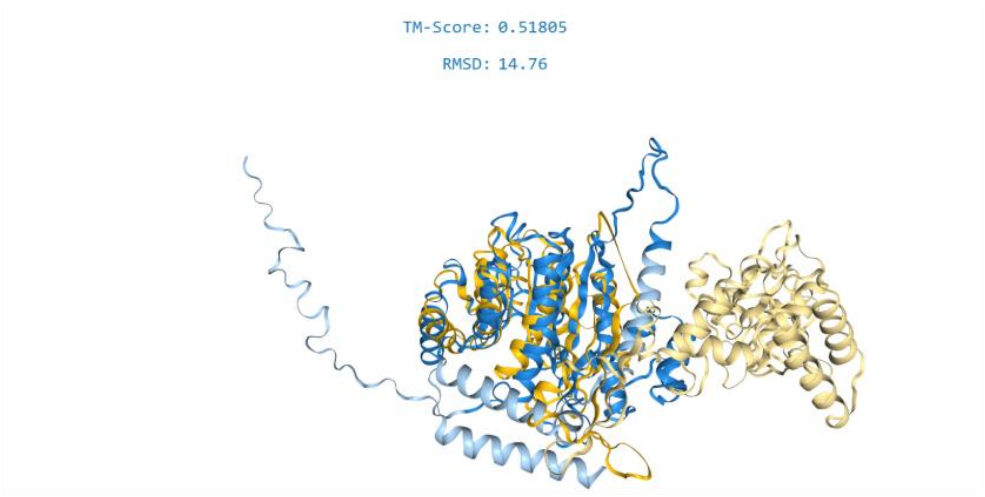

Figure *14*. **Structural alignment of the AlphaFold model of *P. putida M2* (*PP_5651*) with its top Foldseek hit from the PDB database: an adenine-specific DNA methyltransferase from Geobacillus stearothermophilus; PDBID: 7QW8 (TM-score 0.52)**. The structural overlap is shown in ribbon representation, where *M2* protein is shown in blue and the PDB hit protein is in yellow. Despite the low sequence identity, the shared structural fold suggests *M2* belongs to the adenine methyltransferase family.

Additionally, a midpoint-rooted phylogenetic tree was generated from the UniProt sequence alignment, which supports the relationships between compared genes (Figure 15). This tree demonstrates that *M1* of *P. putida* is most related to *M1.HphI*, while *M2* of *P. putida* remains entirely unrelated to the rest of the analyzed genes.

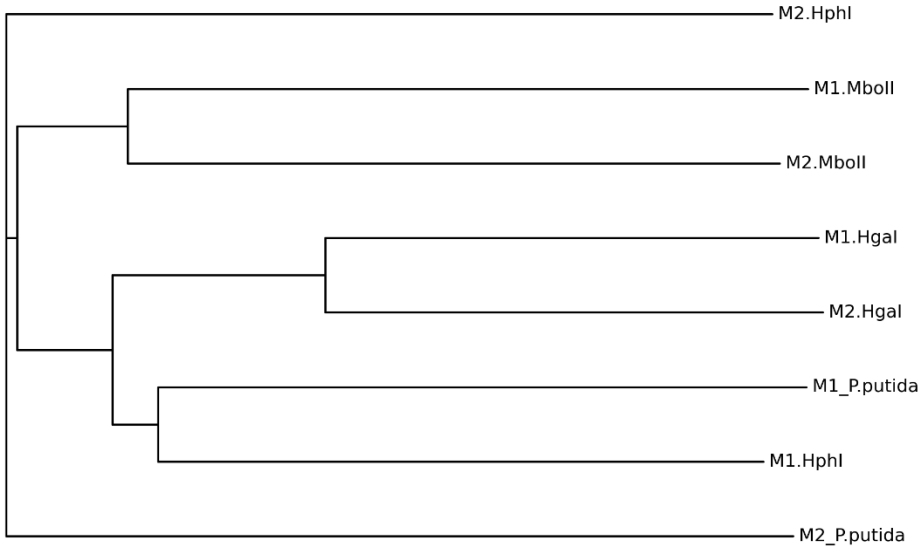

Figure *15*. **Midpoint-rooted phylogenetic tree generated from UniProt sequence alignment**. The aligned sequences include *P. putida M1* and *M2* methylases together with methylase pairs from selected type IIA(S) RM systems: MboII (*Moraxella bovis*), HphI (*Haemophilus parahaemolyticus*), and HgaI (*Haemophilus gallinarum*). Branch lengths reflect sequence divergence. The tree shows that *M1* of *P. putida* clusters with *M1.HphI* and is related to the HgaI methylases, while *M2* of *P. putida* forms a separate, distant branch.

3.3 DISCUSSION

3.3.1 Functional insights of the RMII system in *Pseudomonas putida* PaW85

Genomic analysis of the restriction-modification (RM) locus in *Pseudomonas putida* PaW85 revealed its non-canonical architecture. The system consists of a putative restrictase (*PP_3988*, *R*) and two distinct methylases (*PP_3989*, *M1* and *PP_5651*, *M2*). Growth curve analysis of mutant strains clearly showed that the restrictase becomes toxic to *P. putida* only when both methylases are deleted. This confirmed that *M1* and *M2* are each capable of providing genomic protection independently. The presence of two functional methylase genes suggests redundancy. However, when the recognition sequence is asymmetric, each strand requires a separate enzyme for methylation. If only a single methylase is present, a fully unmethylated site is produced after DNA replication or repair when the strand lacking the corresponding methylase is used as a template. This site could then be targeted by the restrictase until the methylase restores its modification. Therefore, two strand-specific methylases are common in type IIA systems (Loenen et al., 2014; Pingoud & Jeltsch, 2001; Roberts et al., 2003). For example, in the MboII system of *Moraxella bovis*, which recognizes the asymmetric sequence 5′-GAAGA-3′, M1.MboII methylates the top-strand adenine and M2.MboII modifies the bottom-strand cytosine (Furmanek-Blaszk et al., 2009). The same principle applies to the HphI system of *Haemophilus parahaemolyticus* and the HgaI system of *Haemophilus gallinarum* (Lubys et al., 1996; Sugisaki et al., 1991). In the case of *P. putida,* the strains retaining only one methylase (ΔRΔM1 or ΔRΔM2) grew similarly to the wild type when expression of the restrictase was induced. However, assuming the RMII system in *P. putida* recognizes an asymmetric sequence, the growth of these strains is expected to be slightly impaired. Nevertheless, the possibility that the RMII system has an asymmetric target site cannot be ruled out, as several factors may explain the result. Firstly, when bacteria are grown under non-stressful laboratory conditions in rich medium, most fully unmethylated sites are generated due to DNA replication. Therefore, the sites may be vulnerable for a shorter period of time, and even one active methylase may be sufficient to restore the methylation status of the target strand. Additionally, the R gene was expressed from the low-copy babyTOL plasmid, so restrictase levels may have been insufficient to inhibit the growth through self-digestion (Lewańczyk, unpublished).

3.3.2 Future directions

Even though *M1* and *M2* are structurally and evolutionarily unrelated, they perform the same function, which protects the genome from self-restriction. The parallels can be drawn with M1.HphI and M2.HphI genes in the HphI system, which are also structurally distinct but serve the same purpose to protect the asymmetric recognition site (Lubys et al., 1996). Therefore, these genes were likely functionally linked to a single system through horizontal gene transfer (Makarova et al., 2011).

Analysis of methylases alone is insufficient to determine which subtypes the RMII system in *P. putida* belongs to. Since the classification relies on the properties of the restriction enzyme, which include recognition sequence symmetry, cleavage position, cofactor requirements, and oligomeric state (Loenen et al., 2014; Pingoud et al., 2014; Pingoud & Jeltsch, 2001). Therefore, the R protein needs to be purified to experimentally determine these properties.

3.3.3 Conclusions

In this work, a functional antiphage defense system (RMII) was identified in *P. putida* PaW85, which reduces phage infection efficiency by up to 200-fold for phages from genus clusters G16 and G17. The RMII locus encodes a restriction enzyme (*R/PP_3988*) and two methylases (*M1/PP_3989* and *M2/PP_5651*). The expression of the restrictase seems to be toxic in *E. coli*, but it was demonstrated that each of the methylases is capable of protecting the *P. putida* genome from self-restriction. Despite their structural and evolutionary divergence, *M1* and *M2* carry out the same protective function, mirroring the HphI system. The RMII system subtype cannot be determined solely from methylase analysis; biochemical characterization of the *R* gene product, including its recognition sequence, cleavage pattern, and cofactor requirements, is required for definitive classification.

SUMMARY

Phages are a constant threat to bacterial survival and are a major concern in industrial fermentation, where contamination can significantly interfere with production. In response, bacteria have evolved a wide array of anti-phage defense systems, the most prevalent of which are Restriction-Modification (RM) systems. Despite the biotechnological importance of *Pseudomonas putida*, its anti-phage defense systems are not well-described. This thesis set out to identify functional defense systems in *P. putida* PaW85 and to characterize the non-canonical type II RM system (RMII) encoded in this bacterium.

Using an expanded CEPEST phage collection and comparative infection assays between *P. putida* deletion strains, the RMII system was validated as a functional anti-phage defense system. Genomic analysis of the RMII locus revealed a non-canonical architecture: in addition to the previously predicted restrictase (*R, PP_3988*) and methylase (*M1, PP_3989*), a third gene, *PP_5651*, encodes a second methylase (*M2*) in the same operon. To determine whether both methylases are functional, deletion strains were constructed lacking either or both methylase genes. Growth curve analysis while expressing the restrictase in these strains demonstrated that the restrictase is toxic only when both methylases are absent, establishing that M1 and M2 each independently protect the *P. putida* genome from self-digestion. Bioinformatic and structural analyses revealed that M1 is structurally similar to cytosine methyltransferases from well-characterized type IIA systems and shares a conserved PPCQ catalytic motif with them. M2, in contrast, is structurally and evolutionarily unrelated to M1 and to other canonical type II methylases, with its closest structural match being an adenine-specific methyltransferase. Overall, these results suggest that M1 and M2 were functionally coupled through horizontal gene transfer rather than gene duplication, forming a non-canonical system in *P. putida* with two methylases.

These findings present the first experimental characterization of the RMII system in *P. putida* PaW85 and reveal its non-canonical architecture. The confirmed defense activity and the functional independence of each methylase provide a foundation for future biochemical purification and characterization of the restrictase. More broadly, characterizing active defense systems in *P. putida* will facilitate more precise strain engineering for industrial applications.

REFERENCES

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., … Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, *630*(8016), 493–500. https://doi.org/10.1038/s41586-024-07487-w

Bayley, S. A., Duggleby, C. J., Worsey, M. J., Williams, P. A., Hardy, K. G., & Broda, P. (1977). Two modes of loss of the tol function from Pseudomonas putida mt-2. *Molecular and General Genetics MGG*, *154*(2), 203–204. https://doi.org/10.1007/BF00330838

Belda, E., van Heck, R. G. A., José Lopez-Sanchez, M., Cruveiller, S., Barbe, V., Fraser, C., Klenk, H.-P., Petersen, J., Morgat, A., Nikel, P. I., Vallenet, D., Rouy, Z., Sekowska, A., Martins dos Santos, V. A. P., de Lorenzo, V., Danchin, A., & Médigue, C. (2016). The revisited genome of Pseudomonas putida KT2440 enlightens its value as a robust metabolic chassis. *Environmental Microbiology*, *18*(10), 3403–3424. https://doi.org/10.1111/1462-2920.13230

Bernheim, A., & Sorek, R. (2020). The pan-immune system of bacteria: Antiviral defence as a community resource. *Nature Reviews Microbiology*, *18*(2), 113–119. https://doi.org/10.1038/s41579-019-0278-2

Bittrich, S., Segura, J., Duarte, J. M., Burley, S. K., & Rose, Y. (2024). RCSB protein Data Bank: Exploring protein 3D similarities via comprehensive structural alignments. *Bioinformatics*, *40*(6), btae370. https://doi.org/10.1093/bioinformatics/btae370

Borodovich, T., Shkoporov, A. N., Ross, R. P., & Hill, C. (2022). Phage-mediated horizontal gene transfer and its implications for the human gut microbiome. *Gastroenterology Report*, *10*, goac012. https://doi.org/10.1093/gastro/goac012

Brauer, A., Rosendahl, S., Kängsep, A., Lewańczyk, A. C., Rikberg, R., Hõrak, R., & Tamman, H. (2024). Isolation and characterization of a phage collection against Pseudomonas putida. *Environmental Microbiology*, *26*(6), e16671. https://doi.org/10.1111/1462-2920.16671

Callahan, S. J., Luyten, Y. A., Gupta, Y. K., Wilson, G. G., Roberts, R. J., Morgan, R. D., & Aggarwal, A. K. (2016). Structure of Type IIL Restriction-Modification Enzyme MmeI in Complex with DNA Has Implications for Engineering New Specificities. *PLOS Biology*, *14*(4), e1002442. https://doi.org/10.1371/journal.pbio.1002442

Chevallereau, A., Pons, B. J., van Houte, S., & Westra, E. R. (2022). Interactions between bacterial and phage communities in natural environments. *Nature Reviews Microbiology*, *20*(1), 49–62. https://doi.org/10.1038/s41579-021-00602-y

de Lorenzo, V., Pérez-Pantoja, D., & Nikel, P. I. (2024). Pseudomonas putida KT2440: The long journey of a soil-dweller to become a synthetic biology chassis. *Journal of Bacteriology*, *206*(7), e00136-24. https://doi.org/10.1128/jb.00136-24

Elsener, T. A., Cehovin, A., Yee, W.-X., Forrow, A., Palmer, S., & Tang, C. M. (2026). *Phase-variable restriction–modification systems dynamically regulate plasmid transmission in Neisseria gonorrhoeae* (p. 2026.01.10.698759). bioRxiv. https://doi.org/10.64898/2026.01.10.698759

Fuerst, T. R., Niles, E. G., Studier, F. W., & Moss, B. (1986). Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*, *83*(21), 8122–8126. https://doi.org/10.1073/pnas.83.21.8122

Furmanek-Blaszk, B., Boratynski, R., Zolcinska, N., & Sektas, M. (2009). M1.MboII and M2.MboII type IIS methyltransferases: Different specificities, the same target. *Microbiology*, *155*(4), 1111–1121. https://doi.org/10.1099/mic.0.025023-0

Georjon, H., & Bernheim, A. (2023). The highly diverse antiphage defence systems of bacteria. *Nature Reviews Microbiology*, *21*(10), 686–700. https://doi.org/10.1038/s41579-023-00934-x

Göse, M., Magill, E. E., Hughes-Games, A., Shaw, S. J., Diffin, F. M., Rawson, T., Nagy, Z., Seidel, R., & Szczelkun, M. D. (2024). Short-range translocation by a restriction enzyme motor triggers diffusion along DNA. *Nature Chemical Biology*, *20*(6), 689–698. https://doi.org/10.1038/s41589-023-01504-1

Grafakou, A., Mosterd, C., de Waal, P. P., van Rijswijck, I. M. H., van Peij, N. N. M. E., Mahony, J., & van Sinderen, D. (2024). Functional and practical insights into three lactococcal antiphage systems. *Applied and Environmental Microbiology*, *90*(9), e01120-24. https://doi.org/10.1128/aem.01120-24

Güemes, A. G. C., Youle, M., Cantú, V. A., Felts, B., Nulton, J., & Rohwer, F. (2016). Viruses as Winners in the Game of Life. *Annual Review of Virology*, *3*(Volume 3, 2016), 197–214. https://doi.org/10.1146/annurev-virology-100114-054952

Hille, F., Richter, H., Wong, S. P., Bratovič, M., Ressel, S., & Charpentier, E. (2018). The Biology of CRISPR-Cas: Backward and Forward. *Cell*, *172*(6), 1239–1259. https://doi.org/10.1016/j.cell.2017.11.032

Jędrzejewski, M., Szeleszczuk, Ł., & Pisklak, D. M. (2025). Mechanistic Insights into SAM-Dependent Methyltransferases: A Review of Computational Approaches. *International Journal of Molecular Sciences*, *26*(18), 9204. https://doi.org/10.3390/ijms26189204

Jurkowski, T. P., Meusburger, M., Phalke, S., Helm, M., Nellen, W., Reuter, G., & Jeltsch, A. (2008). Human DNMT2 methylates tRNAAsp molecules using a DNA methyltransferase-like catalytic mechanism. *RNA*, *14*(8), 1663–1670. https://doi.org/10.1261/rna.970408

Kommireddy, V., & Nagaraja, V. (2013). Diverse Functions of Restriction-Modification Systems in Addition to Cellular Defense. *Microbiology and Molecular Biology Reviews : MMBR*, *77*, 53–72. https://doi.org/10.1128/MMBR.00044-12

Kusano, K., Naito, T., Handa, N., & Kobayashi, I. (1995). Restriction-modification systems as genomic parasites in competition for specific sequences. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(24), 11095–11099. https://doi.org/10.1073/pnas.92.24.11095

Leavitt, J. C., Woodbury, B. M., Gilcrease, E. B., Bridges, C. M., Teschke, C. M., & Casjens, S. R. (2024). Bacteriophage P22 SieA-mediated superinfection exclusion. *mBio*, *15*(2), e02169-23. https://doi.org/10.1128/mbio.02169-23

Loenen, W. A. M., Dryden, D. T. F., Raleigh, E. A., Wilson, G. G., & Murray, N. E. (2014). Highlights of the DNA cutters: A short history of the restriction enzymes. *Nucleic Acids Research*, *42*(1), 3–19. https://doi.org/10.1093/nar/gkt990

Loenen, W. A. M., & Raleigh, E. A. (2014). The other face of restriction: Modification-dependent enzymes. *Nucleic Acids Research*, *42*(1), 56–69. https://doi.org/10.1093/nar/gkt747

Lozupone, C., & Knight, R. (2007). Global pattern in bacterial diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 11436–11440. https://doi.org/10.1073/pnas.0611525104

Lubys, A., Lubienė, J., Kulakauskas, S., Stankevicius, K., Timinskas, A., & Janulaitis, A. (1996). Cloning and Analysis of the Genes Encoding the Type IIS Restriction-Modification System Hph I from Haemophilus Parahaemolyticus. *Nucleic Acids Research*, *24*(14), 2760–2766. https://doi.org/10.1093/nar/24.14.2760

Makarova, K. S., Wolf, Y. I., Snir, S., & Koonin, E. V. (2011). Defense Islands in Bacterial and Archaeal Genomes and Prediction of Novel Defense Systems. *Journal of Bacteriology*, *193*(21), 6039–6056. https://doi.org/10.1128/jb.05535-11

Manning, A. J., & Kuehn, M. J. (2011). Contribution of bacterial outer membrane vesicles to innate bacterial defense. *BMC Microbiology*, *11*(1), 258. https://doi.org/10.1186/1471-2180-11-258

Martínez-García, E., & de Lorenzo, V. (2011). Engineering multiple genomic deletions in Gram-negative bacteria: Analysis of the multi-resistant antibiotic profile of Pseudomonas putida KT2440. *Environmental Microbiology*, *13*(10), 2702–2716. https://doi.org/10.1111/j.1462-2920.2011.02538.x

Messling, J.-E., & Williams, A. B. (2016). Chapter 4—Genome Instability in Bacteria and Archaea: Strategies for Maintaining Genome Stability. In I. Kovalchuk & O. Kovalchuk (Eds.), *Genome Stability* (pp. 51–67). Academic Press. https://doi.org/10.1016/B978-0-12-803309-8.00004-5

Murray, N. E. (2000). Type I Restriction Systems: Sophisticated Molecular Machines (a Legacy of Bertani and Weigle). *Microbiology and Molecular Biology Reviews*, *64*(2), 412–434. https://doi.org/10.1128/mmbr.64.2.412-434.2000

Nabergoj, D., Modic, P., & Podgornik, A. (2017). Effect of bacterial growth rate on bacteriophage population growth rate. *MicrobiologyOpen*, *7*(2), e00558. https://doi.org/10.1002/mbo3.558

Nidhanapati K, R., Bheemanaik, S., & Rao, D. (2012). Mechanistic insights into type III restriction enzymes. *Frontiers in Bioscience : A Journal and Virtual Library*, *17*, 1094–1107. https://doi.org/10.2741/3975

Nikel, P. I., & de Lorenzo, V. (2018). *Pseudomonas putida* as a functional *chassis* for industrial biocatalysis: From native biochemistry to *trans*-metabolism. *Metabolic Engineering, Metabolic Engineering Host Organism Special Issue*, *50*, 142–155. https://doi.org/10.1016/j.ymben.2018.05.005

Pinarbaşi, H., PINARBAŞI, E., & HORNBY, D. (2001). Substitution of the Conserved Cysteine With Glycine (Cys82Gly) of Agmenellum quadruplicatum Methylase Aqul (M. Aqul) is not Cytotoxic to E. coli. *Turkish Journal of Biology*, *25*(2), 177–184. https://doi.org/-

Pingoud, A., & Jeltsch, A. (2001). Structure and function of type II restriction endonucleases. *Nucleic Acids Research*, *29*(18), 3705–3727. https://doi.org/10.1093/nar/29.18.3705

Pingoud, A., Wilson, G. G., & Wende, W. (2014). Type II restriction endonucleases—A historical perspective and more. *Nucleic Acids Research*, *42*(12), 7489–7527. https://doi.org/10.1093/nar/gku447

Platt, R., Drescher, C., Park, S.-K., & Phillips, G. J. (2000). Genetic System for Reversible Integration of DNA Constructs and *lacZ* Gene Fusions into the *Escherichia coli* Chromosome. *Plasmid*, *43*(1), 12–23. https://doi.org/10.1006/plas.1999.1433

Roberts, R. J., Belfort, M., Bestor, T., Bhagwat, A. S., Bickle, T. A., Bitinaite, J., Blumenthal, R. M., Degtyarev, S. Kh., Dryden, D. T. F., Dybvig, K., Firman, K., Gromova, E. S., Gumport, R. I., Halford, S. E., Hattman, S., Heitman, J., Hornby, D. P., Janulaitis, A., Jeltsch, A., … Xu, S. (2003). A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Research*, *31*(7), 1805–1812. https://doi.org/10.1093/nar/gkg274

Saha, C. K., Sanches Pires, R., Brolin, H., Delannoy, M., & Atkinson, G. C. (2021). FlaGs and webFlaGs: Discovering novel biology through the analysis of gene neighbourhood conservation. *Bioinformatics*, *37*(9), 1312–1314. https://doi.org/10.1093/bioinformatics/btaa788

Sharma, S., Chatterjee, S., Datta, S., Prasad, R., Dubey, D., Prasad, R. K., & Vairale, M. G. (2017). Bacteriophages and its applications: An overview. *Folia Microbiologica*, *62*(1), 17–55. https://doi.org/10.1007/s12223-016-0471-x

Srinivasiah, S., Bhavsar, J., Thapar, K., Liles, M., Schoenfeld, T., & Wommack, K. E. (2008). Phages across the biosphere: Contrasts of viruses in soil and aquatic environments. *Research in Microbiology, Exploring the Prokaryotic Virosphere*, *159*(5), 349–357. https://doi.org/10.1016/j.resmic.2008.04.010

Sugisaki, H., Yamamoto, K., & Takanami, M. (1991). The HgaI restriction-modification system contains two cytosine methylase genes responsible for modification of different DNA strands. *Journal of Biological Chemistry*, *266*(21), 13952–13957. https://doi.org/10.1016/S0021-9258(18)92795-X

Suttle, C. A. (2007). Marine viruses—Major players in the global ecosystem. *Nature Reviews Microbiology*, *5*(10), 801–812. https://doi.org/10.1038/nrmicro1750

Tesson, F., Planel, R., Egorov, A., Georjon, H., Vaysset, H., Brancotte, B., Néron, B., Mordret, E., Atkinson, G., Bernheim, A., & Cury, J. (2024). *A Comprehensive Resource for Exploring*

Antiphage Defense: DefenseFinder Webservice, Wiki and Databases (p. 2024.01.25.577194). bioRxiv. https://doi.org/10.1101/2024.01.25.577194

The UniProt Consortium. (2025). UniProt: The Universal Protein Knowledgebase in 2025. *Nucleic Acids Research*, *53*(D1), D609–D617. https://doi.org/10.1093/nar/gkae1010

van Kempen, M., Kim, S. S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C. L. M., Söding, J., & Steinegger, M. (2024). Fast and accurate protein structure search with Foldseek. *Nature Biotechnology*, *42*(2), 243–246. https://doi.org/10.1038/s41587-023-01773-0

Volke, D. C., Friis, L., Wirth, N. T., Turlin, J., & Nikel, P. I. (2020). Synthetic control of plasmid replication enables target- and self-curing of vectors and expedites genome engineering of *Pseudomonas putida*. *Metabolic Engineering Communications*, *10*, e00126. https://doi.org/10.1016/j.mec.2020.e00126

Volke, D. C., Wirth, N. T., & Nikel, P. (2021). Rapid Genome Engineering of Pseudomonas Assisted by Fluorescent Markers and Tractable Curing of Plasmids. *BIO-PROTOCOL*, *11*(4). https://doi.org/10.21769/BioProtoc.3917

Waldor, M. K., & Mekalanos, J. J. (1996). Lysogenic Conversion by a Filamentous Phage Encoding Cholera Toxin. *Science*, *272*(5270), 1910–1914. https://doi.org/10.1126/science.272.5270.1910

Winsor, G. L., Griffiths, E. J., Lo, R., Dhillon, B. K., Shay, J. A., & Brinkman, F. S. L. (2016). Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. *Nucleic Acids Research*, *44*(D1), D646-653. https://doi.org/10.1093/nar/gkv1227

Xu, Z., Huang, Y., Dong, Y., An, X., Tong, Y., & Li, M. (2025). Engineering broad-spectrum phage-resistant Escherichia coli via adaptive and programmable defense strategies. *Applied and Environmental Microbiology*, *91*(11), e01596-25. https://doi.org/10.1128/aem.01596-25

Zaleski, P., Wojciechowski, M., & Piekarowicz, A. (2005). The role of Dam methylation in phase variation of Haemophilus influenzae genes involved in defence against phage infection. *Microbiology*, *151*(10), 3361–3369. https://doi.org/10.1099/mic.0.28184-0

NON-EXCLUSIVE LICENCE TO REPRODUCE THE THESIS AND MAKE THE THESIS PUBLIC

I, Oleksandr Syritka ,
(author´s name)

grant the University of Tartu a free permit (non-exclusive licence) to

reproduce, for the purpose of preservation, including for adding to the digital archives of the University of Tartu until the expiry of the term of copyright, my thesis

"Describing the non-canonical Restriction-Modification type II system (RMII) in Pseudomonas putida PaW85" ,
(title of thesis)

supervised by Andres Ainelo and Hedvig Tamman ;
(supervisor´s name)

grant the University of Tartu a permit to make the thesis specified in point 1 available to the public via the web environment of the University of Tartu, including via the digital archives, under the Creative Commons licence CC BY NC ND 4.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright;

am aware of the fact that the author retains the rights specified in points 1 and 2;

confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Oleksandr Syritka

20/05/2026