

DISARMing viral invaders of bacteria

Aparicio Maldonado, C.

DOI

[10.4233/uuid:ac1af560-c551-46ef-94b8-8c2dbca21176](https://doi.org/10.4233/uuid:ac1af560-c551-46ef-94b8-8c2dbca21176)

Publication date

2023

Document Version

Final published version

Citation (APA)

Aparicio Maldonado, C. (2023). *DISARMing viral invaders of bacteria*. [Dissertation (TU Delft), Delft University of Technology]. <https://doi.org/10.4233/uuid:ac1af560-c551-46ef-94b8-8c2dbca21176>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

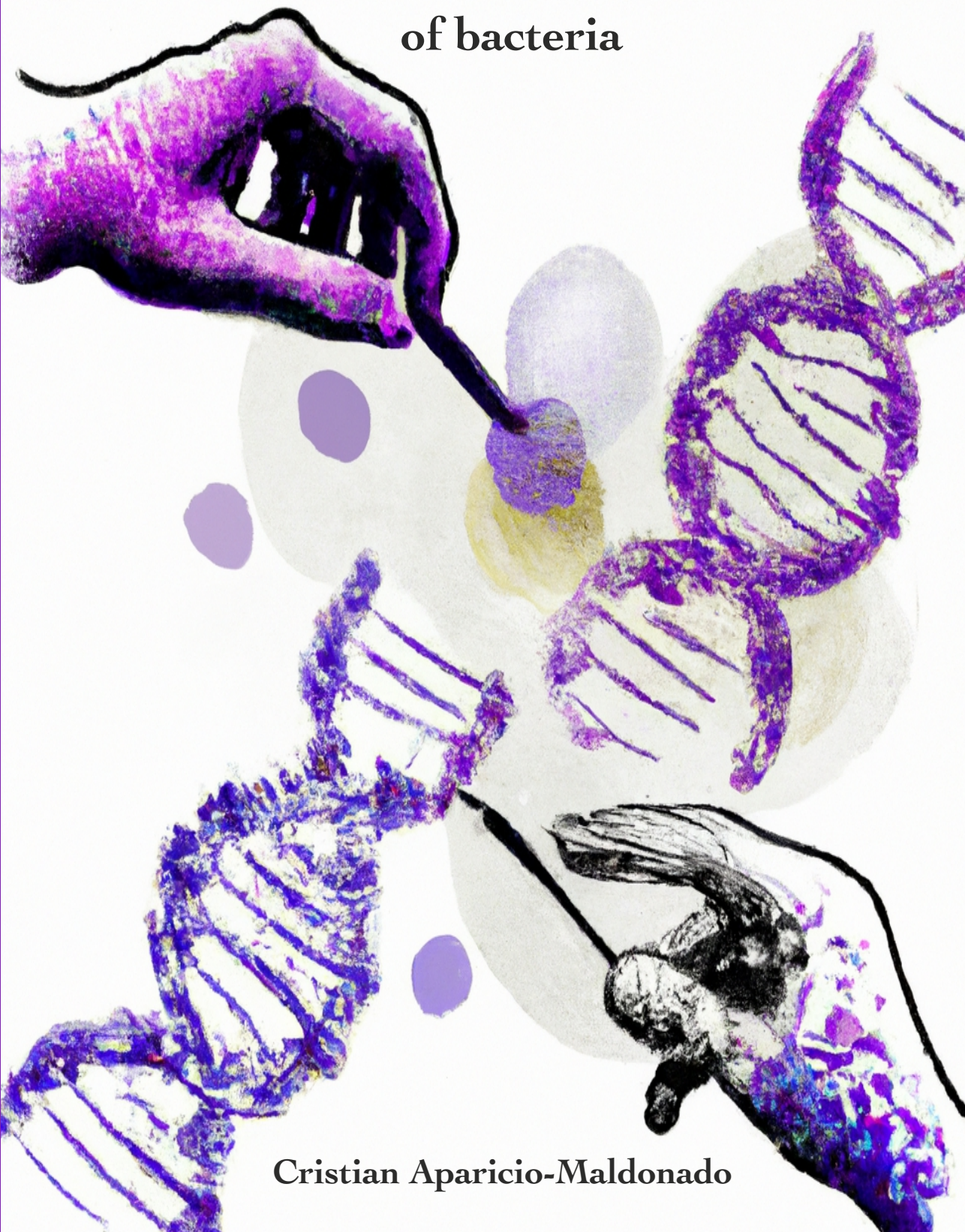
Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

DISARMin^g viral invaders of bacteria



Cristian Aparicio-Maldonado

DISARMing viral invaders of bacteria

Dissertation

for the purpose of obtaining the degree of doctor

at Delft University of Technology

by the authority of the Rector Magnificus, Prof.dr.ir. T.H.J.J. van der Hagen

chair of the Board for Doctorates

to be defended publicly on

Friday 27, January 2023 at 12:30 o'clock

by

Cristian APARICIO-MALDONADO

Master en Virologia, Universidad Complutense de Madrid, Spain

born in Almagro, Spain.

This dissertation has been approved by the promotor.

Composition of the doctoral committee:

Rector Magnificus	chairperson
Dr.ir. S.J.J. Brouns	Delft University of Technology, promotor
Dr.ir. F.L. Nobrega	University of Southampton, United Kingdom, copromoter

Independent members:

Prof.dr. M. Clokie	University of Leicester, United Kingdom
Dr. R.H.J. Staals	University of Wageningen
Prof.dr. A. Briegel	University of Leiden
Prof.dr. C. Joo	Delft University of Technology

Reserve member:

Prof.dr. C. Dekker	Delft University of Technology
--------------------	--------------------------------



Bionanoscience Department
Think big about life at the smallest scale

© 2023 Cristian Aparicio-Maldonado

Casimir PhD series 2022-39

ISBN: 978-90-8593-548-3

An electronic version of this dissertation is

available at: <http://repository.tudelft.nl>



Para mi Tito Agustín,
te fuiste demasiado rápido,
pero siempre perdurarás en nuestra memoria.

Y para Mamá y Papá,
por estar siempre ahí, pase lo que pase.
Hacéis todo posible.

Table of Contents

Summary		7
Samenvatting (Dutch)		9
Resumen (Español)		12
Chapter 1	General Introduction	15
Chapter 2	An overview of bacterial defense mechanisms	39
Chapter 3	Characterization of Class 1 DISARM as bacterial defense system	87
Chapter 4	Structural characterization of Class 1 DISARM mechanism	123
Chapter 5	Short pAgos trigger cell death upon detection of invading DNA	167
Chapter 6	Phage Nami and Shinka: isolation, characterization, and sequencing	233
Appendix	Acknowledgements – About the author – List of publications – and others	251

Summary

The research performed in this thesis focuses on the understanding on the interactions of bacteria and phages occurring via anti-phage defense system mechanisms. This involved a set of literature research and experimental studies that provide an overview of the diverse defense systems described at the time of submission.

The thesis begins with **Chapter 1 (General introduction)**, which provides context to understanding the content appearing on subsequent chapters. This chapter discusses the basics of bacteriophage biology, provides an overview of defense systems and addresses the beginnings of phage therapy.

Chapter 2 (An overview of bacterial defense mechanisms) reviews the known anti-phage resistance mechanisms. It starts with a brief description of the phage infection cycle, highlighting those steps important to understand the resistance mechanisms. Subsequently, the mechanisms of phage resistance are described, including those happening in the bacterial surface and those occurring inside the cell once the infection has begun. The latter are divided in defense systems evolved by the host and those derived from phages and later acquired by the host. This chapters also describes some mechanisms that phages have evolved to counteract the host defense systems. Finally, it explores the importance and role of these resistance mechanisms in a clinical context, with the perspective of the therapeutical use of phages.

The study of anti-phage defense systems starts with **Chapter 3 (Characterization of Class 1 DISARM as bacterial defense system)**, a characterization of the novel defense system DISARM. This study focusses on the Class 1 DISARM system from *Serratia* sp. SCBI, and uses different experimental approaches to understand the system at different levels. The defense activity of this system is studied for a broad range of virulent phages of the *Caudovirales* order as well as for the chronic-infecting phage M13. In addition, it shows the anti-conjugation activity of the system, a feature not commonly investigated for

other defense systems. Altogether, this chapter informs about the role and essentiality of the DISARM components.

Chapter 4 (Structural characterization of Class 1 DISARM mechanism) can be interpreted as the follow-up study of the Class 1 DISARM system, in which the molecular mechanism of the main essential components of DISARM (DrmA and DrmB) is explored through structural and biochemical studies. These proteins form a 220 kDa complex that is responsible for the DISARM defense activity. For this, the complex is able to recognize foreign molecules by detecting a methylation pattern in DNA molecules presenting a single-stranded 5' end, common in actively-replicating invading DNA.

Chapter 5 (Short prokaryotic Argonaute systems trigger cell death upon detection of invading DNA) studies the SPARTA prokaryotic Argonaute system using structural, biochemical and phage biology assays. This work expanded the current understanding of the NAD(P)ase activity associated to the RNA-guided DNA-binding of the heterodimeric complex formed by both proteins of the system. The SPARTA system targets multi-copy DNA invaders, mainly those of plasmid nature. The anti-phage activity was not as strong as for other pAgo systems described previously.

The thesis finalizes with **Chapter 6 (Phage Nami and Shinka: isolation, characterization, and sequencing)**, a study in the field of phage biology. Two novel Tequatroviruses were isolated against the clinical uropathogenic *E. coli* strain R10256. Phages were coined as Nami and Shinka based on the features they showed. Nami (“wave” in Japanese) showed a wavy infecting profile of bacterial cultures. Shinka (“evolution” in Japanese) was used in evolutionary studies of phage resistance and its overcoming. Both phages were characterized by the study of their infection cycle and morphology, as well as by genome sequencing. In addition, these phages were used, together with other phages, to study the different defense systems explored in the previous chapters of this thesis. Finally, a phage resistance study was performed with Nami and Shinka to explore their potential in clinical use and broaden our knowledge in this field.

Samenvatting

Het onderzoek dat in dit proefschrift wordt uitgevoerd, richt zich op het begrip van de interacties van bacteriën en bacteriofagen die plaatsvinden via antifaag afweermechanismen. Dit omvat een reeks literatuuronderzoeken en experimentele studies die een overzicht geven van de diverse verdedigingssystemen die op het moment van inleveren beschreven zijn.

Het proefschrift begint met **Hoofdstuk 1 (General introduction)**, dat inhoudelijke informatie geeft voor het begrijpen van volgende hoofdstukken. Het hoofdstuk bespreekt de basisprincipes van bacteriofaagbiologie, geeft een overzicht van afweersystemen en behandelt het begin van faagtherapie.

Hoofdstuk 2 (An overview of bacterial defense mechanisms) bevat een gedetailleerde beschrijving van de bekende anti-faag resistentie mechanismen. Het begint met een korte beschrijving van de faaginfectiecyclus, waarbij de stappen worden benadrukt die belangrijk zijn om de resistentiemechanismen te begrijpen. Vervolgens worden de mechanismen van faagresistentie beschreven, inclusief de mechanismen die op het bacteriële oppervlak en binnen in de cel optreden zodra de infectie begonnen is. Deze twee mechanismen zijn onderverdeeld in afweersystemen die zijn ontwikkeld door de gastheer, en die zijn afgeleid van fagen en later door de gastheer zijn verworven. Verder beschrijft dit hoofdstuk ook mechanismen die fagen hebben ontwikkeld om afweersystemen van de gastheer te onderdrukken. Ten slotte wordt het belang en de rol van deze resistentiemechanismen in een klinische context onderzocht, met het perspectief op het therapeutisch gebruik van fagen.

De studie van anti-faag afweersystemen begint met **Hoofdstuk 3 (Characterization of Class 1 DISARM as bacterial defense system)**, waarin de karakterisering van het nieuwe afweersysteem DISARM beschreven wordt. Deze studie richt zich op het Class 1-DISARM systeem van *Serratia* sp. SCBI, waarbij verschillende experimentele benaderingen gebruikt worden om de functie van het systeem op verschillende niveaus te begrijpen. De anti-faag activiteit van dit systeem wordt bestudeerd op een breed scala van virulente fagen behorende tot

de Caudovirales-order, en voor de chronisch infecterende faag M13. Bovendien toont het de anti-conjugatie-activiteit van het systeem, een functie die niet vaak wordt onderzocht voor andere bacteriële verdedigingssystemen. In het geheel, geeft dit hoofdstuk informatie over de rol en essentie van de DISARM-componenten.

Hoofdstuk 4 (Structural characterization of Class 1 DISARM mechanism)

kan worden geïnterpreteerd als de vervolgstudie van het Class 1-DISARM systeem, waarin het moleculaire mechanisme van de essentiële componenten van DISARM (DrmA en DrmB) wordt onderzocht door middel van structurele en biochemische studies. Samen vormen deze eiwitten een 220 kDa-complex dat verantwoordelijk is voor de DISARM-afweeractiviteit. Door middel van een methylerings patroon kan het complex niet-eigen DNA moleculen herkennen met een enkelstrengs 5'-uiteinde, gebruikelijk in actief replicerend binnenvallend DNA.

Hoofdstuk 5 (Short prokaryotic Argonaute systems trigger cell death upon detection of invading DNA) bestudeert het SPARTA prokaryotische argonaute-systeem met behulp van structurele, biochemische en faagbiologische experimenten. Dit werk breidt het huidige begrip uit van de NAD(P)ase-activiteit die geassocieerd wordt met de RNA-geleide DNA-binding van het heterodimere complex dat gevormd wordt door beide eiwitten van het systeem. Het SPARTA-systeem richt zich op DNA-indringers met meerdere kopieën, voornamelijk die van plasmide-aard. De vertoonde anti-faagactiviteit was niet even sterk als beschreven voor andere pAgo-systemen.

Het proefschrift wordt afgesloten met **Hoofdstuk 6 (Phage Nami and Shinka: isolation, characterization, and sequencing)**. Er werden twee nieuwe Tequatrovirussen geïsoleerd tegen de klinische uropathogene *E. coli*-stam R10256. Fagen werden benoemd als Nami en Shinka op basis van de kenmerken die ze vertoonden. Nami ("golf" in het Japans) vertoonde een golvend infecterend profiel van bacterieculturen. Shinka ("evolutie" in het Japans) werd gebruikt in evolutionaire studies van faagresistentie en het overwinnen ervan. Beide fagen werden gekarakteriseerd door de studie van hun infectiecyclus en morfologie,

evenals door genoom sequencing. Verder werden deze fagen, samen met andere fagen, gebruikt om de verschillende verdedigingssystemen te bestuderen die in de bovengenoemde hoofdstukken van dit proefschrift zijn onderzocht. Ten slotte werd een faagresistentiestudie uitgevoerd met Nami en Shinka om hun potentieel in klinisch gebruik te onderzoeken en onze kennis op dit gebied te verbreden.

Resumen

El trabajo de investigación de esta tesis se centra en entender las interacciones entre bacterias y los virus que les infectan (bacteriófagos o fagos), principalmente aquellas interacciones a través de sistema de defensa anti-fagos. Para ello, la tesis incluye un extensa revisión de literatura así como estudios experimentales, que, en conjunto, resultaron en una visión general de los sistema de defensa descritos en el momento de terminar la tesis.

Esta tesis comienza con el **Capítulo 1 (General introduction)**, el cuál aporta el contexto necesario para entender los capítulos subsecuentes. Aquí se exponen los aspectos básicos de la biología de fagos, un resumen de los sistemas de defensa y los inicios de la terapia de fagos como tratamiento de infecciones bacterianas.

El **Capítulo 2 (An overview of bacterial defense mechanisms)** consiste en una revisión de los sistemas resistencia en bacterias contra fagos conocidos previamente. Comienza con una descripción breve del ciclo de infección de los fagos, destacando aquellos aspectos de más importancia para entender los mecanismos de resistencia. A continuación, se describen estos mecanismos de resistencia, incluyendo aquellos que ocurren en la superficie bacteriana y aquellos que tienen lugar dentro de la célula una vez que la infección ha comenzado. Estos últimos se dividen en aquellos sistemas evolucionados por el hospedador bacteriano, y aquellos derivados de bacteriófagos que se incorporaron con posterioridad al hospedador. En este capítulo también se describen ciertos mecanismos que los bacteriófagos han desarrollado para contratacar dichos sistemas de defensa de la bacteria hospedadora. Finalmente, se explora la importancia de estos sistemas de defensa en un contexto clínico, con la perspectiva del uso terapéutico de los bacteriófagos para tratar infecciones.

Específicamente, el estudio de sistemas de defensa anti-fago comienza en el **Capítulo 3 (Characterization of Class 1 DISARM as bacterial defense system)**, con una caracterización del novedoso sistema de defensa DISARM. Este estudio se centra en el sistema de defensa DISARM clase 1 de la bacteria *Serratia* sp.

SCBI, para el que se usan diferentes técnicas experimentales . La actividad defensiva de este sistema se estudia para un gran abanico de fagos virulentos del orden *Caudovirales*, así como para el fago M13 con un ciclo de infección crónico. Asimismo, se estudia la actividad anti-conjugativa de este sistema de defensa; una característica no investigada corrientemente en otros estudios de sistemas de defensa. Todo esto conlleva a entender los componentes esenciales de DISARM, así como su funcionalidad.

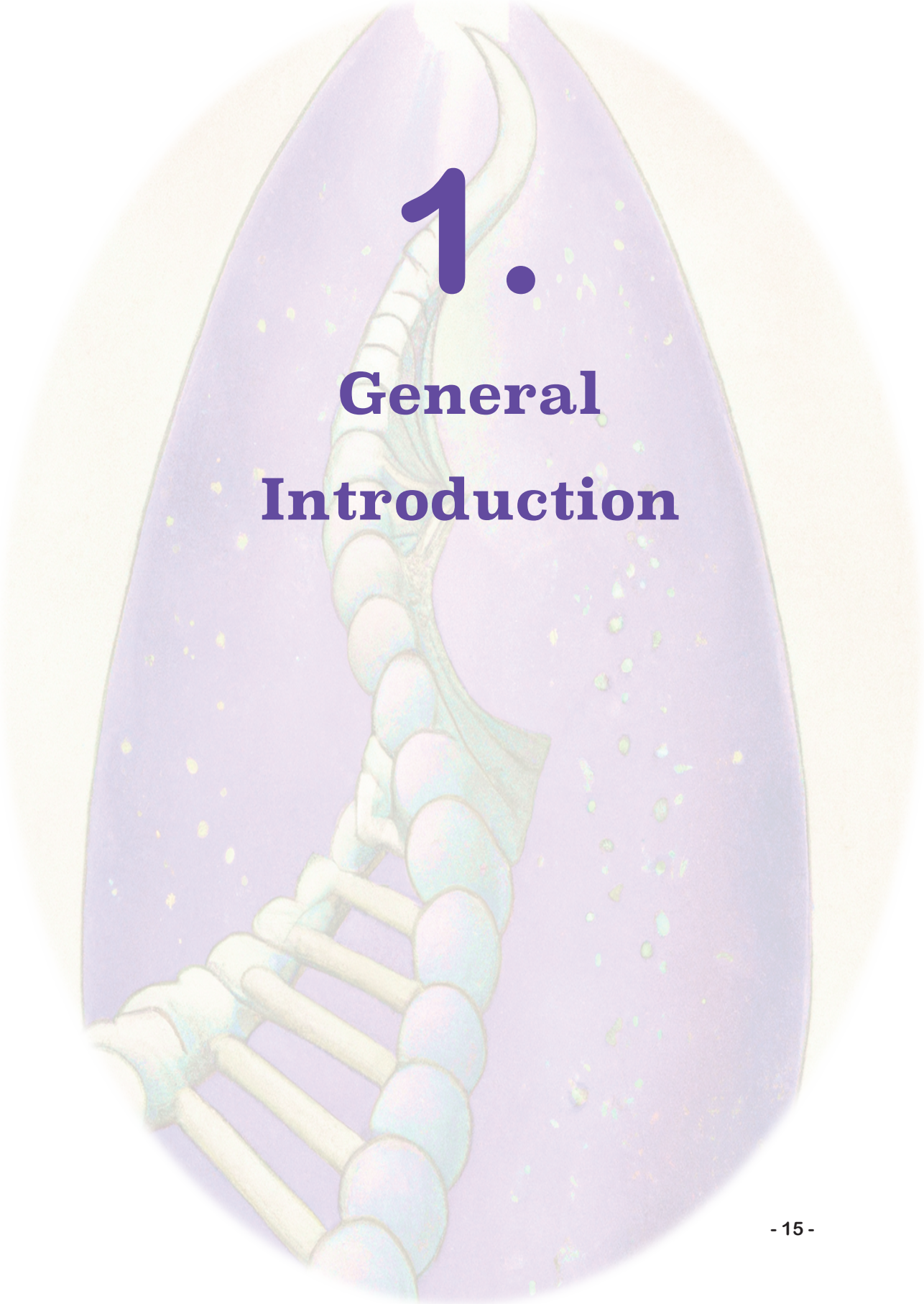
El **Capítulo 4 (Structural characterization of Class 1 DISARM mechanism)** puede interpretarse como una continuación en el estudio del sistema DISARM, centrándose en entender el mecanismo molecular de los principales componentes efectores del sistema (DrmA y DrmB) a través de estudios bioquímicos y estructurales. Dichos componentes forman un complejo proteínico de 220 KDa, el cual es responsable de la actividad defensiva de DISARM. Para ello, el complejo reconoce moléculas extrañas a través de detectar el patrón de metilación en el ADN que estas presentan, identificando y escaneando en primer lugar un extremo monocatenario 5' que se encuentra comúnmente en aquellos ADN invasores que replican activamente.

En el **Capítulo 5 (Short prokaryotic Argonaute systems trigger cell death upon detection of invading DNA)** se estudia el Sistema de defensa SPARTA, un tipo de sistema Argonauta procariota. Para ello se usan análisis estructurales, bioquímicos y de biología de fagos. Este estudio amplía el conocimiento actual de la actividad NAD(P)asa asociada con la unión a ADN guiada por ARN del sistema heterodimérico formado por ambas proteínas del sistema. El sistema SPARTA defiende contra invasores ADN de copias múltiple, principalmente aquellos de naturaleza plásmida. La actividad anti-fago no es tan potente como en otros sistemas Argonauta procariota descritos con anterioridad.

Esta tesis concluye con el **Capítulo 6 (Phage Nami and Shinka: isolation, characterization, and sequencing)**, un estudio centrado en el campo de la biología de fagos. Dos nuevos fagos *Tequatrovirus* fueron aislados para la cepa bacteriana R10256, una *E. coli* uropatógena de relevancia clínica. Los fagos fueron nombrados por el autor como Nami y Shinka, basando en las características

relevantes que manifestaron. Nami (de *ola* u *onda*, en japonés) mostró un perfil de infección ondulante en cultivos bacterianos. Shinka (de *evolución*, en japonés) fue empleado en estudios evolutivos de resistencia de fagos. Ambos fagos fueron caracterizados a través del estudio de su ciclo de infección, su morfología y su secuencia genómica. Además, se emplearon, junto con otros fagos, en el estudio de varios sistemas de defensa de los capítulos anteriores. Finalmente, se llevó a cabo un estudio de resistencia a fagos con Nami y Shinka para explorar su uso potencial en el contexto clínico y así ampliar el conocimiento en este campo.

Al final del libro, se puede encontrar un **Apéndice** con información sobre el autor, su *Curriculum Vitae*, una lista de las publicaciones en la fecha de entrega de la tesis, y los agradecimientos dedicados muchas de las personas que contribuyeron a poder finalizar este capítulo de la vida.



1.

General Introduction

Any encounter between biological entities might result in an interaction with a positive outcome or in a competitive encounter, broadly defined as mutualism or commensalism, respectively ¹⁻³. The former might benefit both parts, making them evolutionary advantageous compared to other organisms in that ecosystem. However, a competitive or negative interaction might result in the extinction of one or both parts, reinforcing the need of protection against the opponent. This last scenario is the reason for defense mechanisms to have been developed in any kingdom of life ⁴⁻⁶; such as poisons or spines by plants to chase herbivores away, shields or camouflage by animals to hide, or an immune system to fight small, infectious pathogens. Thus, the composition of an ecosystem is a result of all the interactions taking place ⁷.

In the early 20th century, among many scientific discoveries, an unexpected one ⁸ would begin a learning journey about how the astonishing microbiological world is no exception (see in **Section 1.4**). In these interactions, prokaryotic cells have developed several defense mechanisms to combat their main intracellular invaders, usually referred to as bacteriophages or phages ⁹. These are viral pathogens hijack the bacterial machinery and, in last instance, lyse it to achieve the progeny release. Probably these fierce interactions between bacteria and the phages is one of the earliest occurring in life. This explains the great amount and complexity of the molecular mechanisms through which they interact ¹⁰⁻¹³.

1.1. Bacteriophages. The strategy of viral attackers

Although *what exactly is a virus* remains to be solved, different approaches attempt to determine what entities should be included on this definition. The *virocell* concept was proposed to distinct the infected cell from the virion, which is the extracellular, potentially-infectious particle ¹⁴. In other model, the many features defining or not a virus lead to the concept of virosphere, proposed as a dynamic group of replicators that encode structural proteins and transmit among hosts ¹⁵. From all the possible classification and definitions, it is clear the great diversity which viruses might present.

Bacterial and archaeal viruses are no exception for this, favoring a huge variety interactions between invaders and hosts which is rather expanded by the diversity of the formers. Among all bacteriophages, there are plenty of families and genera in which they are classified (**Box 1**), each with specific features that influence and diversify the interactions with the bacterial host. They present numerous virion morphologies that relate with the diverse ways in which they externally interact with the potential host. But also, they contain different genetic materials, challenging the host with different replication strategies. In addition, the same virus can follow different replication strategies depending on the circumstances; for example some might follow different replication strategies depending on factors like the cell density ¹⁶ or number of infections events in the surroundings ¹⁷. Therefore, this diversity that phages present increase even further the possible interactions that might take place. To understand this, we will first describe the infection process and classification of the main phages used in this thesis.

1.1.1. Caudovirales replication cycle

Phages belonging to the *Caudovirales* order ¹⁸ are the most well-understood and the most widely used in clinical applications ¹⁹. Known as tailed phages, they all contain a dsDNA genome and a head-tail structure (**Fig. 1**). Until recently, 3 families (based on structural differences) were considered to comprise this order ²⁰: *Podoviridae* with a short tail; *Siphoviridae* with a long, flexible tail;

and *Myoviridae* with a long, contractile tail. This taxonomic classification was based on phage particle morphology. But recently, a new, extended classification based on genomic approaches was proposed (Box 1). However, even with the differences leading to this diverse classification, the replication cycle of *Caudovirales* have the main replication steps in common .

The receptor binding proteins (RBPs) are involved in first step of their replication cycle. In *Caudovirales*, they are localized in the tail spikes and tail fibers of the phage tail ²¹. RBPs interact with the host, recognizing specific receptors on the surface of bacteria, leading to a subsequent adsorption of the phage ²¹. Phage receptors are present in the bacterial surface, typically found on the cell wall, like peptides or polysaccharides, or in protruding structures, like flagella, pili or capsules ^{22,23}. The first interaction of RBPs with receptors occurs through a weak, reversible attachment, that might be followed by a stronger, irreversible binding to either the same or a secondary receptor ²². In most cases, the adsorption step by recognition of the receptors is highly specific, so this is generally the most limiting factor for the phage host range ²⁴.

After the phage adsorbs to its precise receptor, it triggers the genetic material ejection into the bacterial cytoplasm. Although this ejection mechanism is complex and not fully understood for many phage types ²⁵, the *Caudovirales* ejection mechanism is the best well-known. The RBP-receptor interaction triggers conformational changes in the virion that result in the opening of a channel that enables DNA release from the over-pressured capsid ^{26,27}. In some cases, the channel is created by the ejection of the spike or the tape measure protein, depending on each case, that creates a hole in membrane ^{28–32}. In other cases, like *Podoviruses* with a short tail, depolymerases (lytic enzymes) are released first to degrade the cell wall, as well as some proteins that form a channel for the DNA entry into the cytoplasm ^{33,34}. There are different factors and models proposed to be involved in the process of DNA ejection. The most accepted model is a first release caused by thermodynamic and compressing pressures, followed by further DNA motion due to the interaction with bacterial proteins and hydrodynamic forces due to the capsid emptying ^{35–39}.

Once the DNA is inside, next steps vary depending if a lytic or a lysogenic cycle follows. For obligately virulent phages (lytic cycle), the virus hijacks the cell machinery, converting the bacteria into a phage factory. For this, the invader shuts down the expression of host genes in order to replicate its own genome and express its own proteins using bacterial resources ⁴⁰. Some phages rely on the bacterial RNA polymerase ⁴¹, while others might co-reject with the genome or early express their own RNA polymerase ⁴². Towards the end of the cycle, structural proteins are expressed from late genes, which will form virion particles inside the cell. To finalize the lytic cycle, phage-produced endolysins degrade the bacterial membrane from within, leading to the release of viral particles to the environment ⁴³.

Temperate phages are capable of establishing a lysogenic cycle under specific conditions. In this case, the genome will not replicate and produce virions, but it will integrate its DNA into the bacterial genome, becoming a prophage ⁴⁴. The prophage is transferred vertically with bacterial replication and remains inactive. Certain conditions, like UV light ⁴⁵ or certain physico-chemical factors ⁴⁶, induce the prophage. The induction leads to excision from the host genome, initiating the phage lytic cycle ⁴⁴. The cycle will finish with the viral progeny being released. This process has a great impact in bacterial horizontal gene transfer (HGT), as they can take some bacterial DNA during excision apart of their own genome ^{47–49}, and transfer it to the new host. This process is known as transduction ^{50,51}, and can be generalized, in which any genome portion will be randomly taken ⁵², or specialized. In this last one, phages will take only specific regions of the host genome, commonly those in the proximity of regions of DNA with a phage packaging signal sequence ^{53,54}. The decision between the lytic-lysogenic cycle was not fully understood, but recent insights into this showed that it can be regulated by phage-triggered communication peptides between infected cells ⁵⁵ or by a quorum-sensing system based on host repressor genes ⁵⁶.

Box 1 – Phage structure and taxonomy

For over a century of their study, bacteriophages followed a basic classification based on morphology and/or genomic material type. In this classification, the most common phage morphology was that of tailed phages of the Caudovirales order, presenting a head-tail structure, which included the families *Myoviridae*, *Siphoviridae*, and *Podoviridae*. These families were grouped depending on the length of the tail and its contractile capability. Interestingly, Caudovirales' morphogenesis modules are related to those of herpesviruses¹²⁹, indicating an evolutionary relationship of viruses infecting very distant hosts. Other group recently enlarged with metagenomic analysis⁶⁵, are filamentous phages. Their virion is formed by an elongated filament capsid comprised of proteins that contain the genomic material. A less numerous group, are non-tailed icosahedral phages, which have a protein capsid that contains the genome inside, in some cases including an additional lipid membrane. Some of these viruses share a structural features close to the eukaryotic adenoviruses¹³⁰, remarking evolutionary relationships of viruses from different domains of life. **Fig. 1** displays a schematic representation of the morphologies for the main phages used in this thesis, indicating families that present that morphology.

In addition to the morphological classification, viruses can also be classified based on their genomic material. The early Baltimore classification was focused on the molecular composition of the genome^{131,132}. Following this approach for phage classification¹³³, the main group with dsDNA genome is mainly comprised of Caudovirales, but also families like *Tectiviridae* and *Plasmaviridae*. Recently, starting with the analysis of spounaviruses¹³⁴, the ICTV reorganized the Caudovirales classification in 14 families: *Ackermannviridae*, *Autographiviridae*, *Chaseviridae*, *Demereciviridae*, *Drexelvriidae*, *Gueliniviridae*, *Herelleviridae*, *Myoviridae*, *Podoviridae*, *Rountreeviridae*, *Salasmaviridae*, *Schitoviridae*, *Siphoviridae*, *Zobellviridae*. The group with ssDNA genome include the families *Microviridae* or *Inoviridae*, while the *Leviviruses* have ssRNA genomes. Some phages containing dsRNA have also been described, like in the *Cystoviridae* family. However, new approaches re-classified them based on their evolutionary relationship and the genomic sequence similarity¹³⁵.

Therefore, phages, and mainly prophages, are part of the big group of mobile genetic elements (MGE) that shape bacterial ecosystems and evolution ⁵⁷. Some MGEs, like prophages and conjugative plasmids, typically perform HGT. Transposons and insertion sequences are DNA segments that re-locate in random sequences of the same or different DNA molecule. On the other hand, integrons relocated to site-specific regions. Importantly, MGEs have a great impact on ecosystems and evolution, affecting the transfer, for example, of defense systems and antibiotic resistance genes ⁵⁸.

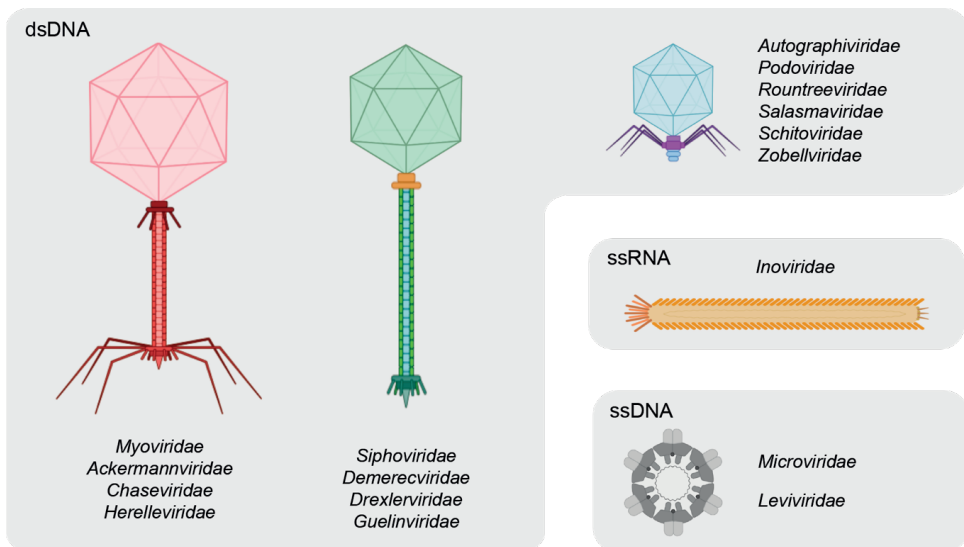


Figure 1. Schematic representation of phage morphologies. Families belonging to each morphological classification are indicated and genetic material type that they contain. Created with BioRender.com.

1.1.2. Inovirus replication cycle

The *Inoviridae* family contains a total of 21 genera ⁵⁹. Usually exemplified by phage M13, these phages are non-enveloped, protein filaments (**Fig. 1**) that contain a circular, ssDNA genome varying from 5.5 to 12.4 kb ⁶⁰. They code for 7-17 proteins, which are sufficient to hijack the host machinery and transform the cell into a phage-releasing factory. This infectious process differs in many aspects from the common *Caudovirales* described previously.

To begin the replication cycle, M13 virions bind specifically to the bacterial pili, and eject the DNA through the membrane into the cytoplasm. Once inside, host proteins synthesize a complementary strand of the ssDNA, creating a circular dsDNA molecule. Next, this double-strand form replicates to create multiple copies of the genome. Simultaneously, the genome is used to express protein products that will form new virions. Later, ssDNA genomes are produced from the dsDNA form via rolling-circle replication ⁶¹. This DNA will associate with binding proteins, whose role is directing each genomic molecule to the membrane. Once there, binding proteins is exchanged by coat proteins and then the virions extruded out of the bacteria ^{62,63}. Unlike virulent phages, M13 infection does not lead to lysis or death of the host, but a permanent or chronic infection is established. Therefore, virions are constantly released to the media while bacteria remain intact and even replicate. In some cases, the genome can integrate in the bacterial chromosome ⁶⁴, instead of persisting episomally like in the case of M13.

Inoviruses have been found to be pervasive in bacteria in many environments ⁶⁵, even favoring the host in some conditions ⁶⁶. Their spreading mechanism differs totally from *Caudovirales*; instead of breaking the cell to pursue progeny release and production, they reprogram the cell machinery to produce them through long-time periods, becoming a constant factory of phage particles. This has allowed developing many applications ⁶⁰ such as molecular tools, protein production by phage display, or even their use as metal-like components in batteries.

1.2. Defense systems. The bacterial arsenal to counter-attack

Similarly to other domains of life, the interaction between bacteria and phages might have evolved in some cases for the benefit of both parts. This is the case of some prophages whose presence in the genome might provide higher fitness for survival to those cells containing it ⁶⁷. For example, in some pathogenic strains, the prophage presence increases their virulence and survival when facing the immune system ⁶⁸. In other cases, the prophage helps to avoid the infection by other invaders, known as superinfection exclusion, such as other virulent phages whose aim is to lyse the cell ⁶⁹. However, even though these positive interactions take place, up to date the most described and understood ones are those of negative nature. Given the need of survival in hostile environments, several defense systems have developed through evolution ^{9,70}, as their usage might result in survival or death.

1.2.1. Defense system classification

The vast diversity of defense systems in the microbial community and their different features make their classification challenging. Defense systems can be grouped based on which step of phage replication cycle is targeted. Some of them might block early stages like adsorption ^{71–73} or DNA entry ⁷⁴. But the most abundant ones discovered to date act on or interfere with phage DNA replication ⁹. As last resource, it is possible to speculate that some defense system might act late on infection, when other defense systems have failed, sacrificing the cell for the benefit of the bacterial population. All these systems differ on their *modus operandi*. Some act on an indirect form, in which bacteria avoid contact with the phage particles, therefore acting as prevention of the infection. For this, the host bacteria will just play unrecognizable for the phage, as by hiding themselves in extracellular polymer matrix ⁷⁵, blocking or not expressing the phage receptor, or just by spontaneous mutations in the population that will affect the phage interaction with the surface. In other cases, the systems defend directly blocking the viral infection once it has started, interfering with the

replication process. For this case, the bacteria will use molecular mechanisms or metabolic changes to actively interfere with the viral infection ⁷⁶.

These systems can originate in the host through evolution to prevent or suppress phage infections, or they can be acquired by MGE ^{77,78}. As described previously, MGEs might help in exchange of defense systems between different species or genus ⁵⁷. Therefore, some systems evolved, to fight a certain invader, might be transferred and used against other invader which has not evolved to counter-attack it. In addition, in this category falls the interesting case of systems originally developed temperate phages to outcompete other phages during co-infection, but acquired by bacteria for their benefit ⁷⁹.

Additionally, defense systems can act as innate or adaptive forms of immunity ⁸⁰. In the innate immunity, systems will be based on blocking early stages of the infection or in recognizing self vs non-self DNA ¹¹. The only adaptive defense system discovered up to date targets specific sequences of phages that bacteria have previously encountered. Then it needs a previous exposure to the invader in order to create immune memory ¹².

1.2.2. Mechanisms of action

This thesis focuses on studying the mechanism of defense systems that act directly to clear infections once the DNA enters the cell. This group also presents a big diversity at the molecular level.

The most widespread defense system ⁸¹ and known for longest are Restriction-Modification (R-M) systems. There are four types of R-M systems, classified depending on subunit composition and sequence recognition, among others ⁸². The most studied ones are Type II, comprised of a methyltransferase that modifies self-DNA, and a restriction endonuclease. The latest recognizes specific sites and cleaves foreign DNA based on the methylation pattern ⁸³. Analogously, Dnd and Ssp systems have similar mechanisms, they perform DNA phosphorothioation of the backbone phosphate instead of methylation of the nucleobase ^{84,85}. In these cases, the absence of the DNA modification enzyme leads to toxicity due to the endonuclease.

Another system in which methylases play a role is the bacteriophage exclusion (BREX) system. In this case, the system will not cleave, but hamper the replication of unmethylated DNA^{86,87}. However, the absence of the methylase does not present a toxic effect for the bacteria⁸⁸.

Our knowledge on R-M-related defense systems is continuously expanding, and linked to them more systems are being discovered through analysis of bacterial genomes. A recent example is DISARM (defense island system associated with restriction–modification). These systems, although not abundant, are widespread in bacteria and archaea. They have shown in experimental assays a broad anti-phage activity⁸⁹. They contain a core-gene cluster formed by the gene *drmA* coding for a predicted DNA helicase, the gene *drmB* with a DUF1998 domain, and the gene *drmC* coding for a predicted phospholipase D. In addition, these systems often include MTase-coding genes (*drmMI* or *drmMII*) and/or other genes coding for proteins with predicted RNA-helicase (*drmD*) or unknown (*drmE*) functions. Although the functional domains present in DISARM systems suggested until now a mechanism similar to RM systems, it has shown to be mechanistically distinct from other well-known systems^{90,91}.

Other defense systems interfere with the infection based on recognition of specific DNA sequences. In eukaryotes, Argonaute proteins mediate the degradation of exogenous RNA using small interfering-RNAs (siRNA) as guides to recognize their targets⁹². In some bacteria, prokaryotic homologues of the Argonaute (pAgo) proteins can also interfere against invader DNA using similar mechanisms^{93–95}. While this process is not as well studied in prokaryotic cells as it is in eukaryotes, the pAgo mechanism has been described in *Thermus thermophilus* (TtAgo) and in *Rhodobacter sphaeroides* (RsAgo)^{96,97}. The TtAgo mechanism of action is based on DNA-DNA interference rather than the RNA-RNA interference of eukaryotic Ago⁹⁸. This protein also has an endonuclease (slicer) domain that allows it to cleave both single-stranded DNA and negatively supercoiled double-stranded DNA, normally of plasmid origin. Although it is unclear how the DNA guides used by TtAgo are formed, it appears that the

activity of the pAgo protein itself is necessary for the production of DNA guides. RsAgo, in contrast, uses small RNA molecules as guides to target foreign DNA molecules ⁹⁹. Of note, RsAgo lacks the slicer domain, meaning that DNA interference is caused by binding the target rather than by cleaving it. Recently, the short pAgo system from *Maribacter polysiphoniae* (MapAgo) showed to hamper phage lysis in cultures ¹⁰⁰. Short pAgo systems encode for pAgo and a protein containing an Analog of PAZ (APAZ) domain and a toll-interleukin receptor (TIR) domain, also present in the defense system Thoeris ^{101,102}. The MapAgo complex uses RNA guides to bind ssDNA targets, which triggers the NAD(P)ase activity of the TIR domain. This leads to cell death due to NAD(P)⁺ depletion. However, MapAgo was still able to provide anti-phage activity independently of the NAD(P)ase activity.

Another defense system that targets specific invader sequences are CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats - CRISPR associated proteins) systems. They are present in 39% of the analyzed bacterial genomes ⁸¹ and are the only adaptive form of immunity in prokaryotes described up to date ^{103,104}. These systems are composed of the CRISPR array and a set of *cas* genes. The former contains the memory sequences to identify invaders (spacers), that alternate with repeats to create the array ¹⁰⁵. The spacers will code for a guide RNA that will bind the Cas proteins to target invader sequences and cleave them. Other Cas proteins are specialized in the acquisition of new memory spacers to target future infections ^{106,107}. Up to date, based on different mechanism, CRISPR-Cas systems are classified in two classes, including six types and 33 subtypes ¹⁰⁸.

1.3. Phage therapy. The human vessel as the battlefield

In 1915, Frederick Twort, a physician and microbiologist interested in the study of the vaccinia virus, observed for the first time what we consider today as a plaque or centre of infection. In an attempt to cultivate vaccinia, he observed glassy and transparent areas in agar cultures containing micrococci bacteria ¹⁰⁹. Without having much knowledge about the “contagious agents” that he studied, he described many features that we still use to define a bacteriophage: they pass a porcelain filter, they multiply in presence of the host micro-organism and live at their expense, they lyse this micro-organism impeding its growth, they lyse better young cultures and not well-grown ones, they can be transmitted from culture to culture, and they are host-dependent ^{110,111}. Therefore, Twort is considered by many the father of the bacteriophage field. Two years later, the microbiologist Felix D’Hérelle isolated and described the phenomena of bacteria lysis by this agents ¹¹², that he coined as bacteriophages (from the Greek *βακτήρια-φάγειν*, or “*bacteria eaters*”). In further investigations, D’Hérelle proceeded to test their use as therapeutic agents to treat bacterial infections, for which he can be considered as the father of phage therapy.

Both of them discovered this phenomenon independently and described with great detail some key features of bacteriophage infection, opening the doors to an outstanding field that many others followed in the 1920s and 1930s. However, the interest of using phages as anti-microbials quickly faded away, due to the discovery of penicillin in 1929 ^{19,113}. The former Soviet republics like Georgia and Russia continued making use of phage therapy ¹¹⁴, which has slightly contributed to their research and study after the abandoning of phage therapy. In recent years, Western medicine has started to reconsider the therapeutic use of phages due to the alarming rise in infections caused by multidrug resistant (MDR) bacteria ¹¹⁵. Although still much needs to be understood about phage-bacteria interactions, clinical trials are recently providing insights on it ¹¹⁶. However, the current bottleneck for phage therapy seem to be external factors like regulations and policy for its clinical use ¹¹⁷.

In this everlasting battle, multiple mechanisms by which bacteria defend against phages have been developed and recently uncovered ^{70,118}. But phages have also counter-developed many mechanisms to counteract these defense systems or the acquired resistance ¹⁰. For example, some of these strategies include the expression of proteins that inhibit defensive host proteins ^{119,120} or the acquisition of DNA mutations in the target sequence target ¹²¹. Therefore, a constant evolutionary race takes place with new mechanisms developed. This is a big advantage of phages versus antibiotics, as the former can adapt and/or deploy anti-defense systems of their own to overcome the defense mechanisms of bacteria ^{122–127}, while antibiotics are static chemical molecules.

With all this considered, it is clear that a proper understanding of the complex interactions between phages and bacteria ¹³ is required in order to successfully implement phage therapy ¹²⁸. All the different outcomes that might result due to the presence of all defense systems and how they might impact phage resistance still need to be fully studied in clinical settings. Therefore, studies like those contained in this thesis are of great importance to contribute to this knowledge and bring closer the proper use of phage therapy.

1.4. Thesis outline

The main goal of this thesis is understanding the interactions between phages and bacteria, specifically those in which defense systems are involved. This goal correlates with the use of bacteriophages in phage therapy applications. Due to the dynamic and unpredictable nature of these interactions, a deep understanding of all the possible ways of interaction is required before the application of phages as a treatment. Phage therapy is a promising alternative to solve the antibiotic resistance crisis worldwide, being used independently or in combinatory approaches to cure infections.

As an overview of this field at the moment that this thesis is published, **Chapter 2** offers an extensive review of defense systems focusing on a clinical point of view. Here we analyze the forms by which bacteria might develop bacteriophage resistance as well as the mechanisms of action of known defense systems.

The next two chapters comprise the core of the PhD thesis, consisting on a detailed characterization of a recently discovered innate defense system; DISARM. In **Chapter 3**, we studied DISARM from the biological point of view. DISARM system Class 1 from *Serratia* sp. SCBI as characterized in its role as a defense system. For this, we studied the essentiality of its components to provide anti-phage and anti-conjugative defense. We observed a broad protection anti-phage effect, not only for diverse virulent phages, but also effective to clear chronic infection by the inovirus M13. In addition, we proved its activity against invasive conjugative plasmids, and describe the role of methylation patterns for this defense.

In **Chapter 4**, we provided a structural characterization of the DISARM effector complex formed by the proteins DrmA and DrmB. We reconstructed a 3D-model of this protein complex using Cryo-EM, obtaining a resolution of 2.8 Å. Using this model, together with biochemical assays, we described the molecular mechanism of action of the DISARM core complex, which includes the DUF1998 domain. This domain is found in many defense systems of bacteria but its role in defense was unknown. We show that the DrmAB complex binds to 5' ssDNA ends

in order to scan the methylation pattern of DNA molecules and distinguish self- vs non-self DNA. It is regulated by a trigger loop that occludes the DNA entry site until a 5' end is found, avoiding autoimmunity due to action of the complex when not required. In this protein complex, DUF1998 coordinates big conformational changes occurring for the recognition of the methylation pattern.

In **Chapter 5**, we studied a pAgo system; an innate defense system which uses sequence recognition to target the invader. Specifically we studied the short prokaryotic Argonaute TIR-APAZ (SPARTA) systems from *Crenotalea thermophila* 70 (CrtAgo) and *Maribacter polysiphoniae* (MapAgo). We observed that none of them provided a strong anti-phage protection in the first encounter with the phages. However, the replication efficiency overtime of both virulent phages and chronic-infecting M13 were reduce in presence of the system.

Finally, the thesis concludes with **Chapter 6**, an isolation, sequencing and characterization of two coliphages that, in addition to other model phages, were used to characterize the defense systems in previous chapters. Bacteriophages Nami and Shinka are Tequatroviruses isolated for the antibiotic-resistant, uropathogenic *E. coli* isolate R10256. We provide a description of both phages at the genomic, morphological and infective levels.

1.5. References

1. Lang,J.M. and M. Eric Benbow (2013) Species Interactions and Competition. Nat. Educ. Knowl.
2. Ricklefs,R.E. (2008) Disintegration of the ecological community. Am. Nat., 172.
3. Brooker,R.W., Callaway,R.M., Cavieres,L.A., Kikvidze,Z., Lortie,C.J., Michalet,R., Pugnaire,F.I., Valiente-Banuet,A. and Whitham,T.G. (2009) Don't diss integration: A comment on Ricklefs's disintegrating communities. Am. Nat., 174.
4. Muthamilarasan,M. and Prasad,M. (2013) Plant innate immunity: An updated insight into defense mechanism. J. Biosci., 38.
5. Campo,S., Gilbert,K.B. and Carrington,J.C. (2016) Small RNA-Based Antiviral Defense in the Phytopathogenic Fungus *Colletotrichum higginsianum*. PLoS Pathog., 12.
6. Swevers,L., Liu,J. and Smagghe,G. (2018) Defense mechanisms against viral infection in *Drosophila*: RNAi and non-RNAi. Viruses, 10.
7. Harrison,S. and Cornell,H. (2008) Toward a better understanding of the regional causes of local community richness. Ecol. Lett., 11.
8. Twort,F.W. (1926) The Twort-D'Herelle Phenomenon. Lancet, 207.
9. Bernheim,A. and Sorek,R. (2020) The pan-immune system of bacteria: antiviral defence as a community resource. Nat. Rev. Microbiol., 18.
10. Samson,J.E., Magadán,A.H., Sabri,M. and Moineau,S. (2013) Revenge of the phages: Defeating bacterial defences. Nat. Rev. Microbiol., 11.
11. Isaev,A.B., Musharova,O.S. and Severinov,K. V. (2021) Microbial Arsenal of Antiviral Defenses – Part I. Biochem., 86.
12. Isaev,A.B., Musharova,O.S. and Severinov,K. V. (2021) Microbial Arsenal of Antiviral Defenses – Part II. Biochem., 86.
13. Egado,J.E., Costa,A.R., Aparicio-Maldonado,C., Haas,P.-J. and Brouns,S.J.J. (2021) Mechanisms and clinical importance of bacteriophage resistance. FEMS Microbiol. Rev., fuab048.
14. Forterre,P. (2013) The virocell concept and environmental microbiology. ISME J., 7.
15. Koonin,E. V and Kuhn,J.H. (2021) Viruses Defined by the Position of the Virosphere within the Replicator Space. Microbiol. Mol. Biol. Rev., 85, e00193-20.
16. Silpe,J.E., Bassler,B.L., Silpe,J.E. and Bassler,B.L. (2019) A Host-Produced Quorum-Sensing Autoinducer Article A Host-Produced Quorum-Sensing Autoinducer Controls a Phage Lysis-Lysogeny Decision. Cell, 176.
17. Erez,Z., Steinberger-levy,I., Shamir,M., Doron,S., Stokar-avihail,A., Peleg,Y. and Melamed,S. (2017) Communication between viruses guides lysis – lysogeny decisions. Nature, nature21049.
18. Ackermann,H.-W. (1998) Tailed bacteriophages: the order Caudovirales. In Advances in Virus Research. 2.
19. Wittebole,X., De Roock,S. and Opal,S.M. (2013) A historical overview of bacteriophage therapy as an alternative to antibiotics for the treatment of bacterial pathogens. Virulence, 5.

20. Fields,B.N., Knipe,D.M. David M. and Howley,P.M. (1996) *Virology* 3rd. Lippincott-Raven Publishers.
21. King,A. (2012) *Caudovirales*. In *Virus Taxonomy*. Elsevier.
22. Bertozzi Silva,J., Storms,Z. and Sauvageau,D. (2016) Host receptors for bacteriophage adsorption. *FEMS Microbiol. Lett.*, 363, fnw002.
23. Nobrega,F.L., Vlot,M., Jonge,P.A., Dreesens,L.L., Beaumont,H.J.E., Lavigne,R., Dutilh,B.E. and Brouns,S.J.J. (2018) Targeting mechanisms of tailed bacteriophages. *Nat. Rev. Microbiol.*, 16.
24. Jonge,P.A. De, Nobrega,F.L., Brouns,S.J.J. and Dutilh,B.E. (2018) Molecular and Evolutionary Determinants of Bacteriophage Host Range. *Trends Microbiol.*
25. Evilevitch,A. (2018) The mobility of packaged phage genome controls ejection dynamics. *Elife*, 7.
26. González-García,V.A., Pulido-Cid,M., Garcia-Doval,C., Bocanegra,R., Van Raaij,M.J., Martín-Benito,J., Cuervo,A. and Carrascosa,J.L. (2015) Conformational changes leading to T7 DNA delivery upon interaction with the bacterial receptor. *J. Biol. Chem.*, 290.
27. Wang,C., Tu,J., Liu,J. and Molineux,I.J. (2019) Structural dynamics of bacteriophage P22 infection initiation revealed by cryo-electron tomography. *Nat. Microbiol.*, 4.
28. Boulanger,P., Jacquot,P., Plançon,L., Chami,M., Engel,A., Parquet,C., Herbeuval,C. and Letellier,L. (2008) Phage T5 straight tail fiber is a multifunctional protein acting as a tape measure and carrying fusogenic and muralytic activities. *J. Biol. Chem.*, 283.
29. Cumby,N., Reimer,K., Mengin-Lecreulx,D., Davidson,A.R. and Maxwell,K.L. (2015) The phage tail tape measure protein, an inner membrane protein and a periplasmic chaperone play connected roles in the genome injection process of E.coli phage HK97. *Mol. Microbiol.*, 96.
30. Taylor,N.M.I., Prokhorov,N.S., Guerrero-Ferreira,R.C., Shneider,M.M., Browning,C., Goldie,K.N., Stahlberg,H. and Leiman,P.G. (2016) Structure of the T4 baseplate and its function in triggering sheath contraction. *Nature*, 533.
31. Browning,C., Shneider,M.M., Bowman,V.D., Schwarzer,D. and Leiman,P.G. (2012) Phage pierces the host cell membrane with the iron-loaded spike. *Structure*, 20.
32. Wenzel,S., Shneider,M.M., Leiman,P.G., Kuhn,A. and Kiefer,D. (2020) The central spike complex of bacteriophage T4 contacts PpiD in the periplasm of Escherichia coli. *Viruses*, 12.
33. Leptihn,S., Gottschalk,J. and Kuhn,A. (2016) T7 ejectosome assembly: A story unfolds. *Bacteriophage*, 6, e1128513.
34. Molineux,I.J. (2001) No syringes please, ejection of phage T7 DNA from the virion is enzyme driven. *Mol. Microbiol.*, 40.
35. Molineux,I.J. and Panja,D. (2013) Popping the cork: Mechanisms of phage genome ejection. *Nat. Rev. Microbiol.*, 11.
36. Smith,D.E., Tans,S.J., Smith,S.B., Grimes,S., Anderson,D.L. and Bustamante,C. (2001) The bacteriophage ϕ 29 portal motor can package DNA against a large internal force. *Nature*, 413.

37. Panja,D. and Molineux,I.J. (2010) Dynamics of bacteriophage genome ejection *in vitro* and *in vivo*. *Phys. Biol.*, 7.
38. Kemp,P., Gupta,M. and Molineux,I.J. (2004) Bacteriophage T7 DNA ejection into cells is initiated by an enzyme-like mechanism. *Mol. Microbiol.*, 53.
39. Choi,K.H., McPartland,J., Kaganman,I., Bowman,V.D., Rothman-Denes,L.B. and Rossmann,M.G. (2008) Insight into DNA and Protein Transport in Double-Stranded DNA Viruses: The Structure of Bacteriophage N4. *J. Mol. Biol.*, 378.
40. De Smet,J., Hendrix,H., Blasdel,B.G., Danis-Wlodarczyk,K. and Lavigne,R. (2017) *Pseudomonas* predators: Understanding and exploiting phage-host interactions. *Nat. Rev. Microbiol.*, 15.
41. Hinton,D.M. (2010) Transcriptional control in the prereplicative phase of T4 development. *Viol. J.*, 71.
42. Drobysheva,A. V., Panafidina,S.A., Kolesnik,M. V., Klimuk,E.I., Minakhin,L., Yakunina,M. V., Borukhov,S., Nilsson,E., Holmfeldt,K., Yutin,N., et al. (2021) Structure and function of virion RNA polymerase of a crAss-like phage. *Nature*, 589.
43. Hobbs,Z. and Abedon,S.T. (2016) Diversity of phage infection types and associated terminology: the problem with 'Lytic or lysogenic'. *FEMS Microbiol. Lett.*, 363, fnw047.
44. Nash,H.A. (1981) Integration and excision of bacteriophage lambda: the mechanism of conservation site specific recombination. *Annu. Rev. Genet.*, 15.
45. Barnhart,B.J., Cox,S.H. and Jett,J.H. (1976) Prophage induction and inactivation by UV light. *J. Virol.*, 18.
46. Choi,J., Kotay,S.M. and Goel,R. (2010) Various physico-chemical stress factors cause prophage induction in *Nitrosospira multiformis* 25196- an ammonia oxidizing bacteria. *Water Res.*, 44.
47. Canchaya,C., Fournous,G., Chibani-Chennoufi,S., Dillmann,M.L. and Brüssow,H. (2003) Phage as agents of lateral gene transfer. *Curr. Opin. Microbiol.*, 6.
48. Eggers,C.H., Gray,C.M., Preisig,A.M., Glenn,D.M., Pereira,J., Ayers,R.W., Alshahrani,M., Acabbo,C., Becker,M.R., Bruenn,K.N., et al. (2016) Phage-mediated horizontal gene transfer of both prophage and heterologous DNA by ϕ BB-1, a bacteriophage of *Borrelia burgdorferi*. *Pathog. Dis.*, 74.
49. Wang,G.H., Sun,B.F., Xiong,T.L., Wang,Y.K., Murfin,K.E., Xiao,J.H. and Da Huang,W. (2016) Bacteriophage WO can mediate horizontal gene transfer in endosymbiotic *Wolbachia* genomes. *Front. Microbiol.*, 7.
50. Chiang,Y.N., Penadés,J.R. and Chen,J. (2019) Genetic transduction by phages and chromosomal islands: The new and noncanonical. *PLoS Pathog.*, 15, 1–7.
51. Penadés,J.R., Chen,J., Quiles-Puchalt,N., Carpena,N. and Novick,R.P. (2015) Bacteriophage-mediated spread of bacterial virulence genes. *Curr. Opin. Microbiol.*, 23.
52. Schmiegier,H. and Buch,U. (1975) Appearance of transducing particles and the fate of host DNA after infection of *Salmonella typhimurium* with P22-mutants with increased transducing ability (HT-mutants). *MGG Mol. Gen. Genet.*, 140.
53. Kwoh,D.Y. and Kemper,J. (1978) Bacteriophage P22-mediated specialized transduction in *Salmonella typhimurium*: high frequency of aberrant prophage excision. *J. Virol.*, 27.

54. Morse, M.L., Lederberg, E.M. and Lederberg, J. (1956) Transduction in *Escherichia coli* K-12. *Genetics*, 41.
55. Erez, Z., Steinberger-Levy, I., Shamir, M., Doron, S., Stokar-Avihail, A., Peleg, Y., Melamed, S., Leavitt, A., Savidor, A., Albeck, S., et al. (2017) Communication between viruses guides lysis-lysogeny decisions. *Nature*, 541.
56. Silpe, J.E. and Bassler, B.L. (2019) A Host-Produced Quorum-Sensing Autoinducer Controls a Phage Lysis-Lysogeny Decision. *Cell*, 176, 268-280.e13.
57. Koonin, E. V., Makarova, K.S., Wolf, Y.I. and Krupovic, M. (2020) Evolutionary entanglement of mobile genetic elements and host defence systems: guns for hire. *Nat. Rev. Genet.*, 21.
58. Partridge, S.R., Kwong, S.M., Firth, N. and Jensen, S.O. (2018) Mobile genetic elements associated with antimicrobial resistance. *Clin. Microbiol. Rev.*, 31.
59. Knezevic, P., Adriaenssens, E.M., Siddell, S.G., Lefkowitz, E.J., Sabanadzovic, S., Simmonds, P., Zerbini, F.M., Smith, D.B. and Orton, R.J. (2021) ICTV virus taxonomy profile: Inoviridae. *J. Gen. Virol.*, 102.
60. Mai-Prochnow, A., Hui, J.G.K., Kjelleberg, S., Rakonjac, J., McDougald, D. and Rice, S.A. (2015) Big things in small packages: The genetics of filamentous phage and effects on fitness of their host. *FEMS Microbiol. Rev.*, 39.
61. Nomura, N. and Ray, D.S. (1980) Replication of bacteriophage M13. XV. Location of the specific nick in M13 replicative form II accumulated in *Escherichia coli* polAex1. *J. Virol.*, 34.
62. Rasched, I. and Oberer, E. (1986) Ff coliphages: Structural and functional relationships. *Microbiol. Rev.*, 50.
63. Zinder, N.D. and Horiuchi, K. (1985) Multiregulatory element of filamentous bacteriophages. *Microbiol. Rev.*, 49.
64. King, A.M.Q., Adams, M.J., Carstens, E.B. and Lefkowitz, E.J. eds. (2012) Family - Inoviridae. In *Virus Taxonomy*. Elsevier, San Diego.
65. Roux, S., Krupovic, M., Daly, R.A., Borges, A.L., Nayfach, S., Schulz, F., Sharrar, A., Matheus Carnevali, P.B., Cheng, J.F., Ivanova, N.N., et al. (2019) Cryptic inoviruses revealed as pervasive in bacteria and archaea across Earth's biomes. *Nat. Microbiol.*, 4.
66. Tarafder, A.K., von K  gelgen, A., Mellul, A.J., Schulze, U., Aarts, D.G.A.L. and Bharat, T.A.M. (2020) Phage liquid crystalline droplets form occlusive sheaths that encapsulate and protect infectious rod-shaped bacteria. *Proc. Natl. Acad. Sci.*
67. Wendling, C.C., Refardt, D. and Hall, A.R. (2020) Fitness benefits to bacteria of carrying prophages and prophage-encoded antibiotic-resistance genes peak in different environments. *Evolution* (N. Y.). evo.14153.
68. Fortier, L.C. and Sekulovic, O. (2020) Importance of prophages to evolution and virulence of bacterial pathogens. *Virulence*, 4.
69. Bondy-Denomy, J., Qian, J., Westra, E.R., Buckling, A., Guttman, D.S., Davidson, A.R. and Maxwell, K.L. (2016) Prophages mediate defense against phage infection through diverse mechanisms. *ISME J.*, 10.

70. Doron,S., Melamed,S., Ofir,G., Leavitt,A., Lopatina,A., Keren,M., Amitai,G. and Sorek,R. (2018) Systematic discovery of antiphage defense systems in the microbial pangenome. *Science*, 80.
71. Moak,M. and Molineux,I.J. (2000) Role of the Gp16 lytic transglycosylase motif in bacteriophage T7 virions at the initiation of infection. *Mol. Microbiol.*, 37.
72. Lu,M. and Henning,U. (1994) Superinfection exclusion by T-even-type coliphages. *Trends Microbiol.*, 2.
73. Manning,A.J. and Kuehn,M.J. (2011) Contribution of bacterial outer membrane vesicles to innate bacterial defense. *BMC Microbiol.*, 11.
74. Lu,M.J., Stierhof,Y.D. and Henning,U. (1993) Location and unusual membrane topology of the immunity protein of the *Escherichia coli* phage T4. *J. Virol.*, 67.
75. Labrie,S.J., Samson,J.E. and Moineau,S. (2010) Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.*, 8.
76. Seed,K.D. (2015) Battling Phages: How Bacteria Defend against Viral Attack. *PLoS Pathog.*, 11.
77. Forsberg,K.J. and Malik,H.S. (2018) Microbial Genomics: The Expanding Universe of Bacterial Defense Systems. *Curr. Biol.*, 28.
78. Hussain,N.A.S., Kirchberger,P.C., Case,R.J. and Boucher,Y.F. (2021) Modular Molecular Weaponry Plays a Key Role in Competition Within an Environmental *Vibrio cholerae* Population. *Front. Microbiol.*, 12.
79. Rousset,F., Dowding,J., Bernheim,A., Rocha,E.P.C. and Bikard,D. (2021) Prophage-encoded hotspots of bacterial immune systems. *bioRxiv*.
80. Hampton,H.G., Watson,B.N.J. and Fineran,P.C. (2020) The arms race between bacteria and their phage foes. *Nature*, 577.
81. Tesson,F., Hervé,A., Touchon,M., Humières,C. and Cury,J. (2021) Systematic and quantitative view of the antiviral arsenal of prokaryotes. *bioRxiv*.
82. Vasu,K. and Nagaraja,V. (2013) Diverse Functions of Restriction-Modification Systems in Addition to Cellular Defense. *Microbiol. Mol. Biol. Rev.*, 77.
83. Tock,M.R. and Dryden,D.T.F. (2005) The biology of restriction and anti-restriction. *Curr. Opin. Microbiol.*, 8.
84. Xu,T., Yao,F., Zhou,X., Deng,Z. and You,D. (2010) A novel host-specific restriction system associated with DNA backbone S-modification in *Salmonella*. *Nucleic Acids Res.*, 38.
85. Xiong,X., Wu,G., Wei,Y., Liu,L., Zhang,Y.Y., Su,R., Jiang,X., Li,M., Gao,H., Tian,X., et al. (2020) SspABCD–SspE is a phosphorothioation-sensing bacterial defence system with broad anti-phage activities. *Nat. Microbiol.*
86. Goldfarb,T., Sberro,H., Weinstock,E., Cohen,O., Doron,S., Charpak-Amikam,Y., Afik,S., Ofir,G., Sorek,R., Charpak-Amikam,Y., et al. (2015) BREX is a novel phage resistance system widespread in microbial genomes. *EMBO J.*, 34.
87. Picton,D.M., Luyten,Y.A., Morgan,R.D., Nelson,A., Smith,D.L., Dryden,D.T.F., Hinton,J.C.D. and Blower,T.R. (2021) The phage defence island of a multidrug resistant plasmid uses both BREX and type IV restriction for complementary protection from viruses. *Nucleic Acids Res.*, 49.

88. Gordeeva,J., Morozova,N., Sierro,N., Isaev,A., Sinkunas,T., Tsvetkova,K., Matlashov,M., Truncaite,L., Morgan,R.D., Ivanov,N. V., et al. (2019) BREX system of *Escherichia coli* distinguishes self from non-self by methylation of a specific DNA site. *Nucleic Acids Res.*, 47.
89. Ofir,G., Melamed,S., Sberro,H., Mukamel,Z., Silverman,S., Yaakov,G., Doron,S. and Sorek,R. (2018) DISARM is a widespread bacterial defence system with broad anti-phage activities. *Nat. Microbiol.*, 3.
90. Bravo,J.P.K., Aparicio-Maldonado,C., Nobrega,F.L., Brouns,S.J.J. and Taylor,D.W. (2022) Structural basis for broad anti-phage immunity by DISARM. *Nat. Commun.*, 13.
91. Aparicio-Maldonado,C., Ofir,G., Salini,A., Sorek,R., Nobrega,F.L. and Brouns,S.J.J. (2022) Class I DISARM provides anti-phage and anti-conjugation activity by unmethylated DNA recognition. *bioRxiv*.
92. Ma,X., Zuo,Z., Shao,W., Jin,Y. and Meng,Y. (2018) The expanding roles of Argonautes: RNA interference, splicing and beyond. *Brief. Funct. Genomics*, 17.
93. Kuzmenko,A., Oguienko,A., Esyunina,D., Yudin,D., Petrova,M., Kudinova,A., Maslova,O., Ninova,M., Ryazansky,S., Leach,D., et al. (2020) DNA targeting and interference by a bacterial Argonaute nuclease. *Nature*, 587.
94. Kuzmenko,A., Oguienko,A., Esyunina,D., Yudin,D., Kudinova,A., Maslova,O., Ninova,M., Ryazansky,S., Leach,D., Aravin,A.A., et al. (2020) pAgo-induced DNA interference protects bacteria from invader DNA. *bioRxiv*.
95. Olovnikov,I., Chan,K., Sachidanandam,R., Newman,D.K. and Aravin,A.A. (2013) Bacterial Argonaute samples the transcriptome to identify foreign DNA. *Mol. Cell*, 51.
96. Willkomm,S., Makarova,K.S. and Grohmann,D. (2018) DNA silencing by prokaryotic Argonaute proteins adds a new layer of defense against invading nucleic acids. *FEMS Microbiol. Rev.*, 42.
97. Wu,J., Yang,J., Cho,W.C. and Zheng,Y. (2020) Argonaute proteins: Structural features, functions and emerging roles. *J. Adv. Res.*, 24.
98. Swarts,D.C., Jore,M.M., Westra,E.R., Zhu,Y., Janssen,J.H., Snijders,A.P., Wang,Y., Patel,D.J., Berenguer,J., Brouns,S.J.J., et al. (2014) DNA-guided DNA interference by a prokaryotic Argonaute. *Nature*, 507.
99. Miyoshi,T., Ito,K., Murakami,R. and Uchiumi,T. (2016) Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. *Nat. Commun.*, 7.
100. Koopal,B., Potocnik,A., Mutte,S.K., Aparicio-Maldonado,C., Lindhoud,S., Vervoort,J.J.M., Brouns,S.J.J. and Swarts,D.C. (2022) Short prokaryotic Argonaute systems trigger cell death upon detection of invading DNA. *Cell*.
101. Ka,D., Oh,H., Park,E., Kim,J. and Bae,E. (2020) Structural and functional evidence of bacterial antiphage protection by Thoeis defense system via NAD⁺ degradation. *Nat. Commun.*
102. Ofir,G., Herbst,E., Baroz,M., Cohen,D., Millman,A., Doron,S., Tal,N., Malheiro,D.B.A., Malitsky,S., Amitai,G., et al. (2021) Antiviral activity of bacterial TIR domains via signaling molecules that trigger cell death. *bioRxiv*.

103. Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J. and Soria, E. (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.*, 60.
104. Brouns, S.J.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J.H., Snijders, A.P.L., Dickman, M.J., Makarova, K.S., Koonin, E. V. and Van Der Oost, J. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, 80.
105. Hille, F., Richter, H., Wong, S.P., Bratovič, M., Ressel, S. and Charpentier, E. (2018) The Biology of CRISPR-Cas: Backward and Forward. *Cell*, 172.
106. Al-Attar, S., Westra, E.R., van der Oost, J. and Brouns, S.J.J. (2011) Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. *Biol. Chem.*, 392.
107. McGinn, J. and Marraffini, L.A. (2019) Molecular mechanisms of CRISPR-Cas spacer acquisition. *Nat. Rev. Microbiol.*, 17.
108. Makarova, K.S., Wolf, Y.I., Iranzo, J., Shmakov, S.A., Alkhnbashi, O.S., Brouns, S.J.J., Charpentier, E., Cheng, D., Haft, D.H., Horvath, P., et al. (2020) Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.*, 18.
109. Twort, F.W. (1915) An investigation on the nature of ultra-microscopic viruses. *Lancet*, 186.
110. Twort, F.W. (1925) THE DISCOVERY OF THE "BACTERIOPHAGE". *Lancet*, 148.
111. Twort, F.W., End, M.R.C. and F.R.S (1930) Filter-passing transmissible bacteriolytic agents (bacteriophage). *Lancet*, xxxi.
112. d'Herelle, F. (1917) Sur un microbe invisible antagoniste des bacilles dysentériques. *Comptes rendus Acad Sci Paris*, 165.
113. Tan, S.Y. and Tatsumura, Y. (2015) Alexander Fleming (1881–1955): Discoverer of penicillin. *Singapore Med. J.*, 56.
114. Chanishvili, N. (2012) Phage Therapy—History from Twort and d'Herelle Through Soviet Experience to Current Approaches. *Adv. Virus Res.*, 83.
115. Moelling, K., Broecker, F. and Willy, C. (2018) A Wake-Up Call: We Need Phage Therapy Now. *Viruses*, 10.
116. Górski, A., Borysowski, J. and Międzybrodzki, R. (2020) Phage therapy: Towards a successful clinical trial. *Antibiotics*, 9.
117. Furfaro, L.L., Payne, M.S. and Chang, B.J. (2018) Bacteriophage Therapy: Clinical Trials and Regulatory Hurdles. *Front. Cell. Infect. Microbiol.*, 8.
118. Gao, L., Altae-Tran, H., Böhning, F., Makarova, K.S., Segel, M., Schmid-Burgk, J.L., Koob, J., Wolf, Y.I., Koonin, E. V. and Zhang, F. (2020) Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science*, 1084.
119. Pawluk, A., Davidson, A.R. and Maxwell, K.L. (2017) Anti-CRISPR: discovery, mechanism and function. *Nat. Rev. Microbiol.*, nrmicro.2017.120.
120. Chaudhary, K., Chattopadhyay, A. and Pratap, D. (2017) Anti-CRISPR proteins: Counterattack of phages on bacterial defense (CRISPR/Cas) system. *J. Cell. Physiol.*, jcp.25877.

121. Pleška,M. and Guet,C.C. (2017) Effects of mutations in phage restriction sites during escape from restriction–modification. *Biol. Lett.*, 13.
122. S,I., MB,S., TA,B. and W,A. (1987) Two DNA antirestriction systems of bacteriophage P1, darA, and darB: characterization of darA⁻ phages. *Virology*, 157.
123. C,A., TJ,S., SS,S. and DT,D. (2002) Interaction of the ocr gene 0.3 protein of bacteriophage T7 with EcoKI restriction/modification enzyme. *Nucleic Acids Res.*, 30.
124. Otsuka,Y. and Yonesaki,T. (2012) Dmd of bacteriophage T4 functions as an antitoxin against *Escherichia coli* LsoA and RnIA toxins. *Mol. Microbiol.*, 83.
125. Isaev,A., Drobiazko,A., Sierro,N., Gordeeva,J., Qimron,U., Ivanov,N. V, Severinov,K. and Yosef,I. (2020) Phage T7 DNA mimic protein Ocr is a potent inhibitor of BREX defence. *Nucleic Acids Res.*, gkaa290.
126. Malone,L.M., Warring,S.L., Jackson,S.A., Warnecke,C., Gardner,P.P., Gumy,L.F. and Fineran,P.C. (2020) A jumbo phage that forms a nucleus-like structure evades CRISPR–Cas DNA targeting but is vulnerable to type III RNA-based immunity. *Nat. Microbiol.*, 5.
127. Wiegand,T., Karambelkar,S., Bondy-Denomy,J. and Wiedenheft,B. (2020) Structures and Strategies of Anti-CRISPR-Mediated Immune Suppression. *Annu. Rev. Microbiol.*, 74.
128. Roach,D.R. and Debarbieux,L. (2017) Phage therapy: awakening a sleeping giant. *Emerg. Top. Life Sci.*, 1.
129. Rixon,F.J. and Schmid,M.F. (2014) Structural similarities in DNA packaging and delivery apparatuses in Herpesvirus and dsDNA bacteriophages. *Curr. Opin. Virol.*, 5.
130. Benson,S.D., Bamford,J.K.H., Bamford,D.H. and Burnett,R.M. (1999) Viral evolution revealed by bacteriophage PRD1 and human adenovirus coat protein structures. *Cell*, 98.
131. Koonin,E. V., Krupovic,M. and Agol,V.I. (2021) The Baltimore Classification of Viruses 50 Years Later: How Does It Stand in the Light of Virus Evolution? *Microbiol. Mol. Biol. Rev.*, 85.
132. Baltimore,D. (1971) Expression of animal virus genomes. *Bacteriol. Rev.*, 35, 235–241.
133. Chibani,C.M., Farr,A., Klama,S., Dietrich,S. and Liesegang,H. (2019) Classifying the Unclassified: A Phage Classification Method. *Viruses*, 11.
134. Lavigne,R., Darius,P., Summer,E.J., Seto,D., Mahadevan,P., Nilsson,A.S., Ackermann,H.W. and Kropinski,A.M. (2009) Classification of Myoviridae bacteriophages using protein sequence similarity. *BMC Microbiol.*, 9.
135. Turner,D., Kropinski,A.M. and Adriaenssens,E.M. (2021) A roadmap for genome-based phage taxonomy. *Viruses*, 13.



2.

An overview of bacterial defense mechanisms

Published as: *Mechanisms and clinical importance of bacteriophage resistance.* Julia E. Egido, Ana Rita Costa, Cristian Aparicio-Maldonado, Pieter-Jan Haas, Stan J.J. Brouns (2022). **FEMS Microbiology Reviews.** doi: 10.1093/femsre/fuab048

Abstract

We are in the midst of a golden age of uncovering defense systems against bacteriophages. Apart from the fundamental interest in these defense systems, and revolutionary applications that have been derived from them (e.g. CRISPR-Cas9 and restriction endonucleases), it is unknown how defense systems contribute to resistance formation against bacteriophages in clinical settings. Bacteriophages are now being reconsidered as therapeutic agents against bacterial infections due the emergence of multidrug resistance. However, bacteriophage resistance through defense systems and other means could hinder the development of successful phage-based therapies. Here we review the current state of the field of bacteriophage defense, highlight the relevance of bacteriophage defense for potential clinical use of bacteriophages as therapeutic agents, and suggest new directions of research.

2.1. Introduction

The use of bacteriophages, or phages, as therapeutic agents to treat bacterial infections began immediately after phage discovery in 1917 ^{1,2}. The initial interest in phages as anti-bacterial agents faded quickly following the discovery of penicillin two decades later ³, although phage therapy remained in use in former Soviet republics like Georgia and Russia ⁴. In recent years, Western medicine has started to reconsider the therapeutic use of phages due to the alarming rise in infections caused by multidrug resistant (MDR) bacteria ⁵. However, the success of phage therapy might be limited by the development of phage resistance by bacteria, much akin to the resistance developed towards antibiotics. Recently, multiple mechanisms by which bacteria defend against phages have been uncovered ^{6,7}, some specific for certain species or strains, others more widespread. Unlike antibiotics, phages can adapt and/or deploy anti-defense systems of their own to overcome the defense mechanisms of bacteria ^{8–14}.

The evident complexity of phage-bacteria interactions needs to be considered for phage therapy to be implemented successfully ¹⁵. It is unknown how defense systems contribute to and impact resistance formation against phages in clinical settings, and this could be a bottleneck in the development of successful phage-based therapies when left without consideration.

Here, we provide an overview of the current state of the field of natural and acquired phage resistance, highlight the relevance of phage defense for potential clinical use of phages as therapeutic agents, and suggest new directions of research.

2.1.1. The multistep process of bacteriophage infection

There are multiple families of bacteriophages, each with specific features that influence their process of infection of a bacterial host. For the purposes of this review, we will focus on phages belonging to the order *Caudovirales* ¹⁶, which are known as tailed phages and are the most widely used in clinical applications ³. Tailed phages have double-stranded DNA genomes and a structure made up of

an icosahedral head and a tail, which usually incorporates receptor binding proteins (RBPs) such as tail spikes and tail fibers at the distal end ¹⁷. These elements are responsible for the first step of infection, i.e. recognition of specific receptors on the surface of bacteria, and subsequent adsorption of the phage ¹⁷. Phage receptors on the bacterial surface are typically peptide sequences or polysaccharides present on the bacterial cell wall, as well as protruding structures such as capsules, pili or flagella ^{18,19}. Phage attachment to the host surface often occurs first through a reversible interaction with a receptor, which is then followed by an irreversible binding event to the same or a second receptor ¹⁸. Generally, phages recognize receptors with a great degree of specificity, meaning that the host range of a certain phage is often limited at the adsorption stage by the receptors available on the cell surface ²⁰.

Adsorption of the phage to its native receptor on the cell triggers ejection of the genetic material of the phage into the host cytoplasm. The mechanism of this complex phenomenon is not yet completely understood for many phage types, but in *Caudovirales* it commonly involves conformational changes of the phage triggered by binding of the phage receptor binding proteins (RBPs) to the receptor that result in the opening of the channel required for DNA release from the capsid ^{21,22}. In some phages, these conformational changes lead also to the ejection of the tape measure protein that reconfigures into a channel through which the genome translocates into the cell cytoplasm ^{23–25}. In phages with short tails (e.g. *Podoviridae*) proteins ejected together with the genome can work to form a similar channel for genome passage ²⁶. The forces behind phage genome ejection into the cell cytoplasm are still unclear, with different models proposed ²⁷. It seems that thermodynamic and compressing pressures cause the initial release of DNA ²⁸, with complete ejection being achieved by further hydrodynamic forces and/or bacterial proteins involved in transcription of the initial segment of the phage genome ^{29–31}.

If the infecting phage is obligately virulent, which is preferred for phage therapy applications ^{32,33}, the infection follows a lytic cycle once the phage genome is inside the cell. In this case, the phage hijacks the cellular machinery of the

bacteria, shutting off the expression of host genes and achieving the replication of its genome and the expression of its own genes ³⁴. For this purpose, some phages rely on the bacterial RNA polymerase ³⁵, while others encode and/or co-inject their own ³⁶. The lytic cycle culminates with the expression of late genes, which encode structural proteins and proteins necessary for bacterial host lysis. This ultimately leads to the production of more viral particles that will in the end burst out of the host cell ³⁷. However, if the infecting phage is temperate, it may also follow a lysogenic cycle. In this case, the viral genome persists within the host cell, either introduced in the bacterial chromosome as a prophage or in the bacterial cytoplasm as a plasmid. The lysis-lysogeny decision may depend on peptide-based communication between the viruses ³⁸ or on host repressor genes that form part of a quorum-sensing system ³⁹.

2.2. Mechanisms of phage resistance

Bacteria evade phage infections in different ways. Here, we classify different resistance mechanisms in three main categories:

- Receptor adaptations: random mutations or phenotypical variations in bacteria which result in decreased phage adsorption (**Fig. 1**).
- Host defense systems: molecular pathways that have specifically evolved in bacteria to prevent or suppress phage infections (**Fig. 2**).
- Phage-derived phage defense systems: molecular pathways encoded by phages to compete with other phages to the benefit of the host (**Fig. 3**).

2.2.1. Receptor adaptations leading to phage resistance

In their natural environments, bacteria are subjected to constant selective pressure, which has driven bacteria and phages into an arms race to evolve defense systems and to counter them. The arms race is characterized by high mutation rates and horizontal gene transfer, and leads to rapid evolution of genetic traits and genetic diversity ^{40–42}. Mutations that cause cell surface alterations can result in blockage of phage adsorption, and are therefore directly beneficial to the host.

Bacteria can pose a barrier to phage adsorption by decreasing the availability of the receptors to which phages bind. The acquisition of point mutations in their genome (**Fig. 1a**) is probably the simplest way by which bacteria can become fully resistant to phages. In fact, mutations in the receptor genes or their regulation have been a common way to identify the receptor of a phage ^{19,43}. These mutations occur often upon phage challenges, and can lead to a loss or decrease in the gene expression of certain receptors, or to modifications of their structure ^{44,45}. For example, *Escherichia coli* mutates *tolC* and LPS genes to resist infection by phage U136B ⁴⁶. Similarly, *Acinetobacter baumannii* mutates genes involved in the biosynthesis of capsular polysaccharides to avoid infection by phages øFG02 and øCO01 ⁴⁷.

In *Listeria monocytogenes*, loss or deficiency of wall teichoic acid rhamnosylation leads to resistance to a wide range of phages ⁴⁸, and results also in serovar diversification ⁴⁹. Other proteins involved in phage adsorption and DNA injection, like the phage infection protein from *Enterococcus faecalis* (PIP_{EF}), can also mutate as a response to phage challenges ⁵⁰.

Bacteria may also block phage adhesion by producing proteins that mask or block the phage receptors on the cell surface (**Fig. 1b**). An example of this is F plasmid-encoded protein TraT, which localizes at the cell outer membrane and binds surface-exposed regions of the outer membrane protein OmpA in *E. coli* ⁵¹. This makes this common phage receptor inaccessible for phage binding. Masking molecules, such as lipoproteins, that bind phage receptors can also be produced by bacteria under stress conditions and are released during bacterial lysis ⁵². Some bacteria also produce and release outer membrane vesicles (**Fig. 1c**) that act as cell decoys that capture and inactivate phages ⁵³. Another mechanism that can prevent phages from reaching their receptors is up-regulating the production of extra-cellular matrix typically consisting of polysaccharides, proteins, lipids, and extracellular DNA (**Fig. 1d**), in a way that protects the embedded bacteria or subsequent biofilm against phage adsorption ^{54,55}. In *Lactococcus lactis*, plasmids encoding exopolysaccharide biosynthesis genes can also confer protection against phages ⁵⁶.

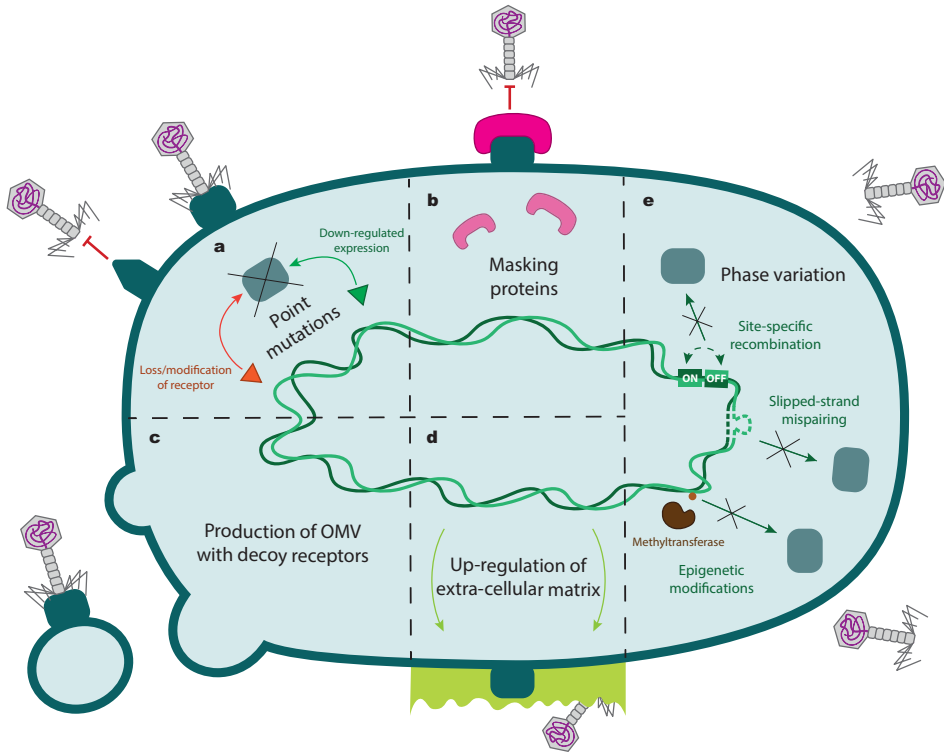


Figure 1. Host adaptations leading to phage resistance. **a)** Point mutations can lead to a loss or modification of the phage receptors (green rectangles), or to down-regulation of their expression. **b)** Receptor masking proteins like TraT of *E. coli* (pink) can bind to the surface-exposed regions of phage receptors, making them unavailable for the phages. **c)** Outer-membrane vesicles (OMV) presenting phage receptors act as decoys to prevent the phages from encountering the bacteria. **d)** An increase in the production of extra-cellular matrix (light green) leads to phage receptors being physically hidden. **e)** Phase variation occurs through three mechanisms: site-specific recombination, slipped-strand mispairing and epigenetic modifications. It can regulate the bacterial phenotype, including the expression of surface proteins like phage receptors.

In addition to this, reversible changes in the regulation of gene expression, a phenomenon known as phase variation (**Fig. 1e**), can lead to a decrease in receptor availability⁵⁷. These changes can be mediated by site-specific recombination⁵⁸, in which inversion of a DNA segment in the promoter or regulatory region of a gene causes its expression to be turned on or off. This is exemplified by the development of flagella in *Salmonella* spp. and fimbriae in *E.*

coli^{59–61}. Other receptors, such as the outer membrane protein Opc of *Neisseria meningitidis* and the subunits of *Bordetella pertussis* fimbriae^{62–64}, are regulated by slipped strand mispairing, i.e. programmed mutations that occur in defined regions during DNA recombination. Epigenetic modifications, such as altered methylation patterns on DNA sequences⁶⁵, also regulate expression of phage receptors, such as the O-antigen chains of lipopolysaccharides (LPS) in *Salmonella enterica*⁶⁶.

All of these alterations act directly on phage receptors and decrease the chances of phage adsorption. However, modifications of surface elements can come with a fitness trade-off for the host bacteria, in terms of reduced virulence or survival ability of the host^{67,68}, limiting the possibility to alter the receptor itself. Due to this, more specific defense systems that target phages within the host cell are also necessary, especially in the context of a complex microbial community^{68,69}.

2.2.2. Host phage defense systems

Bacteria have evolved defense systems dedicated to defense against mobile genetic elements such as phages. Many of them are clustered in regions of the genome known as defense islands⁷⁰, offering an opportunity for discovering new defense systems by analyzing the genetic regions in the proximity of other known defense systems. Such strategy has resulted in a significant and fast expansion of the known arsenal bacteria use to defend against phage infection. We will cover a number of different phage defense systems that have been identified, including those acting on viral nucleic acids and those causing abortive infection of the host.

- **Nucleic acid interference**

The ability to interfere with viral nucleic acids (**Fig. 2a**) is a common strategy that hosts employ to limit phage invasion and propagation. One of the most widespread and longest known examples of phage defense systems are those called Restriction-Modification (R-M) systems which act on phages with DNA genomes^{71,72}. In R-M systems, a methyltransferase (MTase) methylates

endogenous DNA at specific sites, protecting it from cleavage by the restriction endonuclease (REase) that recognizes the foreign, unmodified DNA and cleaves it within, close to, or at a distance from the recognition site ⁷³.

There are four classical types of R-M systems (I-IV), classified according to the characteristics of their specific components ⁷³. The type I R-M system consists of a protein complex of three subunits with distinct activities, the M (MTase), R (REase), and S (specificity) subunits. The S subunit dictates the target sequence specificity of both methylation and restriction by the protein complex. The abundant type II R-M system features a MTase and REase that work independently as separate proteins. These have been the major source for hundreds of commercially available restriction endonucleases used for molecular cloning. The type III R-M system also expresses the two independent MTase and REase proteins, but these exert their function as a complex. The type IV R-M system does not contain an MTase, and is thought to have evolved in response to some phages evading type I-III R-M systems by modifying their genome to evade restriction. Type IV systems overcome this counter-attack by restricting the phage's modified DNA, while the bacterial DNA remains unmethylated ^{74,75}. It is interesting to note that MTases tend to be more conserved than REases, since the latter undergo rapid evolution to keep up with mutations in phage genomes ⁷⁶.

R-M systems typically put epigenetic marks on the nucleobases. However, similar systems have been described that modify the sugar-phosphate backbone by introducing a phosphorothioate (substitution of a non-bridging oxygen with a sulfur) ⁷⁷. The Dnd system works through the double-stranded phosphorothioation of endogenous DNA by proteins DndABCDE and restriction of foreign, unmodified DNA by DndFGH ⁷⁷. Ssp proteins SspABCD also modify the host genome through phosphorothioation, but of only one of the two DNA strands ⁷⁸. This activity couples with that of SspE, which requires sensing of SspABCD to introduce nicks into foreign DNA, or with that of SspFGH, which indiscriminately damages non-phosphorothioated DNA, inhibiting its replication ^{78,79}.

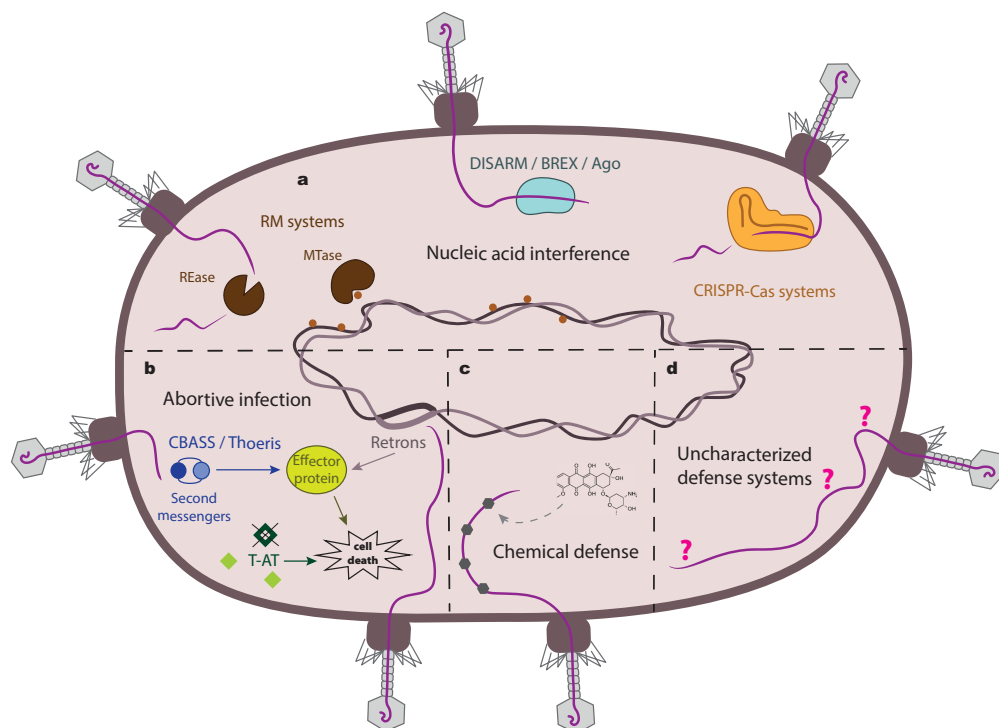


Figure 2. Host phage defense systems. **a)** Multiple defense systems act via nucleic-acid interference. R-M systems are generally composed of an MTase that methylates endogenous DNA to distinguish it from exogenous DNA, and of an REase that cleaves the exogenous, non-methylated DNA. DISARM interacts with phage DNA to prevent its circularization, thereby blocking its replication or lysogeny. BREX or Ago systems interact with phage DNA and prevent it from replicating without necessarily cleaving it. CRISPR-Cas systems are known as the adaptive immune system of bacteria. The CRISPR array contains sequences of foreign origin, that can be transcribed and processed to act as a guide for the Cas endonuclease, which recognizes and cleaves said sequences upon re-entry into the bacteria. **b)** Abortive infection comprises a series of mechanisms that lead to bacterial cell suicide. An example in which this can happen is through an imbalance in the concentration of toxins and antitoxins in a cell. Another example is through the action of effector proteins that might get activated directly, like in the case of retons, or via second messengers, like in the case of CBASS or Thoeris. These effector proteins can lead to cell death in several ways, for instance through inner membrane degradation (CBASS) or through NAD depletion (Thoeris). **c)** Bacteria can produce secondary metabolites such as daunorubicin (depicted) that intercalate phage DNA and prevent it from circularizing and replicating. **d)** Analysis of genetic defense islands has recently led to the discovery of a series of defense systems that are yet to be fully characterized. These include: Hachiman, Shedu, Gabija, Septu, Lamassu, Zorya, Kiwa, Druantia, Wadjet, RADAR, DRTs, AVAST, and pVips, among others.

Our knowledge on R-M-related defense systems is continuously expanding as more systems are being discovered through analysis of bacterial genomes. An example are the DISARM (defense island system associated with restriction–modification) systems ⁸⁰, which include MTases (adenine MTase DrmMI and/or cytosine MTase DrmMII) and proteins with domains of predicted helicase (DrmA, DrmD) and phospholipase D/nuclease (DrmC) activities or of unknown function (DrmB, DrmE). Although the exact mechanism of action of DISARM is not yet understood, it is clear that it involves methylation of the host DNA to distinguish self from non-self, and that it prevents phage DNA circularization, thereby blocking DNA replication and lysogeny at an early stage of the infection. It is also postulated that DISARM might collaborate with different R-M elements, achieving a synergistic effect against phage infection ⁸⁰. The bacteriophage exclusion (BREX) defense system also targets phage DNA upon entrance in the host cell ⁸¹. Similar to R-M systems, BREX methylates host DNA to differentiate it from exogenous DNA. But BREX does not appear to degrade non-methylated phage DNA, and instead seems to hamper replication of the phage DNA without cleavage ⁸¹. Methylated or glycosylated phage DNA is not sensitive to BREX, but deletion of the methylase gene of this system does not have deleterious effects on the bacteria ⁸².

In some bacteria, foreign DNA can also be intercepted by proteins of the Argonaute (Ago) family. These proteins are also present in eukaryotic cells, where they mediate the degradation of exogenous RNA using small interfering- or micro RNAs (siRNA, miRNA) as guides to recognize their targets. While this process is not as well studied in prokaryotic cells as it is in eukaryotes, prokaryotic Ago proteins (pAgo) have been found in *Thermus thermophilus* (TtAgo) and in *Rhodobacter sphaeroides* (RsAgo) ^{83,84}. TtAgo bases its mechanism on DNA-DNA interference rather than the RNA-RNA interference of eukaryotic Ago ⁸⁵. This protein also has an endonuclease (slicer) domain that allows it to cleave both single-stranded DNA and negatively supercoiled double-stranded DNA, normally of plasmid origin. Although it is unclear how the DNA guides used by TtAgo are formed, it appears that the activity of the protein itself is

necessary for their production. RsAgo, in contrast, uses small RNA molecules as guides to target foreign DNA molecules ⁸⁶. Of note, RsAgo lacks the slicer domain, meaning that DNA interference is caused simply by binding the target rather than by cleaving it.

2

A particular form of nucleic acid interference, CRISPR-Cas (clustered regularly interspaced short palindromic repeats and associated proteins) systems constitute the only form of adaptive immunity described in prokaryotes so far ^{87,88}. They are present in many bacterial genomes, and occasionally in plasmids ⁸⁹. A CRISPR locus in a bacterial genome is composed of a CRISPR array and a *cas* gene operon. The CRISPR array contains repeats and sequences of foreign origin called spacers, which form the immunological memory of the defense system. The *cas* operon contains all genes coding for Cas proteins that form the machinery required for immunity. Immunity is achieved via a three-stage process that involves adaptation, expression and interference ^{90–92}. During the adaptation stage, parts of the foreign genetic material are captured and integrated into the CRISPR array as a new ^{93,94}. In DNA targeting CRISPR systems, functional spacers are derived from invader sequences that are flanked by a protospacer adjacent motif (PAM), a short nucleotide sequence that ensures the targeting of foreign invaders rather than the genomic CRISPR locus ⁹⁵. At the expression stage, the CRISPR array serves as a template to transcribe a long precursor CRISPR RNA (crRNA) that is further processed into smaller mature crRNAs. Each crRNA is then loaded into Cas proteins to form an effector complex. At the stage of interference, this effector complex patrols the cell, screening for complementary sequences that are flanked by a PAM. Upon PAM recognition, the foreign genetic material is cleaved by the Cas proteins, and the infection is contained.

While sharing the general stages described above, CRISPR-Cas systems are characterized by mechanistic variability and are currently classified in two classes, six types and 33 subtypes ⁹⁶. Class 1 systems, which include types I, III and IV, are characterized by the presence of a multi-subunit Cas complex that is involved in the recognition of invader DNA (type I, IV) or RNA (type III) during

the interference stage. Class 2 systems, which include types II, V and VI, employ a single-subunit effector protein for recognition and cleavage of the foreign DNA (type II, V) or RNA (type VI) sequence. Of the six types described so far, type II is the best-known due to its applications for genome editing technology ⁹⁷.

In summary, bacteria explore a diverse set of strategies that directly block or cleave phage nucleic acids to survive phage predation.

- **Abortive infection**

Abortive infection (Abi) is a commonly used phage defense strategy (**Fig. 2b**) in which the cells sacrifice themselves before the phage completes its replication cycle to protect the rest of the population ⁹⁸. Many of the Abi systems rely on a toxin-antitoxin (T-AT) mechanism, in which the balance between a stable toxin and an unstable antitoxin determines the fate of the cell ^{99,100}. Infection by a phage triggers repression of the antitoxin promoter or termination of its transcription ¹⁰¹. The result is that the toxin prevails, causing death of the bacterium. Some of these systems, like ToxIN of *Pectobacterium atrosepticum*, are encoded by plasmids ¹⁰⁰.

Other common strategies that lead to abortive infection are characterized by the specific depletion of critical cellular resources upon viral infection, including enzymatic co-factors and nucleotides ¹⁰². Examples in *E. coli* include protease Lit, which is activated by the Gol peptide of the T4 major capsid protein and cleaves translation elongation factor Tu to arrest translation ¹⁰³. Other Abi systems trigger not an individual response but a set of events. For example, exclusion of T7 by the F plasmid-encoded PifA system occurs via reduced synthesis of macromolecules, partially impaired DNA ejection and alteration of membrane permeability ¹⁰⁴. In *Lactococcus* spp., Abi systems that can target phage gene replication and expression are constitutively expressed but are toxic to the cell when overexpressed ¹⁰⁵. Notably, most of these lactococcal defense systems are encoded by plasmids ¹⁰⁶.

Cell suicide upon detection of invading, cytosolic DNA occurs in eukaryotic cells as well. It is mediated by the production of cyclic GMP-AMP, which activates the cGAS-STING pathway ¹⁰⁷, and causes an upregulation of

transcription of inflammatory genes. A similar pathway was found in *Vibrio cholerae* biotype El Tor, where production of cyclic GMP-AMP (cGAMP) activates a phospholipase that degrades the inner membrane leading to cell death ¹⁰⁸. Introduction of the operon encoding this pathway into defective *V. cholerae* and *E. coli* strains conferred resistance to a variety of phages, suggesting an important role of this system in anti-phage defense. The system, called cyclic-oligonucleotide-based anti-phage signaling system (CBASS), has since been found in a broad range of organisms belonging to all major bacterial phyla and at least one archaeal phylum ¹⁰⁹. It is thought to be an ancestor of the eukaryotic cGAS-STING pathway.

Mutations in enzymes involved in protein maturation can also be used to prevent the spread of phage infection. In *Streptococcus thermophilus*, a mutation in the methionine aminopeptidase that impairs its catalytic activity was seen to confer resistance to a broad range of phages, seemingly by hampering virion assembly ¹¹⁰. While not exactly considered an Abi mechanism, this process comes at the cost of impairing bacterial growth for at least several of the strains studied.

In *Vibrio cholerae*, a parasitic phage satellite known as phage-inducible chromosomal island-like element (PLE) defends the bacterial population from phage attack by functioning akin to an Abi system. PLE are found integrated in the *V. cholerae* chromosomes and are excised upon infection by ICP1 phages ^{111–113}. Using both PLE- and phage-encoded products ^{111,114}, PLE replicates and hijacks the structural components of the phage to encapsidate its own genome ¹¹³, and uses protein LidI to disrupt the mechanism of lysis inhibition that would normally give ICP1 phages more time to produce new virions ¹¹⁵. Via a combination of structural hijacking and accelerated cell bursting, PLE prevent phage spreading and efficiently protect the bacterial population while transducing their own genome to other cells.

Recently described Thoeris seems to operate via an Abi mechanism as well ¹¹⁶. It presents a protein with a toll-interleukin receptor (TIR) domain which, upon phage infection, produces an isomer of cyclic ADP-ribose ¹¹⁷. This molecule acts as a second messenger and activates a protein with catalytic NADase

activity, leading to NAD depletion in the infected host. As a result of this, the bacterium presumably dies before phage progeny can mature. TIR domains appear to be specific towards certain phages, and multiple TIR proteins can be present within the same host.

Retrons, bacterial genetic elements composed of a reverse transcriptase (RT) and a non-coding RNA (ncRNA), have also been shown to protect against phage infection via abortive infection ^{7,118}. Effector proteins of multiple functions were found associated with the retons, such as ribosyltransferases, two-transmembrane domain (2TM) genes, and genes with ATPase or HNH endonuclease domains, suggesting a diversity of mechanisms by which abortive infection may be achieved. Characterization of retron Ec48 associated with a 2TM domain gene demonstrates that it acts by sensing inhibition of DNA-repair enzyme RecBCD by proteins of the infecting phage, leading to abortive infection and cell death ¹¹⁸.

More recently, dCTP deaminase and dGTPase proteins have been found to protect bacterial cells from phage infection by degrading deoxynucleotides dCTP and dGTP, efficiently eliminating these from the nucleotide pool ^{119,120}. Depletion of these deoxynucleotides during phage infection halts phage replication and likely leads to cell death ¹²⁰. While abortive infection responses can be encoded by some of the defense systems listed above, some CRISPR-Cas systems have been found to use this strategy as well. Most well-known are the type III CRISPR-Cas systems which produce small signal molecules upon target RNA detection ¹²¹. This molecule then activates unspecific nucleases and other potentially damaging activities in the cell, aborting an infection ¹²². Likewise, some type I CRISPR-Cas systems can function with an Abi mechanism. In *P. atrosepticum*, expression of a type I-F CRISPR-Cas system reduces phage progeny while hampering the survival of infected cells ¹²³.

In summary, many abortive infection-like strategies have been identified in which cells typically detect infection and initiate a self-damaging response that hampers the virus in its infection process, saving the remaining population of cells.

- **Chemical defense**

It is well documented that bacteria produce secondary metabolites amongst which compounds with antimicrobial activity (**Fig. 2c**)¹²⁴. Recently, a panel of bioactive compounds was tested to assess whether they could confer protection to *E. coli* against lysis by phage Lambda¹²⁵. Several compounds were identified that allow bacteria to proliferate in spite of the phage challenge. Most are DNA-intercalating agents, four of which produced by *Streptomyces* spp.: daunorubicin, doxorubicin, epirubicin and idarubicin. These compounds inhibit double-stranded DNA phages targeting *Streptomyces coelicolor*, *E. coli* and *Pseudomonas aeruginosa*. DNA intercalation is thought to prevent the circularization of the phage linear DNA inside the bacterial cytoplasm, or its interaction with proteins involved in replication and transcription.

- **Uncharacterized defense systems**

Bioinformatic analysis of genes in defense system clusters has led to the identification of multiple new defense systems in recent years (**Fig. 2d**). One such approach identified several putative defense systems based on the requirement that each putative system must contain at least one annotated protein domain enriched in defense islands. Some of these were experimentally confirmed to grant protection against at least one phage: Thoeris (now classified as an Abi system), Hachiman, Shedu, Gabija, Septu, Lamassu, Zorya, Kiwa, and Druantia⁶. Zorya contains components that resemble parts of the flagellar motor, and is more abundant in Gram negative species, especially Proteobacteria. Its proposed mechanism of action leads to cell death through membrane depolarization. Another defense system, Wadjet, does not seem to confer phage resistance but seems to target foreign plasmids by a still unknown mechanism⁶.

Additional defense system candidates were identified using an approach independent of domain annotations⁷. These candidates incorporate enzymatic activities not previously thought to be implicated in antiviral defense. Amongst them is the phage restriction by an adenosine deaminase acting on RNA (RADAR) system. RADAR edits RNA transcripts by catalyzing the deamination of adenosine into inosine, seemingly blocking the early stages of the phage

infection cycle. Another candidate system identified in this study is the RT family defense associated reverse transcriptases (DRT). DRTs are not linked to mobile elements, unlike most RTs found in prokaryotes, and they seem to alter phage gene expression in various ways. Anti-phage activity was likewise detected in a group of nucleoside triphosphatases (NTPases) of the STAND (signal transduction ATPases with numerous associated domains) superfamily, which were given the name antiviral ATPases/NTPases of the STAND superfamily (AVAST). Members of this superfamily found in eukaryotes are often involved in programmed cell death, so it was postulated that the AVAST system may constitute an Abi mechanism. Additionally, the study found some other proteins and systems that provided protection against T7-like phages, of which the mechanisms of action need to be further investigated.

Another example of an anti-phage defense mechanism that has recently started to be characterized is prokaryotic viperins (pVips) ¹²⁶. In animals, viperins are interferon-induced proteins that block the replication of several viruses ¹²⁷. In a similar way, pVips produce modified ribonucleotides that inhibit viral polymerase-dependent transcription, thereby protecting against infection by phage T7 ¹²⁶. pVips with anti-phage activity were identified by analyzing prokaryotic homologues of human viperin that are encoded in defense islands.

Finally, some novel defense systems that are still uncharacterized have been identified in T4- and T2-like prophages. These are briefly discussed in the following section.

2.2.3. Phage-derived phage defense systems

Interestingly, phages provide bacteria with defense systems against infection by the same or closely related phage, in a phenomenon known as superinfection exclusion (Sie) (**Fig. 3a**). Some phages produce proteins to mask the cell surface receptors, blocking new infections. This strategy also protects the newly formed phages from being inactivated as a consequence of binding to receptors coming from remains of lysed bacteria. This behavior is observed for example in phage T5, which produces lipoprotein Llp that conceals its own

receptor, outer membrane protein FhuA¹²⁸. Other phages, mostly prophages¹²⁹, use membrane-anchored or membrane-associated proteins to target and block the entry of phage DNA into the bacterial cytoplasm¹³⁰. Such proteins act by inhibiting the formation of the channel through which DNA travels across the cell membrane, by inhibiting the phage lysozyme that degrades the peptidoglycan of the bacterial cell wall, or by changing the conformation of the proteins surrounding the ejection site to prevent translocation¹³¹.

Furthermore, prophages can mediate resistance through non-Sie-like mechanisms as well (**Fig. 3b**). The RexA-RexB system, an Abi system expressed by λ -lysogenic *E. coli*, works by reducing the membrane potential of the cell, leading to a decrease in ATP production that ultimately results in cell death¹³². Another example is the phage Panchino of *Mycobacterium smegmatis*, which provides lysogens with a single subunit R-M system able to recognize a broad range of phages¹³³. Genes encoding repressor proteins that bind phage DNA may also be found in prophages¹³⁴. They are thought to have a role in protecting the viability of the lysogenized bacteria, counteracting accidental prophage transcription events. Prophage-mediated phenotypic changes in bacteria are sometimes encoded in genetic elements called morons, which are flanked by a promoter and a transcriptional terminator and can be transcribed autonomously, independent of prophage activation¹³⁵.

As with host defense systems, the discovery of phage-derived defense systems is ongoing. Analysis of Enterobacteria P4- and P2-like prophages recently led to the discovery of genetic hotspots that encode a variety of bacterial immune mechanisms¹³⁶. Amongst these is the phage anti-restriction-induced system (PARIS). This system triggers an Abi response upon sensing a phage-encoded anti-restriction protein, Ocr, which inhibits R-M systems and BREX¹³⁶. The mechanisms of action of PARIS and the other systems identified in this study remain to be further uncovered.

In summary, once inside the host, phages themselves can provide the bacteria with mechanisms of protection against further phage infection that favor both bacteria and phage.

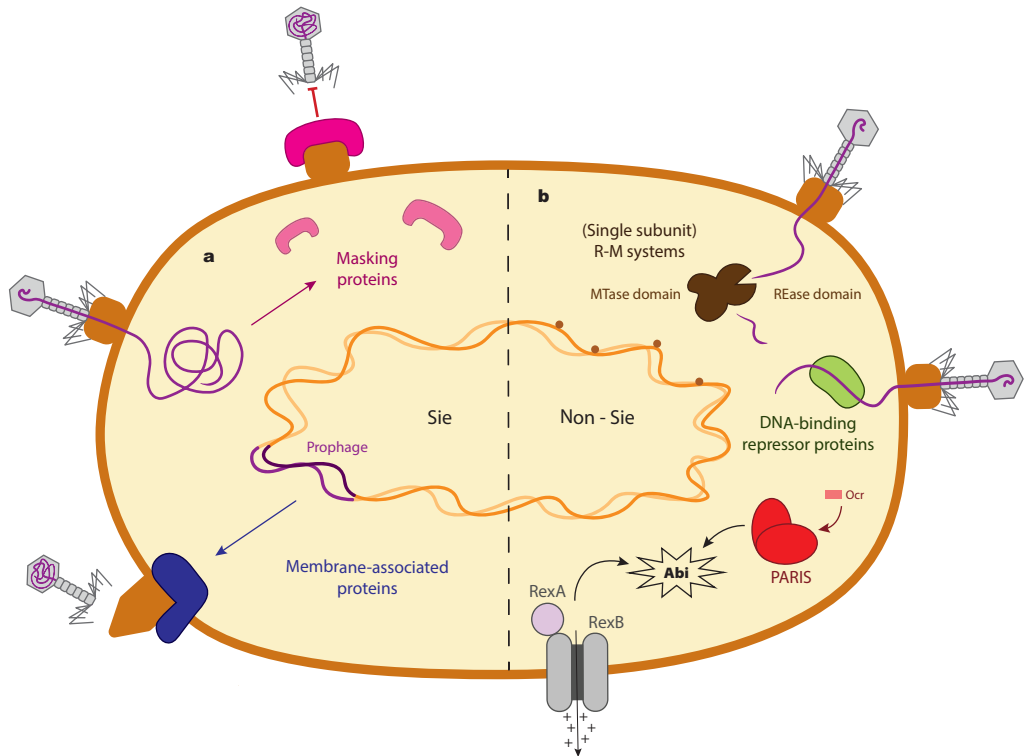


Figure 3. Phage-derived defense systems. **a)** Superinfection exclusion systems (Sie) are encoded by phages to prevent other phages from infecting their host. Some phages like T5 produce proteins that mask their receptor and make it inaccessible. Other phages, especially prophages, encode membrane-associated proteins that interact with the phage receptor, blocking the DNA entry channel, triggering a conformational change, or inhibiting the invading phage's enzymes. **b)** Prophages like Panchino of *M. smegmatis* can confer resistance to their hosts through the expression of R-M systems or DNA-binding repressor proteins that target the DNA of newly infecting phages. Other prophage-encoded systems, like RexA-RexB or the newly characterized PARIS, can trigger an Abi response upon sensing an invasion by a new phage.

2.3. Phage counter-attack strategies

While the mechanisms of phage resistance exhibited by bacteria seem overwhelmingly varied, phages have also developed a broad array of opposing strategies. Just as bacterial defenses target every step in the process of phage infection, every barrier imposed by bacteria has to withstand a phage counterattack ¹³⁷.

In response to variations in the bacterial cell surface receptors, phages are able to change their tropism through mutations in their RBPs. In fact, genes encoding RBPs and other proteins related to host recognition are reported to incorporate mutations at a very high frequency. This is often mediated by the activity of diversity-generating retroelements (DGRs) ¹³⁸. These are regions which are subjected to targeted mutation by means of the exchange of two variable repeats by an error-prone reverse transcriptase ¹³⁹. This type of directed mutagenesis is template dependent and affects determined adenine-specific sites, while a conserved scaffold sequence is retained to ensure stability. This process was first described for the specificity switch of the major tropism determinant protein in *Bordetella* spp. phages. Since then, more phages have been identified which benefit from these systems ¹⁴⁰.

To overcome the barrier imposed by capsules and extracellular layers, some phages became able to bind to these structures ¹⁸, and to degrade them using depolymerases. These enzymes may be either expressed as part of tail spike or tail fiber proteins or released in a soluble form following lysis of infected bacteria ¹⁴¹. A recent review offers an overview of the diversity of phage depolymerases ¹⁴².

Phages have also developed forms of escaping targeting by R-M systems. They can (i) mutate to remove restriction sites from their genome (palindrome avoidance) and therefore avoid recognition by REases ^{143,144}; (ii) modify the sequences recognized by REases (e.g. the glucosyl-hydroxymethylcytosine of T4 that is used instead of the regular cytosine ¹³⁰); (iii) change the distance and orientation of restriction sites to avoid restriction by REases that need to recognize two sequences at a determined distance from each other and in a

specific orientation ¹⁴⁵; (iv) occlude the restriction sites with proteins (e.g. DarA and DarB of P1 phages) that are ejected together with the phage genome ⁸; (v) sequester REases with proteins that mimic the structure of a DNA double helix (e.g. Ocr from T7) ¹⁴⁶; and (vi) acquire genes encoding an MTase that modifies the phage genome ¹⁴⁷, or stimulate the activity of the host MTase for the same purpose ¹⁴⁸.

CRISPR-Cas systems can also be evaded by phages in multiple ways¹⁴⁹. Phages can acquire point mutations or deletions in the PAM sequences or in positions of the protospacer region close to the PAM sequences (i.e. the seed region of the protospacer) ¹⁵⁰. Alternatively, some phages use anti-CRISPR (Acr) proteins, first described in phages of *P. aeruginosa* ¹⁵¹. In general, Acrs work by either preventing recruitment of the crRNA-Cas complex to the target DNA by binding the complex or occluding the PAM sequence, or by inhibiting the endonuclease domain so that cleavage cannot take place ¹⁵². Glucosylation of phage genetic material has also been shown to protect phages against some CRISPR-Cas systems ¹⁵³. A different strategy is employed by the jumbo *Serratia* phage PCH45 and *Pseudomonas* phage ϕ KZ, which form a protein shell that encloses phage DNA in a nucleus-like compartment, physically shielding it from the CRISPR-Cas complexes ^{12,13}. Of note, this mechanism does not protect the phage from RNA-targeting CRISPR-Cas systems, as the transcribed mRNA is not contained within the nucleus-like compartment during translation.

Abi mechanisms can also be outsmarted by phages. Phages can avoid toxin-antitoxin mechanisms by inhibiting the protease that degrades the antitoxin, or by expressing their own antitoxin analogue ^{10,154,155}. Furthermore, mutations in genes involved in the metabolism of nucleic acids also prove to be effective in avoiding toxin-antitoxin systems of some bacteria like *Lactococcus* spp ¹⁵⁶. Mutations in phage genes encoding peptides that activate Abi-associated enzymes, such as the Lit activator Gol peptide in the major head protein of T4, can also result in hindering of the Abi mechanism ¹⁵⁷. Phages ICP1 that infect *V. cholerae* overcome PLE-mediated Abi by using either a phage-encoded CRISPR-Cas system that targets the PLE genome during infection ¹¹², or an endonuclease

that binds and cleaves the PLE origins of replication ¹⁵⁸. It is expected that phages possess counter-measures against the more newly described defense systems as well. The Ocr protein of phage T7, known to inhibit R-M systems, was recently found to also inactivate the BREX system by binding the methyltransferase BrxX ¹¹. The discovery of other new phage counter-attack strategies is likely just a matter of time, as interest in bacterial defense mechanisms and phage anti-defenses continues to grow.

2.4. Phage resistance mechanisms in a clinical context

The number of phage therapy case studies and clinical trials performed in humans has significantly increased in these past years, as the problem of antibiotic resistance aggravates. The efficacy of phage therapy in these studies is quite variable, ranging from negative outcomes to the resolution of severe infections in human patients (**Table 1, Sup. Table S1**). Interestingly, while phage resistance has been shown to develop quickly *in vitro*, studies in humans have described both the presence ¹⁵⁹ and the absence ¹⁶⁰ of phage resistance *in vivo*. As a result of such variable results, there is a lack of consensus in the scientific and medical community about the potential of phages as therapeutic agents.

The human immune response to bacterial infection and the specific phage-resistance mechanisms developed by the bacteria are likely behind the distinct outcomes. The development of an immune response, particularly involving neutrophils, has been shown essential for the success of phage therapy by preventing the outgrowth of phage-resistant mutants ¹⁶¹. Multiple studies have demonstrated that phage-resistant phenotypes often associate with decreased pathogenicity, with the strain becoming more susceptible to the human immune defenses ¹⁶². Receptor adaptations such as mutations in bacterial capsule, LPS, and other surface components are examples of phage-resistance mechanisms that result in increased immune susceptibility ¹⁶³. Importantly, these surface modifications also often associate with increased antibiotic susceptibility. This effect occurs, for example, in cases where the phage interacts with bacterial structures that function as drug efflux pumps ¹⁶⁴. By mutating the

efflux pump to achieve phage resistance, bacteria lose the ability to pump out the antibiotics, thus gaining antibiotic susceptibility as a trade-off (for a review of mechanisms of phage-antibiotic synergism see, for example, ¹⁶⁵). Such interactions have been exploited in therapeutic contexts ^{164,166–168}. However, phage-resistance mutations have also been shown to pleiotropically confer increased antibiotic resistance ⁴⁶, and other mechanisms of phage resistance (e.g. CRISPR-Cas, R-M systems) may lead to a phage-resistance phenotype that does not render the bacteria more susceptible to the immune system or to antibiotics. In such cases, resistance to phages may develop *in vivo* even in the presence of a strong immune response.

Unfortunately, the mechanisms underlying phage resistance are seldom, if ever, investigated in human clinical studies and trials, and represent a clear knowledge gap. Most studies look at the safety and/or clinical outcome of phage therapy, and very few have documented the development of phage resistance (**Table 1**), let alone the mechanisms behind it. There are studies, however, that indicate that addressing and tackling phage resistance can lead to improved treatment outcomes. One such study found phage-resistant clones in a patient suffering from a multidrug resistant *A. baumannii* infection after eight days of treatment with intravenous phage therapy ¹⁶⁹. Phage-resistance was associated with loss of bacterial capsule and increased extracellular polysaccharide production, and was overcome via an iterative process of phage cocktail formulation that resulted in the resolution of the infection. Of relevance, the phage-resistant phenotype was associated with increased antibiotic sensitivity, suggesting a fitness cost of phage-resistant mutations *in vivo*. In another study, the association between phage-resistance and increased antibiotic susceptibility was exploited to treat a patient with a chronic multidrug resistant *P. aeruginosa* infection of an aortic graft ¹⁶⁴. The treatment consisted of a combination of the antibiotic ceftazidime and phage OMKO1, which binds to an outer membrane protein that is part of multidrug efflux systems of *P. aeruginosa*. This combination explored the capacity of the phage to kill the original strain and the

ability of ceftazidime to kill any emerging phage-resistant variants with mutations in the multidrug efflux system, to achieve resolution of the infection.

Characterizing the mechanisms of resistance to phages that target pathogens of interest will inform about the relevance that each phage defense system has in a clinical context, in terms of frequency with which they occur in pathogens and their association with virulence and antibiotic susceptibility of the pathogen. Furthermore, certain natively present defense systems like R-M may affect and reduce the choice of phages available to use in a therapeutic setting. Another issue to consider is that the development of resistance (as well as treatment efficacy) may significantly differ when using single- or multi-phage treatment approaches, and may also vary with the timing and order of phage administration ¹⁷⁰. The more widespread use of bacteriophages for therapeutic purposes could lead to selection for phage resistant phenotypes that arise through horizontal gene transfer of phage defense systems. Understanding the complexity of interactions and mechanisms leading to phage resistance will aid the development of phage-based treatments with better clinical outcomes and to engineered phages that may overcome host defense systems.

Table 1. Case studies and clinical trials of phage therapy in humans, with associated safety, clinical, and phage-resistance outcomes.

Clinical study/case report ^a	Safety	Clinical outcome ^b	Phage resistance ^c	Mechanism of resistance	Ref.
Tibia bone infection with MDR <i>A. baumannii</i> and <i>Klebsiella pneumoniae</i> treated with phages and antibiotics	Well tolerated	Successful	Not found in the patient; found in <i>in vitro</i> experiments with isolates	Mutations in surface adhesin and glycosyl-transferase of the EpsG family (speculated)	171
Necrotizing pancreatitis patient with an MDR <i>A. baumannii</i> infection treated with phages and antibiotics	Well tolerated	Successful	A resistant bacterial isolate was found and used to select a new phage	Loss of bacterial capsule and increased extracellular polysaccharide production	169
Cystic fibrosis patient with MDR <i>P. aeruginosa</i> infection treated with phages and antibiotics	Well tolerated	Successful	One resistant isolate was identified	Not described	172
Patient with <i>K. pneumoniae</i> urinary tract infection treated with non-active antibiotics and phages	Well tolerated	Successful	Bacteria became resistant to two rounds of phage cocktails but were ultimately sensitive to the combination of a third cocktail and previously inactive antibiotics	Not yet elucidated	173
Patient with periprosthetic joint infection and MDR <i>P. aeruginosa</i> chronic osteomyelitis treated with phages and antibiotics	Well tolerated	Successful	Not identified	Not applicable	174
Patient with MDR <i>P. aeruginosa</i> urinary tract infection treated with phages and antibiotics	Well tolerated	Successful	Not identified	Not applicable	160
Aortic graft infection with <i>P. aeruginosa</i> treated with a phage targeting an efflux pump protein and antibiotics	Well tolerated	Clinical improvement	Yes, as expected. Explored to cause sensitization of bacteria to antibiotic.	Mutations in receptor protein (M of mexAB- and mexXY-efflux system)	164
Cystic fibrosis patient with <i>Mycobacterium abscessus</i> and <i>P. aeruginosa</i> infection treated with engineered phages	Well tolerated	Clinical improvement	Resistance detected <i>in vitro</i> for two of the three phages administered	Not described	175

Netherton syndrome patient with MDR <i>Staphylococcus aureus</i> infection and allergy to multiple antibiotics treated with phage cocktails topically and orally	Well tolerated	Clinical improvement	Resistance to a phage cocktail was identified after 3 months of treatment	Not described	159
Thirteen patients with <i>S. aureus</i> bacteremia treated with phages and antibiotics	Well tolerated	Clinical improvement in eight of the patients	Changes in phage susceptibility were detected <i>in vivo</i> in three patients	Single nucleotide polymorphisms were detected in isolates recovered from one of the patients	176
Three lung transplant patients with MDR <i>P. aeruginosa</i> and <i>Burkholderia dolosa</i> infections treated with phages and antibiotics	Well tolerated	Clinical improvement in two of three patients	Resistant isolates were identified post-therapy in one of the patients who had a successful outcome	Not described	177
Patient with bronchiectasis and <i>M. abscessus</i> infection treated with a phage cocktail and antibiotics	Well tolerated	Failure due to anti-phage immune response	Resistant isolates appeared to one out of three phages in the cocktail	Not described	178
Eight cardiothoracic surgery patients treated with antibiotics, phages and fibrin glue	Well tolerated except for an increase in inflammation	Successful in five of the patients; two patients showed improvement but died of unrelated complications; one patient did not experience sufficient bacterial clearance and died of sepsis	Bacterial isolates appeared to have mutated in one of the patients	Not described	179
Ten patients with diverse MDR bacterial infections treated with intravenous phage cocktails and antibiotics	Well tolerated except by one patient, who developed fever, wheezing and shortness of breath after first dose of the cocktail (subsequent doses were well tolerated)	Successful for seven out of ten patients	Resistance developed in three of the patients, but was overcome through administration of phages specific for the resistant isolates (two of these cases still had a successful outcome)	Not described	180

Phase I-II clinical trial to assess the safety and efficacy of a phage cocktail to treat burn wounds infected with <i>P. aeruginosa</i> in 26 patients (13 treated, 13 control)	Fewer adverse events were observed in the phage-treated group than in the group treated with the standard of care	Trial was stopped because of insufficient efficacy (patients were being exposed to 10,000-fold lower doses of phages than originally intended)	Intermediately susceptible and resistant isolates were found	Not described	181
Report of phage therapy performed on 153 patients with different infections	Diverse	Diverse	Resistance and changes in the phage typing profile were detected in multiple cases	Not described	182

^a Only clinical studies/case reports for which phage resistance was investigated are shown. For a full list of clinical studies and case reports with phages, see **Sup. Table S1**.

^b Successful: patient was healed. Clinical improvement: infection and/or associated complications were reduced but not completely resolved.

^c Not described: resistance was not addressed. Not identified: resistance was investigated but not found.

2.5. Concluding remarks

2 In this review, we have provided an overview of the current knowledge of mechanisms behind existing and developing phage resistance, and highlighted the potential knowledge gaps and clinical importance of phage resistance for phage therapeutic strategies. While our understanding of the mechanisms behind phage resistance has expanded in recent years, many defense systems remain uncharacterized or yet undiscovered. As such, the complete picture of phage resistance development remains elusive, especially in the context of the human body.

For phage therapy to move forward, it is imperative that clinical studies and trials also assess the development of resistance in a systematic manner, in which both the emergence of phage resistance and the mechanisms behind it are included in the investigation. Sequencing technologies and genome analysis of both bacterial strains and phages may allow for the identification of defense and anti-defense systems in clinical isolates. Such data will prove invaluable for isolating and selecting candidate phages, as well as for predicting the outcome of the therapeutic intervention.

Improved understanding of how defense systems affect phage therapy, combined with an increased knowledge of the anti-defense strategies employed by phages to counteract bacterial defenses, will greatly contribute to the development of more effective phage-based therapeutic approaches.

2.6. References

1. Twort, F.W. (1915) An investigation on the nature of ultra-microscopic viruses. *Lancet*, 186.
2. d'Herelle, F. (1917) Sur un microbe invisible antagoniste des bacilles dysentériques. *Comptes rendus Acad Sci Paris*, 165.
3. Wittebole, X., De Roock, S. and Opal, S.M. (2013) A historical overview of bacteriophage therapy as an alternative to antibiotics for the treatment of bacterial pathogens. *Virulence*, 5.
4. Chanishvili, N. (2012) Phage Therapy—History from Twort and d'Herelle Through Soviet Experience to Current Approaches. *Adv. Virus Res.*, 83.
5. Moelling, K., Broecker, F. and Willy, C. (2018) A Wake-Up Call: We Need Phage Therapy Now. *Viruses*, 10.
6. Doron, S., Melamed, S., Ofir, G., Leavitt, A., Lopatina, A., Keren, M., Amitai, G. and Sorek, R. (2018) Systematic discovery of antiphage defense systems in the microbial pangenome. *Science*, 4120.
7. Gao, L., Altae-Tran, H., Böhning, F., Makarova, K.S., Segel, M., Schmid-Burgk, J.L., Koob, J., Wolf, Y.I., Koonin, E. V. and Zhang, F. (2020) Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science*, 1084.
8. S, I., MB, S., TA, B. and W, A. (1987) Two DNA antirestriction systems of bacteriophage P1, darA, and darB: characterization of darA phages. *Virology*, 157.
9. C, A., TJ, S., SS, S. and DT, D. (2002) Interaction of the ocr gene 0.3 protein of bacteriophage T7 with EcoKI restriction/modification enzyme. *Nucleic Acids Res.*, 30.
10. Otsuka, Y. and Yonesaki, T. (2012) Dmd of bacteriophage T4 functions as an antitoxin against Escherichia coli LsoA and RnlA toxins. *Mol. Microbiol.*, 83.
11. Isaev, A., Drobiazko, A., Sierro, N., Gordeeva, J., Qimron, U., Ivanov, N. V., Severinov, K. and Yosef, I. (2020) Phage T7 DNA mimic protein Ocr is a potent inhibitor of BREX defence. *Nucleic Acids Res.*, gkaa290.
12. Malone, L.M., Warring, S.L., Jackson, S.A., Warnecke, C., Gardner, P.P., Gumy, L.F. and Fineran, P.C. (2020) A jumbo phage that forms a nucleus-like structure evades CRISPR–Cas DNA targeting but is vulnerable to type III RNA-based immunity. *Nat. Microbiol.*, 5.
13. Mendoza, S.D., Nieweglowska, E.S., Govindarajan, S., Leon, L.M., Berry, J.D., Tiwari, A., Chaikeeratisak, V., Pogliano, J., Agard, D.A. and Bondy-Denomy, J. (2020) A bacteriophage nucleus-like compartment shields DNA from CRISPR nucleases. *Nature*, 577.
14. Wiegand, T., Karambelkar, S., Bondy-Denomy, J. and Wiedenheft, B. (2020) Structures and Strategies of Anti-CRISPR-Mediated Immune Suppression. *Annu. Rev. Microbiol.*, 74, 21–37.
15. Roach, D.R. and Debarbieux, L. (2017) Phage therapy: awakening a sleeping giant. *Emerg. Top. Life Sci.*, 1.
16. Fields, B.N., Knipe, D.M. (David M. and Howley, P.M. (1996) *Virology* 3rd. Lippincott-Raven Publishers.

17. King,A. (2012) Caudovirales. In *Virus Taxonomy*. Elsevier.
18. Bertozzi Silva,J., Storms,Z. and Sauvageau,D. (2016) Host receptors for bacteriophage adsorption. *FEMS Microbiol. Lett.*, 363, fnw002.
19. Nobrega,F.L., Vlot,M., Jonge,P.A., Dreesens,L.L., Beaumont,H.J.E., Lavigne,R., Dutilh,B.E. and Brouns,S.J.J. (2018) Targeting mechanisms of tailed bacteriophages. *Nat. Rev. Microbiol.*, 16.
20. Jonge,P.A. De, Nobrega,F.L., Brouns,S.J.J. and Dutilh,B.E. (2018) Molecular and Evolutionary Determinants of Bacteriophage Host Range. *Trends Microbiol.*
21. González-García,V.A., Pulido-Cid,M., García-Doval,C., Bocanegra,R., Van Raaij,M.J., Martín-Benito,J., Cuervo,A. and Carrascosa,J.L. (2015) Conformational changes leading to T7 DNA delivery upon interaction with the bacterial receptor. *J. Biol. Chem.*, 290.
22. Wang,C., Tu,J., Liu,J. and Molineux,I.J. (2019) Structural dynamics of bacteriophage P22 infection initiation revealed by cryo-electron tomography. *Nat. Microbiol.*, 4.
23. Boulanger,P., Jacquot,P., Plançon,L., Chami,M., Engel,A., Parquet,C., Herbeuval,C. and Letellier,L. (2008) Phage T5 straight tail fiber is a multifunctional protein acting as a tape measure and carrying fusogenic and muralytic activities. *J. Biol. Chem.*, 283.
24. Cumby,N., Reimer,K., Mengin-Lecreulx,D., Davidson,A.R. and Maxwell,K.L. (2015) The phage tail tape measure protein, an inner membrane protein and a periplasmic chaperone play connected roles in the genome injection process of E.coli phage HK97. *Mol. Microbiol.*, 96.
25. Taylor,N.M.I., Prokhorov,N.S., Guerrero-Ferreira,R.C., Shneider,M.M., Browning,C., Goldie,K.N., Stahlberg,H. and Leiman,P.G. (2016) Structure of the T4 baseplate and its function in triggering sheath contraction. *Nature*, 533.
26. Leptihn,S., Gottschalk,J. and Kuhn,A. (2016) T7 ejectosome assembly: A story unfolds. *Bacteriophage*, 6, e1128513.
27. Molineux,I.J. and Panja,D. (2013) Popping the cork: Mechanisms of phage genome ejection. *Nat. Rev. Microbiol.*, 11.
28. Smith,D.E., Tans,S.J., Smith,S.B., Grimes,S., Anderson,D.L. and Bustamante,C. (2001) The bacteriophage ϕ 29 portal motor can package DNA against a large internal force. *Nature*, 413.
29. Kemp,P., Gupta,M. and Molineux,I.J. (2004) Bacteriophage T7 DNA ejection into cells is initiated by an enzyme-like mechanism. *Mol. Microbiol.*, 53.
30. Choi,K.H., McPartland,J., Kaganman,I., Bowman,V.D., Rothman-Denes,L.B. and Rossmann,M.G. (2008) Insight into DNA and Protein Transport in Double-Stranded DNA Viruses: The Structure of Bacteriophage N4. *J. Mol. Biol.*, 378.
31. Panja,D. and Molineux,I.J. (2010) Dynamics of bacteriophage genome ejection *in vitro* and *in vivo*. *Phys. Biol.*, 7.
32. Drulis-Kawa,Z., Majkowska-Skrobek,G., Maciejewska,B., Delattre,A.-S. and Lavigne,R. (2013) Learning from Bacteriophages - Advantages and Limitations of Phage and Phage-Encoded Protein Applications. *Curr. Protein Pept. Sci.*, 13.

33. Petrovic Fabijan,A., Khalid,A., Maddocks,S., Ho,J., Gilbey,T., Sandaradura,I., Lin,R.C., Ben Zakour,N., Venturini,C., Bowring,B., *et al.* (2020) Phage therapy for severe bacterial infections: a narrative review. *Med. J. Aust.*, 212.
34. De Smet,J., Hendrix,H., Blasdel,B.G., Danis-Wlodarczyk,K. and Lavigne,R. (2017) *Pseudomonas* predators: Understanding and exploiting phage-host interactions. *Nat. Rev. Microbiol.*, 15.
35. Hinton,D.M. (2010) Transcriptional control in the prereplicative phase of T4 development. *Virology*, 71.
36. Drobysheva,A. V., Panafidina,S.A., Kolesnik,M. V., Klimuk,E.I., Minakhin,L., Yakunina,M. V., Borukhov,S., Nilsson,E., Holmfeldt,K., Yutin,N., *et al.* (2021) Structure and function of virion RNA polymerase of a crAss-like phage. *Nature*, 589.
37. Hobbs,Z. and Abedon,S.T. (2016) Diversity of phage infection types and associated terminology: the problem with ‘Lytic or lysogenic’. *FEMS Microbiol. Lett.*, 363, fnw047.
38. Erez,Z., Steinberger-levy,I., Shamir,M., Doron,S., Stokar-avihail,A., Peleg,Y., Melamed,S., Leavitt,A., Savidor,A., Albeck,S., *et al.* (2017) Communication between viruses guides lysis – lysogeny decisions. *Nature*, 541.
39. Silpe,J.E. and Bassler,B.L. (2019) A Host-Produced Quorum-Sensing Autoinducer Controls a Phage Lysis-Lysogeny Decision. *Cell*, 176.
40. Takeuchi,N., Wolf,Y.I., Makarova,K.S. and Koonin,E. V. (2012) Nature and intensity of selection pressure on crispr-associated genes. *J. Bacteriol.*, 194.
41. Puigbò,P., Lobkovsky,A.E., Kristensen,D.M., Wolf,Y.I. and Koonin,E. V. (2014) Genomes in turmoil: Quantification of genome dynamics in prokaryote supergenomes. *BMC Med.*, 12.
42. Hampton,H.G., Watson,B.N.J. and Fineran,P.C. (2020) The arms race between bacteria and their phage foes. *Nature*, 577.
43. Kortright,K.E., Chan,B.K. and Turner,P.E. (2020) High-throughput discovery of phage receptors using transposon insertion sequencing of bacteria. *Proc. Natl. Acad. Sci. U. S. A.*, 117.
44. Chapman-McQuiston,E. and Wu,X.L. (2008) Stochastic receptor expression allows sensitive bacteria to evade phage attack. Part I: Experiments. *Biophys. J.*, 94.
45. Chapman-McQuiston,E. and Wu,X.L. (2008) Stochastic receptor expression allows sensitive bacteria to evade phage attack. Part II: Theoretical analyses. *Biophys. J.*, 94.
46. Burmeister,A.R., Fortier,A., Roush,C., Lessing,A.J., Bender,R.G., Barahman,R., Grant,R., Chan,B.K. and Turner,P.E. (2020) Pleiotropy complicates a trade-off between phage resistance and antibiotic resistance. *Proc. Natl. Acad. Sci. U. S. A.*, 117.
47. Gordillo Altamirano,F., Forsyth,J.H., Patwa,R., Kostoulas,X., Trim,M., Subedi,D., Archer,S.K., Morris,F.C., Oliveira,C., Kielty,L., *et al.* (2021) Bacteriophage-resistant *Acinetobacter baumannii* are resensitized to antimicrobials. *Nat. Microbiol.*, 6.
48. Trudelle,D.M., Bryan,D.W., Hudson,L.K. and Denes,T.G. (2019) Cross-resistance to phage infection in *Listeria monocytogenes* serotype 1/2a mutants. *Food Microbiol.*, 84.
49. Eugster,M.R., Morax,L.S., Hüls,V.J., Huwiler,S.G., Leclercq,A., Lecuit,M. and Loessner,M.J. (2015) Bacteriophage predation promotes serovar diversification in *Listeria monocytogenes*. *Mol. Microbiol.*, 97.

50. Duerkop,B.A., Huo,W., Bhardwaj,P., Palmer,K.L. and Hooper,L. V. (2016) Molecular Basis for Lytic Bacteriophage Resistance in Enterococci. *MBio*, 7.
51. Riede,I. and Eschbach,M.L. (1986) Evidence that TraT interacts with OmpA of Escherichia coli. *FEBS Lett.*, 205.
52. Decker,K., Krauel,V., Meesmann,A. and Heller,K.J. (1994) Lytic conversion of Escherichia coli by bacteriophage T5: blocking of the FhuA receptor protein by a lipoprotein expressed early during infection. *Mol. Microbiol.*, 12.
53. Manning,A.J. and Kuehn,M.J. (2011) Contribution of bacterial outer membrane vesicles to innate bacterial defense. *BMC Microbiol.*, 11.
54. Hanlon,G.W., Denyer,S.P., Olliff,C.J. and Ibrahim,L.J. (2001) Reduction in Exopolysaccharide Viscosity as an Aid to Bacteriophage Penetration through Pseudomonas aeruginosa Biofilms. *Appl. Environ. Microbiol.*, 67.
55. Testa,S., Berger,S., Piccardi,P., Oechslin,F., Resch,G. and Mitri,S. (2019) Spatial structure affects phage efficacy in infecting dual-strain biofilms of Pseudomonas aeruginosa. *Commun. Biol.*, 2.
56. Forde,A. and Fitzgerald,G.F. (2003) Molecular organization of exopolysaccharide (EPS) encoding genes on the lactococcal bacteriophage adsorption blocking plasmid, pCI658. *Plasmid*, 49.
57. Gencay,Y.E., Sørensen,M.C.H., Wenzel,C.Q., Szymanski,C.M. and Brøndsted,L. (2018) Phase Variable Expression of a Single Phage Receptor in Campylobacter jejuni NCTC12662 Influences Sensitivity Toward Several Diverse CPS-Dependent Phages. *Front. Microbiol.*, 9.
58. Dybvig,K. (1993) DNA rearrangements and phenotypic switching in prokaryotes. *Mol. Microbiol.*, 10.
59. Heichman,K.A. and Johnson,R.C. (1990) The Hin invertasome: protein-mediated joining of distant recombination sites at the enhancer. *Science*, 249.
60. Abraham,J.M., Freitag,C.S., Clements,J.R. and Eisenstein,B.I. (1985) An invertible element of DNA controls phase variation of type 1 fimbriae of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.*, 82.
61. Choi,Y., Shin,H., Lee,J.-H. and Ryu,S. (2013) Identification and characterization of a novel flagellum-dependent Salmonella-infecting bacteriophage, iEPS5. *Appl. Environ. Microbiol.*, 79.
62. Zhou,K., Aertsen,A. and Michiels,C.W. (2014) The role of variable DNA tandem repeats in bacterial adaptation. *FEMS Microbiol. Rev.*, 38.
63. Willems,R., Paul,A., van der Heide,H.G., ter Avest,A.R. and Mooi,F.R. (1990) Fimbrial phase variation in Bordetella pertussis: a novel mechanism for transcriptional regulation. *EMBO J.*, 9.
64. Sarkari,J., Pandit,N., Moxon,E.R. and Achtman,M. (1994) Variable expression of the Opc outer membrane protein in Neisseria meningitidis is caused by size variation of a promoter containing poly-cytidine. *Mol. Microbiol.*, 13.
65. Casadesús,J. and Low,D. (2006) Epigenetic gene regulation in the bacterial world. *Microbiol. Mol. Biol. Rev.*, 70.

66. Cota,I., Sánchez-Romero,M.A., Hernández,S.B., Pucciarelli,M.G., García-Del Portillo,F. and Casadesús,J. (2015) Epigenetic Control of Salmonella enterica O-Antigen Chain Length: A Tradeoff between Virulence and Bacteriophage Resistance. *PLoS Genet.*, 11.
67. Mangalea,M.R. and Duerkop,B.A. (2020) Fitness Trade-Offs Resulting from Bacteriophage Resistance Potentiate Synergistic Antibacterial Strategies. *Infect. Immun.*
68. Alseth,E.O., Pursey,E., Luján,A.M., McLeod,I., Rollie,C. and Westra,E.R. (2019) Bacterial biodiversity drives the evolution of CRISPR-based phage resistance. *Nature*, 574.
69. Broniewski,J.M., Meaden,S., Paterson,S., Buckling,A. and Westra,E.R. (2020) The effect of phage genetic diversity on bacterial resistance evolution. *ISME J.*, 14.
70. Koonin,E. V., Makarova,K.S. and Wolf,Y.I. (2017) Evolutionary Genomics of Defense Systems in Archaea and Bacteria. *Annu. Rev. Microbiol.*, 71.
71. Luria,S.E. and Human,M.L. (1952) A nonhereditary, host-induced variation of bacterial viruses. *J. Bacteriol.*, 64.
72. Luria,S.E. (1953) Host-induced modifications of viruses. *Cold Spring Harb. Symp. Quant. Biol.*, 18.
73. Tock,M.R. and Dryden,D.T.F. (2005) The biology of restriction and anti-restriction. *Curr. Opin. Microbiol.*, 8.
74. Loenen,W.A.M. and Raleigh,E.A. (2014) The other face of restriction: Modification-dependent enzymes. *Nucleic Acids Res.*, 42.
75. Stewart,F.J., Panne,D., Bickle,T.A. and Raleigh,E.A. (2000) Methyl-specific DNA binding by McrBC, a modification-dependent restriction enzyme. *J. Mol. Biol.*, 298.
76. Gupta,R., Capalash,N. and Sharma,P. (2012) Restriction endonucleases: Natural and directed evolution. *Appl. Microbiol. Biotechnol.*, 94.
77. Xu,T., Yao,F., Zhou,X., Deng,Z. and You,D. (2010) A novel host-specific restriction system associated with DNA backbone S-modification in Salmonella. *Nucleic Acids Res.*, 38.
78. Xiong,X., Wu,G., Wei,Y., Liu,L., Zhang,Y.Y., Su,R., Jiang,X., Li,M., Gao,H., Tian,X., *et al.* (2020) SspABCD–SspE is a phosphorothioation-sensing bacterial defence system with broad anti-phage activities. *Nat. Microbiol.*
79. Wang,S., Wan,M., Huang,R., Zhang,Y., Xie,Y., Wei,Y., Ahmad,M., Wu,D., Hong,Y., Deng,Z., *et al.* (2021) SspABCD-SspFGH Constitutes a New Type of DNA Phosphorothioate-Based Bacterial Defense System. *MBio*, 12.
80. Ofir,G., Melamed,S., Sberro,H., Mukamel,Z., Silverman,S., Yaakov,G., Doron,S. and Sorek,R. (2018) DISARM is a widespread bacterial defence system with broad anti-phage activities. *Nat. Microbiol.*, 3.
81. Goldfarb,T., Sberro,H., Weinstock,E., Cohen,O., Doron,S., Charpak-Amikam,Y., Afik,S., Ofir,G., Sorek,R., Charpak-Amikam,Y., *et al.* (2015) BREX is a novel phage resistance system widespread in microbial genomes. *EMBO J.*, 34.
82. Gordeeva,J., Morozova,N., Sierro,N., Isaev,A., Sinkunas,T., Tsvetkova,K., Matlashov,M., Truncaite,L., Morgan,R.D., Ivanov,N. V., *et al.* (2019) BREX system of

- Escherichia coli* distinguishes self from non-self by methylation of a specific DNA site. *Nucleic Acids Res.*, 47.
83. Willkomm,S., Makarova,K.S. and Grohmann,D. (2018) DNA silencing by prokaryotic Argonaute proteins adds a new layer of defense against invading nucleic acids. *FEMS Microbiol. Rev.*, 42.
 84. Wu,J., Yang,J., Cho,W.C. and Zheng,Y. (2020) Argonaute proteins: Structural features, functions and emerging roles. *J. Adv. Res.*, 24.
 85. Swarts,D.C., Jore,M.M., Westra,E.R., Zhu,Y., Janssen,J.H., Snijders,A.P., Wang,Y., Patel,D.J., Berenguer,J., Brouns,S.J.J., *et al.* (2014) DNA-guided DNA interference by a prokaryotic Argonaute. *Nature*, 507.
 86. Miyoshi,T., Ito,K., Murakami,R. and Uchiumi,T. (2016) Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. *Nat. Commun.*, 7.
 87. Mojica,F.J.M., Díez-Villaseñor,C., García-Martínez,J. and Soria,E. (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.*, 60.
 88. Brouns,S.J.J., Jore,M.M., Lundgren,M., Westra,E.R., Slikhuis,R.J.H., Snijders,A.P.L., Dickman,M.J., Makarova,K.S., Koonin,E. V. and Van Der Oost,J. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, 321.
 89. Millen,A.M., Horvath,P., Boyaval,P. and Romero,D.A. (2012) Mobile CRISPR/Cas-Mediated Bacteriophage Resistance in *Lactococcus lactis*. *PLoS One*, 7.
 90. Koonin,E. V. and Makarova,K.S. (2019) Origins and evolution of CRISPR-Cas systems. *Philos. Trans. R. Soc. B Biol. Sci.*, 374.
 91. Hille,F., Richter,H., Wong,S.P., Bratovič,M., Ressel,S. and Charpentier,E. (2018) The Biology of CRISPR-Cas: Backward and Forward. *Cell*, 172.
 92. Jackson,S.A., McKenzie,R.E., Fagerlund,R.D., Kieper,S.N., Fineran,P.C. and Brouns,S.J.J. (2017) CRISPR-Cas: Adapting to change. *Science*, 356.
 93. McGinn,J. and Marraffini,L.A. (2019) Molecular mechanisms of CRISPR-Cas spacer acquisition. *Nat. Rev. Microbiol.*, 17.
 94. Al-Attar,S., Westra,E.R., van der Oost,J. and Brouns,S.J.J. (2011) Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. *Biol. Chem.*, 392.
 95. Gleditzsch,D., Pausch,P., Müller-Esparza,H., Özcan,A., Guo,X., Bange,G. and Randau,L. (2018) PAM identification by CRISPR-Cas effector complexes: diversified mechanisms and structures. *RNA Biol.*
 96. Makarova,K.S., Wolf,Y.I., Iranzo,J., Shmakov,S.A., Alkhnbashi,O.S., Brouns,S.J.J., Charpentier,E., Cheng,D., Haft,D.H., Horvath,P., *et al.* (2020) Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.*, 18.
 97. Doudna,J.A. and Charpentier,E. (2014) The new frontier of genome engineering with CRISPR-Cas9. *Science*, 346.
 98. Lopatina,A., Tal,N. and Sorek,R. (2020) Abortive Infection: Bacterial Suicide as an Antiviral Immune Strategy. *Annu. Rev. Virol.*, 7.

99. Page,R. and Peti,W. (2016) Toxin-antitoxin systems in bacterial growth arrest and persistence. *Nat. Chem. Biol.*, 12.
100. Fineran,P.C., Blower,T.R., Foulds,I.J., Humphreys,D.P., Lilley,K.S. and Salmond,G.P.C. (2009) The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. *Proc. Natl. Acad. Sci. U. S. A.*, 106.
101. Dy,R.L., Przybilski,R., Semeijn,K., Salmond,G.P.C. and Fineran,P.C. (2014) A widespread bacteriophage abortive infection system functions through a Type IV toxin-antitoxin mechanism. *Nucleic Acids Res.*, 42.
102. Snyder,L. (1995) Phage-exclusion enzymes: a bonanza of biochemical and cell biology reagents? *Mol. Microbiol.*, 15.
103. Levitz,R., Chapman,D., Amitsur,M., Green,R., Snyder,L. and Kaufmann,G. (1990) The optional *E. coli* prr locus encodes a latent form of phage T4-induced anticodon nuclease. *EMBO J.*, 9.
104. Cheng,X., Wang,W. and Molineux,I.J. (2004) F exclusion of bacteriophage T7 occurs at the cell membrane. *Virology*, 326.
105. Chopin,M.-C., Chopin,A. and Bidnenko,E. (2005) Phage abortive infection in lactococci: variations on a theme. *Curr. Opin. Microbiol.*, 8.
106. Mills,S., McAuliffe,O.E., Coffey,A., Fitzgerald,G.F. and Ross,R.P. (2006) Plasmids of lactococci-genetic accessories or genetic necessities? *FEMS Microbiol. Rev.*, 30.
107. Sun,L., Wu,J., Du,F., Chen,X. and Chen,Z.J. (2013) Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science*, 339.
108. Cohen,D., Melamed,S., Millman,A., Shulman,G., Oppenheimer-Shaanan,Y., Kacen,A., Doron,S., Amitai,G. and Sorek,R. (2019) Cyclic GMP-AMP signalling protects bacteria against viral infection. *Nature*, 574.
109. Millman,A., Melamed,S., Amitai,G. and Sorek,R. (2020) Diversity and classification of cyclic-oligonucleotide-based anti-phage signalling systems. *Nat. Microbiol.*, 5.
110. Labrie,S.J., Mosterd,C., Loignon,S., Dupuis,M.È., Desjardins,P., Rousseau,G.M., Tremblay,D.M., Romero,D.A., Horvath,P., Fremaux,C., *et al.* (2019) A mutation in the methionine aminopeptidase gene provides phage resistance in *Streptococcus thermophilus*. *Sci. Rep.*, 9.
111. McKitterick,A.C., LeGault,K.N., Angermeyer,A., Alam,M. and Seed,K.D. (2019) Competition between mobile genetic elements drives optimization of a phage-encoded CRISPR-Cas system: Insights from a natural arms race. *Philos. Trans. R. Soc. B Biol. Sci.*, 374.
112. Seed,K.D., Lazinski,D.W., Calderwood,S.B. and Camilli,A. (2013) A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. *Nature*, 494.
113. O'Hara,B.J., Barth,Z.K., McKitterick,A.C. and Seed,K.D. (2017) A highly specific phage defense system is a conserved feature of the *Vibrio cholerae* mobilome. *PLoS Genet.*, 13.
114. Barth,Z.K., Silvas,T. V., Angermeyer,A. and Seed,K.D. (2020) Genome replication dynamics of a bacteriophage and its satellite reveal strategies for parasitism and viral restriction. *Nucleic Acids Res.*, 48.

115. Hays,S.G. and Seed,K.D. (2020) Dominant vibrio cholerae phage exhibits lysis inhibition sensitive to disruption by a defensive phage satellite. *Elife*, 9.
116. Ka,D., Oh,H., Park,E., Kim,J. and Bae,E. (2020) Structural and functional evidence of bacterial antiphage protection by Thoeris defense system via NAD⁺ degradation. *Nat. Commun.*
117. Ofir,G., Herbst,E., Baroz,M., Cohen,D., Millman,A., Doron,S., Tal,N., Malheiro,D.B.A., Malitsky,S., Amitai,G., *et al.* (2021) Antiviral activity of bacterial TIR domains via signaling molecules that trigger cell death. *bioRxiv*.
118. Millman,A., Bernheim,A., Stokar-Avihail,A., Fedorenko,T., Voichek,M., Leavitt,A., Oppenheimer-Shaanan,Y. and Sorek,R. (2020) Bacterial Retrons Function In Anti-Phage Defense. *Cell*, 183.
119. Severin,G.B., Hsueh,B.Y., Elg,C.A., Dover,J.A., Wessel,A.J., Ridenhour,B.J., Top,E.M., Ravi,J. and Waters,C.M. (2021) A Broadly Conserved Deoxycytidine Deaminase Protects Bacteria from Phage Infection 1 2. *bioRxiv*.
120. Tal,N., Morehouse,B.B., Millman,A., Stokar-Avihail,A., Avraham,C., Fedorenko,T., Yirmiya,E., Herbst,E., Brandis,A., Mehlman,T., *et al.* (2021) Cyclic CMP and cyclic UMP mediate bacterial immunity against phages. *Cell*, 184.
121. Athukoralage,J.S. and White,M.F. (2021) Cyclic oligoadenylate signalling and regulation by ring nucleases during type III CRISPR defence. *RNA*.
122. Makarova,K.S., Timinskas,A., Wolf,Y.I., Gussow,A.B., Siksnys,V., Venclovas,Č. and Koonin,E. V. (2020) Evolutionary and functional classification of the CARF domain superfamily, key sensors in prokaryotic antiviral defense. *Nucleic Acids Res.*, 48.
123. Watson,B.N.J., Vercoe,R.B., Salmond,G.P.C., Westra,E.R., Staals,R.H.J. and Fineran,P.C. (2019) Type I-F CRISPR-Cas resistance against virulent phages results in abortive infection and provides population-level immunity. *Nat. Commun.*, 10.
124. Davies,J. (2013) Specialized microbial metabolites: Functions and origins. *J. Antibiot. (Tokyo)*, 66.
125. Kronheim,S., Daniel-Ivad,M., Duan,Z., Hwang,S., Wong,A.I., Mantel,I., Nodwell,J.R. and Maxwell,K.L. (2018) A chemical defence against phage infection. *Nature*.
126. Bernheim,A., Millman,A., Ofir,G., Meitav,G., Avraham,C., Shomar,H., Rosenberg,M.M., Tal,N., Melamed,S., Amitai,G., *et al.* (2021) Prokaryotic viperins produce diverse antiviral molecules. *Nature*, 589.
127. Helbig,K.J. and Beard,M.R. (2014) The role of viperin in the innate antiviral response. *J. Mol. Biol.*, 426.
128. Pedruzzi,I., Rosenbusch,J.P. and Locher,K.P. (1998) Inactivation *in vitro* of the *Escherichia coli* outer membrane protein FhuA by a phage T5-encoded lipoprotein. *FEMS Microbiol. Lett.*, 168.
129. Van Houte,S., Buckling,A. and Westra,E.R. (2016) Evolutionary Ecology of Prokaryotic Immune Mechanisms. *Microbiology and Mol. Biology Rev.*, 80.
130. Labrie,S.J., Samson,J.E. and Moineau,S. (2010) Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.*, 8.

131. Bondy-Denomy, J., Qian, J., Westra, E.R., Buckling, A., Guttman, D.S., Davidson, A.R. and Maxwell, K.L. (2016) Prophages mediate defense against phage infection through diverse mechanisms. *ISME J.*, 10.
132. Parma, D.H., Snyder, M., Sobolevski, S., Nawroz, M., Brody, E. and Gold, L. (1992) The Rex system of bacteriophage lambda: tolerance and altruistic cell death. *Genes Dev.*, 6.
133. Dedrick, R.M., Jacobs-Sera, D., Bustamante, C.A.G., Garlena, R.A., Mavrich, T.N., Pope, W.H., Reyes, J.C.C., Russell, D.A., Adair, T., Alvey, R., *et al.* (2017) Prophage-mediated defence against viral attack and viral counter-defence. *Nat. Microbiol.*, 2.
134. Pope, W.H., Jacobs-Sera, D., Russell, D.A., Peebles, C.L., Al-Atrache, Z., Alcoser, T.A., Alexander, L.M., Alfano, M.B., Alford, S.T., Amy, N.E., *et al.* (2011) Expanding the Diversity of Mycobacteriophages: Insights into Genome Architecture and Evolution. *PLoS One*, 6.
135. Juhala, R.J., Ford, M.E., Duda, R.L., Youlton, A., Hatfull, G.F. and Hendrix, R.W. (2000) Genomic sequences of bacteriophages HK97 and HK022: Pervasive genetic mosaicism in the lambdoid bacteriophages. *J. Mol. Biol.*, 299.
136. Rousset, F., Dowding, J., Bernheim, A., Rocha, E.P.C. and Bikard, D. (2021) Prophage-encoded hotspots of bacterial immune systems. *bioRxiv*, 10.1101/2021.01.21.427644.
137. Stern, A. and Sorek, R. (2011) The phage-host arms race: shaping the evolution of microbes. *Bioessays*, 33.
138. Guo, H., Arambula, D., Ghosh, P. and Miller, J.F. (2014) Diversity-generating Retroelements in Phage and Bacterial Genomes. *Microbiol. Spectr.*, 2.
139. Paul, B.G., Burstein, D., Castelle, C.J., Handa, S., Arambula, D., Czornyj, E., Thomas, B.C., Ghosh, P., Miller, J.F., Banfield, J.F., *et al.* (2017) Retroelement-guided protein diversification abounds in vast lineages of Bacteria and Archaea. *Nat. Microbiol.*, 2.
140. Benler, S., Cobián-Güemes, A.G., McNair, K., Hung, S.-H., Levi, K., Edwards, R. and Rohwer, F. (2018) A diversity-generating retroelement encoded by a globally ubiquitous Bacteroides phage. *Microbiome*, 6.
141. Latka, A., Maciejewska, B., Majkowska-Skrobek, G., Briers, Y. and Drulis-Kawa, Z. (2017) Bacteriophage-encoded virion-associated enzymes to overcome the carbohydrate barriers during the infection process. *Appl. Microbiol. Biotechnol.*, 101.
142. Knecht, L.E., Veljkovic, M. and Fieseler, L. (2020) Diversity and Function of Phage Encoded Depolymerases. *Front. Microbiol.*, 10.
143. Rocha, E.P.C., Danchin, A. and Viari, A. (2001) Evolutionary role of restriction/modification systems as revealed by comparative genome analysis. *Genome Res.*, 11.
144. Rusinov, I.S., Ershova, A.S., Karyagina, A.S., Spirin, S.A. and Alexeevski, A. V. (2018) Avoidance of recognition sites of restriction-modification systems is a widespread but not universal anti-restriction strategy of prokaryotic viruses. *BMC Genomics*, 19.
145. Golovenko, D., Manakova, E., Tamulaitiene, G., Grazulis, S. and Siksnys, V. (2009) Structural mechanisms for the 5'-CCWGG sequence recognition by the N- and C-terminal domains of EcoRII. *Nucleic Acids Res.*, 37.

146. Zavil'gelskiĭ, G.B. and Kotova, V.I. (2014) Antirestriction activity of T7 Ocr protein in monomeric and dimeric forms. *Mol. Biol. (Mosk)*, 48.
147. Hill, C., Miller, L.A. and Klaenhammer, T.R. (1991) *In vivo* genetic exchange of a functional domain from a type II A methylase between lactococcal plasmid pTR2030 and a virulent bacteriophage. *J. Bacteriol.*, 173.
148. Loenen, W.A. and Murray, N.E. (1986) Modification enhancement by the restriction alleviation protein (Ral) of bacteriophage lambda. *J. Mol. Biol.*, 190.
149. Malone, L.M., Birkholz, N. and Fineran, P.C. (2021) Conquering CRISPR: how phages overcome bacterial adaptive immunity. *Curr. Opin. Biotechnol.*, 68.
150. Tao, P., Wu, X. and Rao, V. (2018) Unexpected evolutionary benefit to phages imparted by bacterial CRISPR-Cas9. *Sci. Adv.*, 4.
151. Bondy-Denomy, J., Pawluk, A., Maxwell, K.L. and Davidson, A.R. (2012) Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature*, 493.
152. Stanley, S.Y. and Maxwell, K.L. (2018) Phage-Encoded Anti-CRISPR Defenses. *Annu. Rev. Genet.*, 52.
153. Vlot, M., Houkes, J., Lochs, S.J.A., Swarts, D.C., Zheng, P., Kunne, T., Mohanraju, P., Anders, C., Jinek, M., Van Der Oost, J., *et al.* (2018) Bacteriophage DNA glucosylation impairs target DNA binding by type I and II but not by type V CRISPR-Cas effector complexes. *Nucleic Acids Res.*, 46.
154. Blower, T.R., Evans, T.J., Przybilski, R., Fineran, P.C. and Salmond, G.P.C. (2012) Viral Evasion of a Bacterial Suicide System by RNA-Based Molecular Mimicry Enables Infectious Altruism. *PLoS Genet.*, 8.
155. Sberro, H., Leavitt, A., Kiro, R., Koh, E., Peleg, Y., Qimron, U. and Sorek, R. (2013) Discovery of Functional Toxin/Antitoxin Systems in Bacteria by Shotgun Cloning. *Mol. Cell*, 50.
156. Samson, J.E., Bélanger, M. and Moineau, S. (2013) Effect of the abortive infection mechanism and type III toxin/antitoxin system AbiQ on the lytic cycle of *Lactococcus lactis* phages. *J. Bacteriol.*, 195.
157. Bingham, R., Ekunwe, S.I.N., Falk, S., Snyder, L. and Kleanthous, C. (2000) The Major Head Protein of Bacteriophage T4 Binds Specifically to Elongation Factor Tu. *J. Biol. Chem.*, 275.
158. Barth, Z.K., Nguyen, M.H.T. and Seed, K.D. (2021) A chimeric nuclease substitutes CRISPR-Cas: A phage weaponizes laterally acquired specificity to destroy subviral parasites. *bioRxiv*.
159. Zhvania, P., Hoyle, N.S., Nadareishvili, L., Nizharadze, D. and Kutateladze, M. (2017) Phage therapy in a 16-year-old boy with netherton syndrome. *Front. Med.*, 4.
160. Khawaldeh, A., Morales, S., Dillon, B., Alavidze, Z., Ginn, A.N., Thomas, L., Chapman, S.J., Dublanchet, A., Smithyman, A. and Iredell, J.R. (2011) Bacteriophage therapy for refractory *Pseudomonas aeruginosa* urinary tract infection. *J. Med. Microbiol.*, 60.
161. Roach, D.R., Leung, C.Y., Henry, M., Morello, E., Singh, D., Di Santo, J.P., Weitz, J.S. and Debarbieux, L. (2017) Synergy between the Host Immune System and Bacteriophage

- Is Essential for Successful Phage Therapy against an Acute Respiratory Pathogen. *Cell Host Microbe*, 22.
162. Sumrall, E.T., Shen, Y., Keller, A.P., Rismondo, J., Pavlou, M., Eugster, M.R., Boulos, S., Disson, O., Thouvenot, P., Kilcher, S., *et al.* (2019) Phage resistance at the cost of virulence: *Listeria monocytogenes* serovar 4b requires galactosylated teichoic acids for InlB-mediated invasion. *PLOS Pathog.*, 15.
 163. Cai, R., Wang, G., Le, S., Wu, M., Cheng, M., Guo, Z., Ji, Y., Xi, H., Zhao, C., Wang, X., *et al.* (2019) Three capsular polysaccharide synthesis-related glucosyltransferases, GT-1, GT-2 and WcaJ, are associated with virulence and phage sensitivity of *Klebsiella pneumoniae*. *Front. Microbiol.*, 10.
 164. Chan, B.K., Turner, P.E., Kim, S., Mojibian, H.R., Elefteriades, J.A. and Narayan, D. (2018) Phage treatment of an aortic graft infected with *Pseudomonas aeruginosa*. *Evol. Med. Public Heal.*, 2018.
 165. Tagliaferri, T.L., Jansen, M. and Horz, H.P. (2019) Fighting Pathogenic Bacteria on Two Fronts: Phages and Antibiotics as Combined Strategy. *Front. Cell. Infect. Microbiol.*, 9.
 166. León, M. and Bastías, R. (2015) Virulence reduction in bacteriophage resistant bacteria. *Front. Microbiol.*, 6.
 167. Chaudhry, W.N., Concepción-Acevedo, J., Park, T., Andleeb, S., Bull, J.J. and Levin, B.R. (2017) Synergy and Order Effects of Antibiotics and Phages in Killing *Pseudomonas aeruginosa* Biofilms. *PLoS One*, 12.
 168. Oechslin, F., Piccardi, P., Mancini, S., Gabard, J., Moreillon, P., Entenza, J.M., Resch, G. and Que, Y.-A. (2017) Synergistic Interaction Between Phage Therapy and Antibiotics Clears *Pseudomonas Aeruginosa* Infection in Endocarditis and Reduces Virulence. *J. Infect. Dis.*, 215.
 169. Schooley, R.T., Biswas, B., Gill, J.J., Hernandez-Morales, A., Lancaster, J., Lessor, L., Barr, J.J., Reed, S.L., Rohwer, F., Benler, S., *et al.* (2017) Development and use of personalized bacteriophage-based therapeutic cocktails to treat a patient with a disseminated resistant *Acinetobacter baumannii* infection. *Antimicrob. Agents Chemother.*, 61.
 170. Wright, R.C.T., Friman, V.P., Smith, M.C.M. and Brockhurst, M.A. (2019) Resistance evolution against phage combinations depends on the timing and order of exposure. *MBio*, 10.
 171. Nir-Paz, R., Gelman, D., Khouri, A., Sisson, B.M., Fackler, J., Alkalay-Oren, S., Khalifa, L., Rimón, A., Yerushalmy, O., Bader, R., *et al.* (2019) Successful Treatment of Antibiotic-resistant, Poly-microbial Bone Infection With Bacteriophages and Antibiotics Combination. *Clin. Infect. Dis.*, 69.
 172. Law, N., Logan, C., Yung, G., Furr, C.L.L., Lehman, S.M., Morales, S., Rosas, F., Gaidamaka, A., Bilinsky, I., Grint, P., *et al.* (2019) Successful adjunctive use of bacteriophage therapy for treatment of multidrug-resistant *Pseudomonas aeruginosa* infection in a cystic fibrosis patient. *Infection*, 47.
 173. Bao, J., Wu, N., Zeng, Y., Chen, L., Li, L., Yang, L., Zhang, Y., Guo, M., Li, L., Li, J., *et al.* (2020) Non-active antibiotic and bacteriophage synergism to successfully treat recurrent urinary tract infection caused by extensively drug-resistant *Klebsiella pneumoniae*. *Emerg. Microbes Infect.*, 9.

174. Tkhilaishvili,T., Winkler,T., Müller,M., Perka,C. and Trampuz,A. (2020) Bacteriophages as Adjuvant to Antibiotics for the Treatment of Periprosthetic Joint Infection Caused by Multidrug-Resistant *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.*, 64.
175. Dedrick,R.M., Guerrero-Bustamante,C.A., Garlena,R.A., Russell,D.A., Ford,K., Harris,K., Gilmour,K.C., Soothill,J., Jacobs-Sera,D., Schooley,R.T., *et al.* (2019) Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant *Mycobacterium abscessus*. *Nat. Med.*, 25.
176. Petrovic Fabijan,A., Lin,R.C.Y., Ho,J., Maddocks,S., Ben Zakour,N.L., Iredell,J.R., Khalid,A., Venturini,C., Chard,R., Morales,S., *et al.* (2020) Safety of bacteriophage therapy in severe *Staphylococcus aureus* infection. *Nat. Microbiol.*, 5.
177. Aslam,S., Courtwright,A.M., Koval,C., Lehman,S.M., Morales,S., Furr,C.L.L., Rosas,F., Brownstein,M.J., Fackler,J.R., Sisson,B.M., *et al.* (2019) Early clinical experience of bacteriophage therapy in 3 lung transplant recipients. *Am. J. Transplant.*, 19.
178. Dedrick,R.M., Freeman,K.G., Nguyen,J.A., Bahadırli-Talbott,A., Smith,B.E., Wu,A.E., Ong,A.S., Lin,C.T., Ruppel,L.C., Parrish,N.M., *et al.* (2021) Potent antibody-mediated neutralization limits bacteriophage treatment of a pulmonary *Mycobacterium abscessus* infection. *Nat. Med.*, 278.
179. Rubalskii,E., Ruemke,S., Salmoukas,C., Boyle,E.C., Warnecke,G., Tudorache,I., Shrestha,M., Schmitto,J.D., Martens,A., Rojas,S. V., *et al.* (2020) Bacteriophage therapy for critical infections related to cardiothoracic surgery. *Antibiotics*, 9.
180. Aslam,S., Lampley,E., Wooten,D., Karris,M., Benson,C., Strathdee,S. and Schooley,R.T. (2020) Lessons Learned From the First 10 Consecutive Cases of Intravenous Bacteriophage Therapy to Treat Multidrug-Resistant Bacterial Infections at a Single Center in the United States. *Open Forum Infect. Dis.*, 7.
181. Jault,P., Leclerc,T., Jennes,S., Pirnay,J.P., Que,Y.A., Resch,G., Rousseau,A.F., Ravat,F., Carsin,H., Le Floch,R., *et al.* (2019) Efficacy and tolerability of a cocktail of bacteriophages to treat burn wounds infected by *Pseudomonas aeruginosa* (PhagoBurn): a randomised, controlled, double-blind phase 1/2 trial. *Lancet Infect. Dis.*, 19.
182. Międzybrodzki,R., Borysowski,J., Weber-Dąbrowska,B., Fortuna,W., Letkiewicz,S., Szufnarowski,K., Pawełczyk,Z., Rogóż,P., Kłak,M., Wojtasik,E., *et al.* (2012) Clinical Aspects of Phage Therapy. *Adv. Virus Res.*, 83.
183. Ferry,T., Leboucher,G., Fevre,C., Herry,Y., Conrad,A., Josse,J., Batailler,C., Chidiac,C., Medina,M., Lustig,S., *et al.* (2018) Salvage debridement, antibiotics and implant retention (“dair”) with local injection of a selected cocktail of bacteriophages: Is it an option for an elderly patient with relapsing *staphylococcus aureus* prosthetic joint infection? *Open Forum Infect. Dis.*, 5.
184. Fish,R., Kutter,E., Bryan,D., Wheat,G. and Kuhl,S. (2018) Resolving digital staphylococcal osteomyelitis using bacteriophage—A case report. *Antibiotics*, 7.
185. Fadlallah,A., Chelala,E. and Legeais,J.-M. (2015) Corneal Infection Therapy with Topical Bacteriophage Administration. *Open Ophthalmol. J.*, 9.

186. Leszczyński,P., Weber-Dąbrowska,B., Kohutnicka,M., Łuczak,M., Górecki,A. and Górski,A. (2006) Successful eradication of methicillin-resistant *Staphylococcus aureus* (MRSA) intestinal carrier status in a healthcare worker - Case report. *Folia Microbiol. (Praha)*, 51.
187. Marza,J.A.S., Soothill,J.S., Boydell,P. and Collins,T.A. (2006) Multiplication of therapeutically administered bacteriophages in *Pseudomonas aeruginosa* infected patients. *Burns*, 32.
188. Letkiewicz,S., Miedzybrodzki,R., Fortuna,W., Weber-Dąbrowska,B. and Górski,A. (2009) Eradication of *enterococcus faecalis* by phage therapy in chronic bacterial prostatitis - Case report. *Folia Microbiol. (Praha)*, 54.
189. Jikia,D., Chkhaidze,N., Imedashvili,E., Mgaloblishvili,I., Tsitlanadze,G., Katsarava,R., Morris,J.G. and Sulakvelidze,A. (2005) The use of a novel biodegradable preparation capable of the sustained release of bacteriophages and ciprofloxacin, in the complex treatment of multidrug-resistant *Staphylococcus aureus*-infected local radiation injuries caused by exposure to Sr90. *Clin. Exp. Dermatol.*, 30.
190. Cano,E.J., Caflisch,K.M., Bollyky,P.L., Van Belleghem,J.D., Patel,R., Fackler,J., Brownstein,M.J., Horne,B., Biswas,B., Henry,M., *et al.* (2020) Phage Therapy for Limb-threatening Prosthetic Knee *Klebsiella pneumoniae* Infection: Case Report and *In Vitro* Characterization of Anti-biofilm Activity. *Clin. Infect. Dis.*
191. Gainey,A.B., Burch,A., Brownstein,M.J., Brown,D.E., Fackler,J., Horne,B., Biswas,B., Bivens,B.N., Malagon,F. and Daniels,R. (2020) Combining bacteriophages with cefiderocol and meropenem/vaborbactam to treat a pan-drug resistant *Achromobacter* species infection in a pediatric cystic fibrosis patient. *Pediatr. Pulmonol.*, 55.
192. Hoyle,N., Zhvaniya,P., Balarjishvili,N., Bolkvadze,D., Nadareishvili,L., Nizharadze,D., Wittmann,J., Rohde,C. and Kutateladze,M. (2018) Phage therapy against *Achromobacter xylosoxidans* lung infection in a patient with cystic fibrosis: a case report. *Res. Microbiol.*, 169.
193. Jennes,S., Merabishvili,M., Soentjens,P., Pang,K.W., Rose,T., Keersebilck,E., Soete,O., François,P.M., Teodorescu,S., Verween,G., *et al.* (2017) Use of bacteriophages in the treatment of colistin-only-sensitive *Pseudomonas aeruginosa* septicaemia in a patient with acute kidney injury-a case report. *Crit. Care*, 21.
194. Ferry,T., Kolenda,C., Batailler,C., Gustave,C.-A., Lustig,S., Malatray,M., Fevre,C., Josse,J., Petitjean,C., Chidiac,C., *et al.* (2020) Phage Therapy as Adjuvant to Conservative Surgery and Antibiotics to Salvage Patients With Relapsing *S. aureus* Prosthetic Knee Infection. *Front. Med.*, 7.
195. Rostkowska,O.M., Międzybrodzki,R., Miszevska-Szyszkowska,D., Górski,A. and Durlik,M. (2020) Treatment of recurrent urinary tract infections in a 60-year-old kidney transplant recipient. The use of phage therapy. *Transpl. Infect. Dis.*
196. Ferry,T., Boucher,F., Fevre,C., Perpoint,T., Chateau,J., Petitjean,C., Josse,J., Chidiac,C., L'hostis,G., Leboucher,G., *et al.* (2018) Innovations for the treatment of a complex bone and joint infection due to XDR *Pseudomonas aeruginosa* including local application of a selected cocktail of bacteriophages. *J. Antimicrob. Chemother.*, 73.

197. Rhoads,D.D., Wolcott,R.D., Kuskowski,M.A., Wolcott,B.M., Ward,L.S. and Sulakvelidze,A. (2009) Bacteriophage therapy of venous leg ulcers in humans: results of a phase I safety trial. *J. Wound Care*, 18.
198. Ferry,T., Batailler,C., Petitjean,C., Chateau,J., Fevre,C., Forestier,E., Brosset,S., Leboucher,G., Kolenda,C., Laurent,F., *et al.* (2020) The Potential Innovative Use of Bacteriophages Within the DAC Hydrogel to Treat Patients With Knee Megaprosthesis Infection Requiring “Debridement Antibiotics and Implant Retention” and Soft Tissue Coverage as Salvage Therapy. *Front. Med.*, 7.
199. Ooi,M.L., Drilling,A.J., Morales,S., Fong,S., Moraitis,S., MacIas-Valle,L., Vreugde,S., Psaltis,A.J. and Wormald,P.J. (2019) Safety and tolerability of bacteriophage therapy for chronic rhinosinusitis due to staphylococcus aureus. *JAMA Otolaryngol. - Head Neck Surg.*, 145.
200. Sarker,S.A., Sultana,S., Reuteler,G., Moine,D., Descombes,P., Charton,F., Bourdin,G., McCallin,S., Ngom-Bru,C., Neville,T., *et al.* (2016) Oral Phage Therapy of Acute Bacterial Diarrhea With Two Coliphage Preparations: A Randomized Trial in Children From Bangladesh. *EBioMedicine*, 4.
201. Khatami,A., Lin,R.C.Y., Petrovic-Fabijan,A., Alkalay-Oren,S., Almuzam,S., Britton,P.N., Brownstein,M.J., Dao,Q., Fackler,J., Hazan,R., *et al.* (2021) Bacterial lysis, autophagy and innate immune responses during adjunctive phage therapy in a child. *EMBO Mol. Med.*, 13.
202. Wright,A., Hawkins,C.H., Änggård,E.E. and Harper,D.R. (2009) A controlled clinical trial of a therapeutic bacteriophage preparation in chronic otitis due to antibiotic-resistant *Pseudomonas aeruginosa* ; a preliminary report of efficacy. *Clin. Otolaryngol.*, 34.
203. Doub,J.B., Ng,V.Y., Johnson,A.J., Slomka,M., Fackler,J., Horne,B., Brownstein,M.J., Henry,M., Malagon,F. and Biswas,B. (2020) Salvage bacteriophage therapy for a chronic MRSA prosthetic joint infection. *Antibiotics*, 9.
204. Duplessis,C., Biswas,B., Hanisch,B., Perkins,M., Henry,M., Quinones,J., Wolfe,D., Estrella,L. and Hamilton,T. (2018) Refractory *Pseudomonas* Bacteremia in a 2-Year-Old Sterilized by Bacteriophage Therapy. *J. Pediatric Infect. Dis. Soc.*, 7.
205. LaVergne,S., Hamilton,T., Biswas,B., Kumaraswamy,M., Schooley,R.T. and Wooten,D. (2018) Phage Therapy for a Multidrug-Resistant *Acinetobacter baumannii* Craniectomy Site Infection. *Open Forum Infect. Dis.*, 5.

2.7. Supplementary material

Supplementary Table S1. Complete list of case studies and clinical trials of phage therapy in humans, with associated safety, clinical, and phage-resistance outcomes.

Clinical study/case report	Safety	Clinical outcome ^a	Phage resistance ^b	Mechanism of resistance	Ref.
Tibia bone infection with MDR <i>A. baumannii</i> and <i>K. pneumoniae</i> treated with phages and antibiotics	Well tolerated	Successful	Not found in the patient; found in <i>in vitro</i> experiments with isolates	Mutations in surface adhesin and glycosyl-transferase of the EpsG family (speculated)	171
Necrotizing pancreatitis patient with an MDR <i>A. baumannii</i> infection treated with phages and antibiotics	Well tolerated	Successful	A resistant bacterial isolate was found and used to select a new phage	Loss of bacterial capsule and increased extracellular polysaccharide production	169
Cystic fibrosis patient with MDR <i>P. aeruginosa</i> infection treated with phages and antibiotics	Well tolerated	Successful	One resistant isolate was identified	Not described	172
Patient with <i>K. pneumoniae</i> urinary tract infection treated with non-active antibiotics and phages	Well tolerated	Successful	Bacteria became resistant to two rounds of phage cocktails but were ultimately sensitive to the combination of a third cocktail and previously inactive antibiotics	Not yet elucidated	173
Patient with periprosthetic joint infection and MDR <i>P. aeruginosa</i> chronic osteomyelitis treated with phages and antibiotics	Well tolerated	Successful	Not identified	Not applicable	174
Patient with MDR <i>P. aeruginosa</i> urinary tract infection treated with phages and antibiotics	Well tolerated	Successful	Not identified	Not applicable	160
Chronic prosthetic joint infection with <i>S. aureus</i> treated with phages	Well tolerated	Successful	Not described	Not described	183
Diabetic foot ulcer infected with <i>S. aureus</i> treated with a phage	Well tolerated	Successful	Not described	Not described	184
Patient with <i>S. aureus</i> corneal abscess and interstitial keratitis	Well tolerated	Successful	Not described	Not described	185

treated with a phage following antibiotic treatment					
Patient with MRSA urinary tract infection and gastrointestinal tract colonization treated with phages	Well tolerated	Successful	Not described	Not described	186
A man with burn wounds and a dog with otitis, both infected with <i>P. aeruginosa</i> and treated with phages	Well tolerated	Successful	Not described	Not described	187
Three patients with chronic bacterial prostatitis treated with phages	Well tolerated	Successful	Not described	Not described	188
Two patients with <i>S. aureus</i> infected radiation burns treated with phages and ciprofloxacin infused in a sustained release preparation	Well tolerated	Successful	Not described	Not described	189
Prosthetic joint infected with <i>K. pneumoniae</i> treated with a phage and minocycline	Well tolerated	Successful	Not described	Not described	190
Pediatric cystic fibrosis patient with an <i>Achromobacter</i> infection treated with phages and antibiotics	Well tolerated	Successful	Not described	Not described	191
Aortic graft infection with <i>P. aeruginosa</i> treated with a phage targeting an efflux pump protein and antibiotics	Well tolerated	Clinical improvement	Yes, as expected. Explored to cause sensitization of bacteria to antibiotic.	Mutations in receptor protein (M of mexAB- and mexXY-efflux system)	164
Cystic fibrosis patient with <i>M. abscessus</i> and <i>P. aeruginosa</i> infection treated with engineered phages	Well tolerated	Clinical improvement	Resistance detected <i>in vitro</i> for two of the three phages administered	Not described	175
Netherton syndrome patient with MDR <i>S. aureus</i> infection and allergy to multiple antibiotics treated with phage cocktails topically and orally	Well tolerated	Clinical improvement	Resistance to a phage cocktail was identified after 3 months of treatment	Not described	159

Cystic fibrosis patient with an <i>A. xylosoxidans</i> infection treated with phages	Well tolerated	Clinical improvement	Not described	Not described	192
Patient with colistin-only-sensitive <i>P. aeruginosa</i> septicemia	Well tolerated	Clinical improvement	Not described	Not described	193
Three patients with <i>S. aureus</i> prosthetic knee infection treated with debridement, a phage cocktail and antibiotics	Well tolerated	Clinical improvement, infection disappeared completely in two of the patients	Not described	Not described	194
Kidney transplant recipient with MDR <i>K. pneumoniae</i>	Well tolerated	Clinical improvement (not clear if due to phage therapy)	Not described	Not described	195
Thirteen patients with <i>S. aureus</i> bacteremia treated with phages and antibiotics	Well tolerated	Clinical improvement in eight of the patients	Changes in phage susceptibility were detected <i>in vivo</i> in three patients	Single nucleotide polymorphisms were detected in isolates recovered from one of the patients	176
Three lung transplant patients with MDR <i>P. aeruginosa</i> and <i>B. dolosa</i> infections treated with phages and antibiotics	Well tolerated	Clinical improvement in two of three patients	Resistant isolates were identified post-therapy in one of the patients who had a successful outcome	Not described	177
Complex bone and joint infection with <i>P. aeruginosa</i> treated with phages and antibiotics	Well tolerated	Clinical improvement; patient died because of unrelated complications	Not described	Not described	196
Phase I clinical trial to assess the safety of a phage preparation in 39 patients (18 treated, 21 control) with chronic venous leg ulcers	Well tolerated	No significant differences found between treated and control groups; trial was not designed to assess efficacy	Not described	Not described	197
Patient with bronchiectasis and <i>M. abscessus</i> infection treated with a phage cocktail and antibiotics	Well tolerated	Failure due to anti-phage immune response	Resistant isolates appeared to one out of three phages in the cocktail	Not described	178

Patient with <i>S. aureus</i> infection of a knee megaprosthesis treated with debridement, a phage cocktail prepared in a hydrogel and antibiotics	Well tolerated (adverse events described are not considered to be related to phage therapy)	Site of surgery did not heal completely and got re-infected (not with the strain targeted by phage therapy)	Not described	Not described	198
Eight cardiothoracic surgery patients treated with antibiotics, phages and fibrin glue	Well tolerated except for an increase in inflammation	Successful in five of the patients; two patients showed improvement but died of unrelated complications; one patient did not experience sufficient bacterial clearance and died of sepsis	Bacterial isolates appeared to have mutated in one of the patients	Not described	179
Phase I clinical trial to assess safety and preliminary efficacy of a phage cocktail in 9 patients with <i>S. aureus</i> chronic rhinosinusitis	Well tolerated, six patients reported mild adverse events	Preliminary efficacy results suggested clinical improvement	Not described	Not described	199
Ten patients with diverse MDR bacterial infections treated with intravenous phage cocktails and antibiotics	Well tolerated except by one patient, who developed fever, wheezing and shortness of breath after first dose of the cocktail (subsequent doses were well tolerated)	Successful for seven out of ten patients	Resistance developed in three of the patients, but was overcome through administration of phages specific for the resistant isolates (two of these cases still had a successful outcome)	Not described	180
Clinical trial to assess the safety and efficacy of a T4-like phage cocktail in pediatric patients with diarrhea	No adverse events were reported in the treated group compared to the control group	No significant differences in treatment outcome were found between the treated and the control groups (phage cocktail had not been tailored to patients' infections)	Not described	Not described	200
Phase I-II clinical trial to assess the safety and efficacy of a phage cocktail to treat burn wounds infected with <i>P. aeruginosa</i> in 26	Fewer adverse events were observed in the phage-treated group than in the group treated	Trial was stopped because of insufficient efficacy (patients were being exposed to 10,000-fold lower	Intermediately susceptible and resistant isolates were found	Not described	181

patients (13 treated, 13 control)	with the standard of care	doses of phages than originally intended)			
Pediatric patient with <i>P. aeruginosa</i> osteoarticular infection treated with phages and antibiotics	Fever and increase in C-reactive protein associated with release of bacterial remains into the bloodstream	Clinical improvement	Not described	Not described	201
Clinical trial evaluating the safety and efficacy of a phage preparation in 24 patients (12 treated, 12 control) with chronic otitis caused by <i>P. aeruginosa</i>	Mild to moderate adverse events were reported, including discomfort, itchiness and wetness	Clinical improvement compared to the control group in all but one patient	Not described	Not described	202
Patient with chronic MRSA prosthetic joint infection treated with phages and antibiotics	Transaminitis was reported following the third intravenous dose of phages	Clinical improvement	Not described	Not described	203
Pediatric patient with congenital heart disease and <i>P. aeruginosa</i> bacteremia treated with phages and antibiotics	Treatment was discontinued due to cardiac decompensation with concerns for anaphylaxis (although subsequently attributed to the pre-existing disease)	Blood sterilization could be achieved; sepsis was developed again after the treatment was interrupted	Not described	Not described	204
Cerebritis patient with <i>A. baumannii</i> infection treated with phages and colistin	Brief hypotension episode	Unsuccessful; therapy was discontinued	Not described	Not described	205
Report of phage therapy performed on 153 patients with different infections	Diverse	Diverse	Resistance and changes in the phage typing profile were detected in multiple cases	Not described	182





3.

Characterization of Class 1 DISARM as bacterial defense system

Preprint as *Class 1 DISARM provides anti-phage and anti-conjugation activity by unmethylated DNA recognition.* Cristian Aparicio-Maldonado, Gal Ofir, Andrea Salini, Rotem Sorek, Franklin L. Nobrega, Stan J.J. Brouns (2021). **BioRxiv.** doi: 10.1101/2021.12.28.474362.

Abstract

Bacteriophages impose a strong evolutionary pressure on microbes for the development of mechanisms of survival. Multiple new mechanisms of innate defense have been described recently, with the molecular mechanism of most of them remaining uncharacterized. Here, we show that a Class 1 DISARM (defense island system associated with restriction-modification) system from *Serratia* sp. provides broad protection from double-stranded DNA phages, and drives a population of single-stranded phages to extinction. We identify that protection is not abolished by deletion of individual DISARM genes and that the absence of methylase genes *drmMI* and *drmMII* does not result in autoimmunity. In addition to antiphage activity we also observe that DISARM limits conjugation, and this activity is linked to the number of methylase cognate sites in the plasmid. Overall, we show that Class 1 DISARM provides robust anti-phage and anti-plasmid protection mediated primarily by *drmA* and *drmB*, which provide resistance to invading nucleic acids using a mechanism enhanced by the recognition of unmethylated cognate sites of the two methylases *drmMI* and *drmMII*.

3.1. Introduction

The constant bacteriophage predation of bacteria drives co-evolutionary dynamics and has led to the evolution of multiple antiviral defense mechanisms in prokaryotes that are collectively known as the prokaryotic immune system ^{1,2}. Most prokaryotic defenses are innate, acting via recognition of general signals that are not shared by the microbial cell. A well-known example of innate defense are the highly abundant restriction-modification (R-M) systems ³ that cleave phage nucleic acids at sequence motifs protected in the host chromosome by epigenetic modifications ⁴. In these systems a methyltransferase marks self-DNA at specific sites, which protects it from the cleave by a restriction endonuclease; in opposition to foreign unmodified DNA that will be cleaved. They can be classified in 4 different types (I-V), depending on their components ⁴. Type II systems are the most abundant, and are formed by a methyltransferase and a nuclease which will perform their action independently as separate proteins. However, in types I and III, both proteins will form a complex and perform their action when bound together. Interestingly, type IV systems don't contain a methyltransferase, and will only cleave modified DNA ^{5,6}. Recently, the analysis of genomic neighborhoods of known defense systems revealed multiple new anti-phage systems ⁷⁻¹⁵, among which is the defense island system associated with restriction-modification (DISARM) ¹⁶.

The DISARM system is composed of three-core genes: gene *drmA* with a helicase domain (pfam00271); gene *drdB* with a DUF1998 domain (pfam09369, helicase-associated); and gene *drmC* with a phospholipase D (PLD) domain (pfam13091). In the uncharacterized Class 1 DISARM systems, this core triplet is preceded by the SNF2-like helicase *drmD* (pfam00176), and the DNA adenine N6 methylase *drmMI* (pfam13659) ¹⁶. Class 2 DISARM systems contain, in addition to the core gene triplet, the DNA 5-cytosine methylase *drmMII* (pfam00145) and, on occasion, the gene of unknown function *drmE*. The Class 2 DISARM system was shown to use methylation of specific host motifs to mark self-DNA, akin to R-M systems, but its molecular mechanism remains largely unknown ¹⁶. Unlike R-M systems, DISARM seems to not fully depend on the

sequence motif identified by the methylase to interfere with the incoming DNA; importantly, the candidate nuclease of the system (*drmO*) was found dispensable for resistance and its activity is yet unknown.

In some cases, Class 1 DISARM systems are accompanied by an additional cytosine methylase *drmMIII* gene¹⁶. Here we characterized such a Class 1 DISARM system present in *Serratia* sp. SCBI (**Fig. 1A**). We found that Class 1 DISARM provides broad anti-phage and anti-conjugation activity independent of methylation status of the incoming DNA, and drives a population of chronic infecting phages to extinction. Unlike Class 2 DISARM and R-M systems, the methylases of Class 1 DISARM only partially methylate adenine and cytosine bases of the host DNA at motifs ACACAG and MTCGAK, and the absence of the methylases does not result in autoimmunity. Overall, our results show that DISARM combines methylation and non-methylation signals to provide protection against invader DNA, establishing a clear distinction from R-M systems.

3.2. Results

3.2.1. Class 1 DISARM protects against widely diverse DNA phages

To determine whether the predicted Class 1 DISARM system from *Serratia* sp. SCBI provides protection from phage infection, we transplanted the six genes of the system into *E. coli* BL21-AI (**Fig. 1A**). We then challenged the DISARM-containing strain (DISARM (+)) with *Caudovirales* of three morphologies (siphon-, myo-, and podophages) ¹⁷ at different multiplicity of infection (MOI). Class 1 DISARM shows clear anti-phage protection against all phages tested, by preventing or delaying the collapse of the bacterial population upon phage infection even at high MOI (**Fig. 1B**, **Sup. Fig. S1**). To quantify the level of protection, we measured the efficiency of plating (EOP) of the same set of phages on the DISARM (+) strain in comparison to control (DISARM (-)) cells. Class 1 DISARM provided significant protection against phages T1, T4, T7, Nami and phiX174 (**Fig. 1C**).

Overall, both liquid and solid media assays demonstrate the broad anti-phage activity provided by Class 1 DISARM.

3.2.2. Class 1 DISARM can drive a phage population with chronic lifestyle to extinction

We next investigated the effects of the Class 1 DISARM on the propagation of a phage and accumulation of active phage in the cell culture. For this, we measured the phage titers periodically upon infection of DISARM (+) or DISARM (-) strains with T1 or Nami phages. The number of infectious T1 phages in the population is reduced by approximately 3 orders of magnitude in DISARM containing strains from 30 min on, reaching a maximum reduction of 5.6×10^4 fold at 180 minutes post-infection (**Fig. 1D**). DISARM also inhibits the propagation of phage Nami, with a maximum reduction of approximately 50-fold at 180 minutes post-infection (**Fig. 1D**).

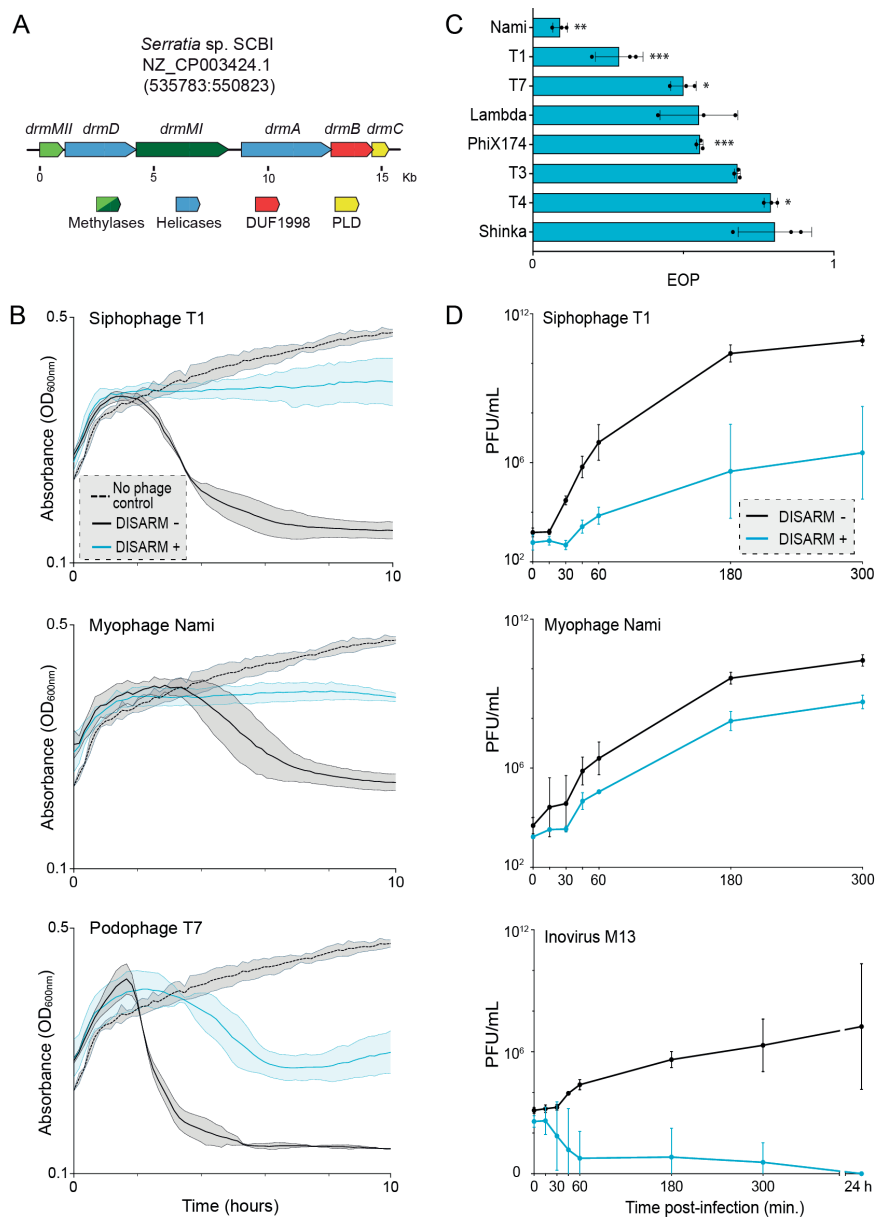


Figure 1. Protection provided by Class 1 DISARM against phage infection. **A)** Gene cluster of Class 1 DISARM from *Serratia* sp. SCBI. **B)** Effect of siphophage T1, podophage T7, and myophage Nami on the growth curve of DISARM (+) or DISARM (-) strains. Uninfected DISARM (-) and DISARM (+) strains have similar growth curves and only uninfected DISARM (-) is displayed. Full results can be seen in **Sup. Fig. S2**. Initial multiplicity of infection (MOD): 5×10^{-4} . Filled areas inside dotted lines indicate standard deviation of three independent replicates.

Virulent *Caudovirales* follow a lytic life cycle where the production of phages occurs typically within 10-30 minutes after the ejection of the phage genome, ultimately resulting in cell death for the release of the newly formed phage particles. Phages that follow a chronic life cycle are able to produce new phages continuously without causing cell death, with the new virions extruding out of the cell ¹⁸. To investigate the effect of Class 1 DISARM on the propagation of a phage population with chronic lifestyle, we monitored the phage titers of a culture infected with the single-stranded DNA (ssDNA) inovirus M13. We observed a rapid decrease of the number of phages, with DISARM containing strains producing no more phage after 24h (**Fig. 1D**), which was not observed for any of the *Caudovirales* tested. To understand if the strong activity of Class 1 DISARM against phage M13 is a consequence of its chronic lifestyle or the type of genetic material (ssDNA versus the double-stranded DNA of *Caudovirales*), we additionally tested DISARM against ssDNA phage phiX174. PhiX174 uses a mechanism of phage DNA replication similar to that of phage M13 ^{19,20}, but follows a lytic life cycle. The protective effect of DISARM against infection by phiX174 was similar to that obtained for the dsDNA phages (**Sup. Fig. S3A**), suggesting that the strong effect of DISARM against phage M13 is not related to the single stranded nature of the phage genome in the phage particle. We further tested the protective effect of Class 1 DISARM against single-stranded RNA (ssRNA) phage MS2 but observed no protection (**Sup. Fig. S3B**), suggesting that Class 1 DISARM is a DNA-directed defense system. In summary, Class 1 DISARM reduces the number of infectious phages produced over time in a bacterial culture, and is able to completely abolish the propagation of chronic phage M13.

[Fig.1]...C) Efficiency of plating (EOP) of a set of DNA phages in a DISARM (+) strain normalized to the DISARM (-) strain. D) Effect of DISARM on the population of siphophage T1, myophage Nami, and inovirus M13 over time. Bacterial cultures of DISARM (-) or DISARM (+) strains were infected with phage at MOI of 2×10^{-6} , 1×10^{-5} and 2×10^{-6} for phages T1, Nami and M13 respectively, and the titer was determined at selected time points. Curves depict the average and standard deviation of three independent experiments. Statistical significance was determined by two-tailed, unpaired t-test and is represented as *, **, or *** for $p < 0.05$, 0.01 or 0.001, respectively.

3.2.3. Class 1 DISARM provides protection independent of methylation status

We next studied the essentiality of the individual genes of the Class 1 DISARM system in anti-phage protection. First, we used EOP assays to challenge cells harboring the complete or partial DISARM system. We observed that DISARM provides protection against infection by dsDNA phages T1 and Nami, and that some of the DISARM genes can be deleted while retaining either full or partial protection (**Fig. 2A, Sup. Fig. S4A**). To better evaluate gene essentiality in Class 1 DISARM, we followed phage propagation over time in liquid culture of cells harboring the complete or partial DISARM system. Results are shown in **Fig. 2B,C and Sup. Fig. S4B** as the increase of phage titer over time at the time point where DISARM achieved the strongest effect compared to the titer at the start of infection. In contrast to what was previously observed for Class 2 DISARM systems ¹⁶, deletion of core genes *drmABC* did not abolish the full protective effect for both phages T1 and M13, and only the additional deletion of *drmMII* resulted in complete loss of protection with full restoration of the phage replication capacity. For phage Nami, deletion of core *drmABC* or deletion of *drmD* and *drmMI* resulted in almost complete loss of protection (**Fig. 2B**), with full restoration of phage replication being achieved with their combined deletion. Importantly, we observed that *drmC* is not required for the anti-phage activity (**Sup. Fig. S5**) of Class 1 DISARM, as previously observed for Class 2 DISARM ¹⁶.

Overall, we observed that Class 1 DISARM phage protection is relatively robust to deletion of individual genes of the operon.

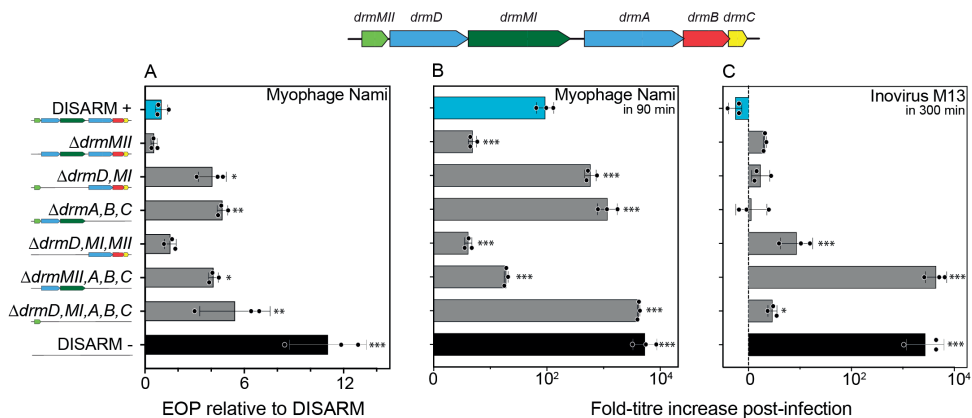


Figure 2. Effect of Class 1 DISARM components on protection against phage infection. **A)** Efficiency of plating (EOP) of myophage Nami on strains containing the modified DISARM system (DISARM (-) strain), normalized to the DISARM strain. **B-C)** Titer-fold increase of myophage Nami (**B**), and inovirus M13 (**C**) upon propagation in cultures of strains containing the complete or modified DISARM system. Graphics represent the time point at which maximum effect on phage replication was achieved. For (**C**), negative values indicate phage titers below the initial phage titer. Bars depict the average and standard deviation of three independent replicates. Statistical significance was determined by one-way ANOVA+ with Tukey post-hoc test and is represented as *, **, or *** for $p < 0.05$, 0.01 or 0.001 , respectively.

3.2.4. Class 1 DISARM of *Serratia* sp. SCBI modifies host DNA with two methylation patterns

To understand if the methylase genes *drmMI* and *drmMII* methylate the host DNA, we sequenced the genomes of the wild type strain, the DISARM (+) strain, and the strain containing one or both methylases using sequencing methods sensitive to epigenetic marks. The adenine methylation by *drmMI* was characterized by PacBio sequencing and revealed an N6-methyladenosine (m6A) modification of the ACACAG motif (methylated base underlined, **Fig. 3A, Sup. Table S5**). The cytosine methylation by *drmMII* was characterized by bisulfite sequencing and revealed a 5-methylcytosine (5mC) modification of MTCGAK motifs (methylated cytosine underlined, **Fig. 3B, Sup. Table S5**), which is a distinct motif from the 5mC modification in CCWGG motifs reported for the Class 2 DISARM system of *Bacillus subtilis*¹⁶. The completeness of modification by the methylases was higher in the presence of the full DISARM system (84.3% m6A, and 67.0% 5mC) than in the presence of the two methylases alone (56.5% m6A, and 34.0% 5mC), suggesting some form of synergetic effect by the methylase pair (**Fig. 3D**). Curiously, the presence of both *drmMI* and *drmMII* increased the m6A methylation ratio of *drmMI* to 73.9%, but had no effect on 5mC methylation by *drmMII*. No relation was observed between the distance of methylated or unmethylated ACACAG and MTCGAK motifs, suggesting that their methylation status was independent of their relative location in the genome. In summary, the Class 1 DISARM system of *Serratia* sp. SCBI uses both adenine and cytosine methylation to modify specific motifs in the bacterial chromosome.

[Fig. 3]...pCONJ with variable numbers of unmethylated motifs ACACAG and/or MTCGAK into the recipient DISARM (+) strain (see **Sup. Fig. S5B** for results in recipient control strain). Control is pCONJ with no motifs in its sequence. F) Conjugation efficiency of plasmid pCONJ with 10 methylated motifs ACACAG or MTCGAK into the recipient DISARM (+) strain, normalized by conjugation efficiency of pCONJ with unmethylated motifs.

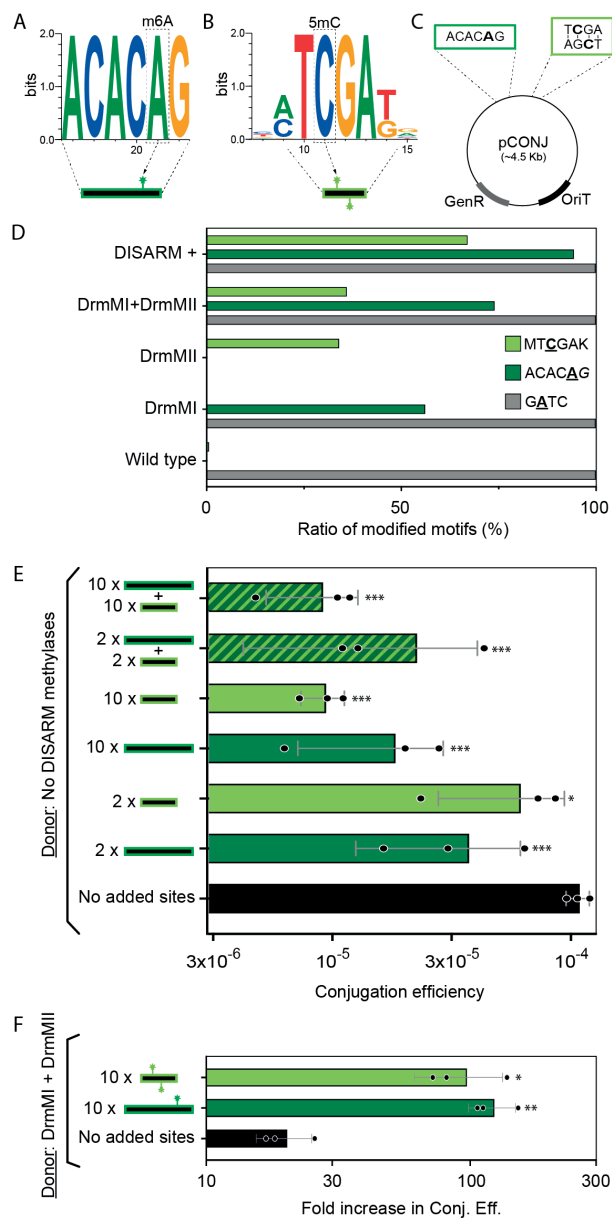


Figure 3. Effect of DNA methylation in Class 1 DISARM. Weblogos of methylation motifs of DISARM methylases (A) DrmMII and (B) DrmMII. C) Schematic representation of pCONJ plasmid with motifs as defined in (A) and (B). D) Relative number of modified sites detected in the genome of *E. coli* expressing the Class 1 DISARM system of *Serratia* sp. SCBI. Three distinct DNA modifications were detected: GATC sites modified by *dam* from *E. coli*; ACACAG sites modified by *drmMI*; and MTCGAK sites modified by *drmMII*. E) Conjugation efficiency of plasmid

3.2.5. Class 1 DISARM displays anti-conjugation activity dependent on number of cognate sites

To determine the influence of the methylation pattern of the invading DNA on the level of protection by Class 1 DISARM, we performed conjugation assays with plasmid pCONJ which does not contain ACACAG or MTCGAK motifs (**Sup. Table S2**). We first assessed the protection provided by DISARM towards conjugation of the unmethylated pCONJ, and observed an approximately 47-fold reduction in conjugation efficiency in the presence of DISARM (**Sup. Fig. S6A**). Next, we engineered different versions of pCONJ with variable numbers of DISARM methylation motifs and performed conjugation assays in the DISARM (+) and DISARM (-) strains as recipients. The conjugation efficiency of non-methylated pCONJ in the DISARM (+) strain decreased moderately with increasing numbers of ACACAG and TCGAK motifs present in the plasmid, up to 12-fold (**Fig. 3E, Sup. Fig. S6B**). The protective effect against plasmid conjugation is only slightly stronger when both types of motifs are present.

Next, we compared the conjugation efficiency of pCONJ originated from cells expressing DISARM methylases DrmMI and DrmMII to that of pCONJ containing non-methylated motifs (**Fig. 3F**). The conjugation efficiency of pCONJ was drastically elevated by 134- and 82-fold when the donor strain contained both DrmMI and DrmMII, supporting the role of methylation as an off switch for DISARM activity. The control plasmid, in which no ACACAG and MTCGAK motifs were added, displayed a 21-fold increase in conjugation efficiency with both methylases present in donor strain possibly due to effects of induction of gene expression of both methylases in the donor. Overall, the conjugation assays demonstrate that Class 1 DISARM displays anti-conjugation activity that is enhanced by the presence of unmethylated forms of both motifs in the plasmid.

3.3. Discussion

Here, we show that Class 1 DISARM of *Serratia* sp. SCBI provides broad protection against phages and plasmid conjugation using a mechanism of incoming nucleic acid detection which is enhanced by the recognition of unmethylated cognate sites of the two methylases *drmMI* and *drmMII*. In Class 2 DISARM systems, the deletion of genes *drmA*, *drdB*, or *drmE* resulted in complete loss of protection ¹⁶. The Class 1 DISARM system of *Serratia* sp. provides a more robust protection, in which the deletion of individual genes is not sufficient to abolish the protective effect of the system against some phages (e.g. T1 and M13). We observed that the number of unmethylated motifs present in the incoming foreign DNA increased the protective effect of DISARM against incoming invaders.

We found that the protection against phages was not scaled to the number of *drmMI* and *drmMII* sites in their genomes (**Sup. Table S6**), suggesting also other factors at play. Similarly, the number of methylation sites on phage genomes did not affect the protection for the BREX ²¹ system, and the same was suggested for Class 2 DISARM ¹⁶. This is markedly distinct from tested R-M systems in which restriction (and therefore protection) is dependent on the number of methylation sites in the invader's DNA ²². This suggests that DISARM and BREX use mechanisms to identify invader DNA distinct from those of R-M systems, and which may prevent strong negative selection for specific methylation motifs. It is also possible that the intrinsic methylation patterns of phage DNA affect defense by DISARM, as observed previously for R-M and BREX systems ^{23,24}.

The use of a molecular mechanism distinct from classical R-M systems is further supported by the lack of autoimmunity in cells in the absence of *drmMI* and *drmMII* (**Sup. Fig. S7**), contrary to R-M systems where this results in cleavage of the bacterial DNA ²⁵. This is also consistent with the fact that not all motifs in the bacterial genome of the transplanted host are methylated by the DISARM methylases, unlike the almost complete motif methylation observed in R-M. This suggests a tight regulation of the defense activity of Class 1 DISARM

that seems to result from the physical occlusion of the DNA entry site of the DrmAB complex by the trigger loop, which is removed upon presence of 5' ssDNA end ²⁶. Interestingly, the methylase dr mMII of Class 2 DISARM provides almost complete methylation of motifs in the bacterial chromosome and its absence was found to be deleterious to the cells ¹⁶. Differences in the molecular mechanisms of Class 1 and Class 2 DISARM likely result from the use of distinct methylases. We found that *Serratia*'s Class 1 DISARM system has the unique feature of combining methylation of both palindromic and non-palindromic motifs in the bacterial chromosome. Akin to BREX and R-M type I and III systems, DrmMI of Class 1 DISARM methylates a non-palindromic site (ACACAG). The modification occurs at the adenine in the fifth position of the recognition site, as previously reported for the BREX site TAGGAG ¹¹. Because the recognition motif of DrmMI is non-palindromic, only one DNA strand will be methylated. Some R-M systems (e.g. type III and type ISP) maintain the epigenetic marks by requiring interactions between different sites (29), but it is unclear if DISARM and BREX use similar mechanisms.

Contrary to DrmMI, the DrmMII methylase of Class 1 DISARM modifies a degenerate palindromic site (MTCGAK) in the bacterial chromosome. The methylation site of DrmMII of the *Bacillus paralicheniformis* Class 2 DISARM system was also shown to be palindromic, although of an unrelated sequence (CCNGG) ¹⁶, much like the methylases of R-M type II systems. Interestingly, DrmMII (palindromic motif) alone was shown to provide anti-phage activity against the chronic infecting phage M13 and phage T1, possibly by interfering with the phage replication cycle as observed previously ²⁸. DrmMI (non-palindromic motif) had no observable impact on anti-M13 activity, as previously observed also for the BREX system (non-palindromic motif). DrmABC without any of the methylases also provides protection against M13. It is possible that the strong activity of DISARM against phage M13 results from the added effect of DrmABC and a possible direct effect of DrmMII methylation of palindromic sites on DNA replication of M13. In conclusion, we show that Class 1 DISARM systems are effective on viral and plasmid DNA, and show enhanced protection

against invader DNA when unmethylated cognate DNA motifs are present (Fig. 4). The mechanism of DISARM is remarkably robust with many of its components playing an enhancing but non-essential role.

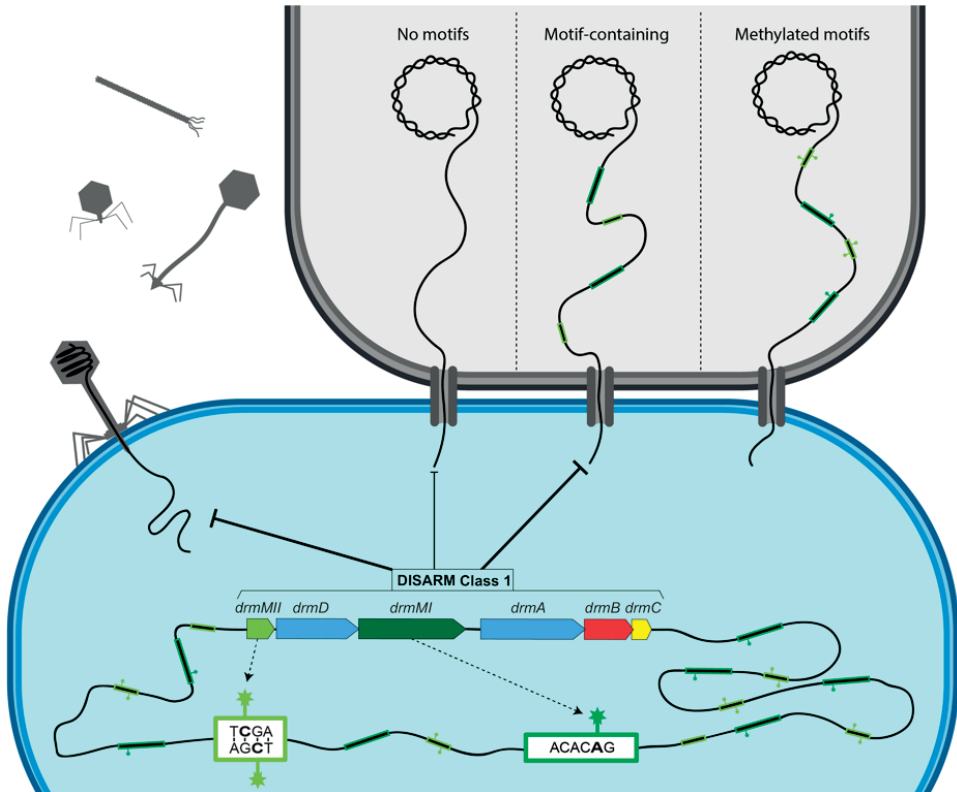


Figure 4. Mechanism of action of Class 1 DISARM of *Serratia sp.* SCBI. The methylases *drmMI* and *drmMII* methylate the host DNA at motifs ACACAG and MTCTGAK, respectively. The DISARM system provides protection from incoming unmethylated plasmid DNA and is less active on incoming DNA with methylated DISARM motifs. The efficiency of the protection increases with the number of unmethylated motifs present in the conjugated plasmid DNA.

3.4. Material and methods

3.4.1. Bacterial strains and growth conditions

Escherichia coli strains DH5 α , BL21-AI, JM109, ER2738, WG5, C3000, and S17-1 were cultured in Lysogeny Broth (LB) media at 37°C. For solid media experiments, LB was supplemented with 1.5% (w/v) agar (LBA) and the cultures incubated overnight at 37°C. When required, media was supplemented with antibiotics at the following final concentrations: 100 μ g/mL ampicillin, 50 μ g/mL kanamycin, 50 μ g/mL streptomycin, 50 μ g/mL gentamicin, and 25 μ g/mL chloramphenicol. Gene expression was induced with 1 mM Isopropyl β -D-1-thiogalactopyranoside (IPTG) and 0.2% (w/v) L-arabinose. The *Serratia* sp. SCBI strain (South African *Caenorhabditis briggsae* Isolate) was grown aerobically in LB media at 30°C with shaking at 180 rpm.

3.4.2. Phage cultivation

E. coli phages were propagated using their host strain (Sup. Table S1) as described previously²⁹. Briefly, bacterial cultures at early exponential growth phase (0.3-0.4 OD₆₀₀) were infected with a phage lysate and incubated overnight at 37°C with shaking. Cultures were spun down and the supernatant filtered (0.2 μ m PES) and stored at 4°C until use. When required, phages were concentrated by adding PEG-8000 and NaCl at final concentrations of 100 mg/mL and 1 M, respectively. The suspension was incubated overnight at 4°C, centrifuged at 11,000 $\times g$ at 4°C for 1h, and the phage-containing pellet was re-suspended in SM buffer (100 mM NaCl, 8 mM MgSO₄·7H₂O, and 50 mM Tris-HCl pH 7.5). The phage titer was determined using the small drop plaque assay method as described by Mazzocco et al.³⁰. *E. coli* strain BL21-AI was used for titering and performing assays of phages T1, T3, T4, T7, Lambda-vir, Myo21S, and Myo22L, strain WG5 was used for MS2, and strain C3000 was used for PhiX174. Phage M13 was titrated using *E. coli* strain ER2738 and 0.6% LBA supplemented with 1 mM IPTG and 200 μ g/mL X-gal. Assays of M13 phage were performed in strain JM109 (DE3).

3.4.3. Cloning of Class 1 DISARM and mutants

Plasmids and primers used in this study are listed in **Sup. Table S2** and **Sup. Table S3**, respectively. The complete Class 1 DISARM system and different combinations of its genes were amplified by PCR from *Serratia* sp. SCBI genomic DNA with primers indicated in **Sup. Table S3** and using Q5 DNA Polymerase (New England Biolabs) according to the manufacturer's instructions. The PCR products were cloned into the plasmid backbones using the NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolabs) following manufacturer's instructions. Plasmid pDIS_3 was used as a template for PCR with primers listed in **Sup. Table S3** to construct pDIS_5, pDIS_6 and pDIS_7 by restriction cloning. All plasmids were confirmed by Sanger sequencing (Macrogen) and transformed into strains *E. coli* BL21-AI and *E. coli* JM109(DE3) for the assays described below.

3.4.4. Phage infection growth curves

Bacterial cultures with the DISARM system or empty vector were grown to early exponential phase ($OD_{600\text{ nm}}$ of 0.2-0.3), induced with IPTG and L-arabinose, and grown for 90 min at 37 °C and 180 rpm. The cultures were normalized to an $OD_{600\text{ nm}}$ of 0.5 (approximately 1×10^8 CFU/ml) and 190 μ l were dispensed into wells of 96-well microtiter plates. Then, 10 μ l of phage suspension were added to the wells at different MOIs, and bacterial growth was followed in an EPOCH2 microplate reader with $OD_{600\text{ nm}}$ measurements every 10 min at 37°C with constant shaking.

3.4.5. Phage replication over time

Bacterial cultures were prepared as above and infected with phages at different MOIs. Cultures were incubated at 37°C with shaking at 180 rpm, and phage titers were measured over time using the small drop plaque assay method using the wild type strain as the host.

3.4.6. Methylation-sensitive DNA sequencing

3 Genomic DNA was extracted from the wild-type strain and strains containing the complete DISARM system or its methylases, using the phenol chloroform method as described before ³¹ with some modifications ³². Briefly, cultures at exponential phase were induced and incubated overnight. Bacterial cells were pelleted, re-suspended in TE buffer, and treated with RNase I and lysozyme at 1 µg/ml for 1h, followed by proteinase K at 50 µg/ml for 1h. DNA was extracted twice with phenol:chloroform (1:1) and precipitated by adding 300 mM sodium acetate pH 5.2 and two volumes of absolute ethanol. After incubation at -20°C for 1h or overnight, the DNA was pelleted by centrifugation at 21,000 × *g*, and the pellet re-suspended in nuclease-free water. The DNA was quantified using Qubit dsDNA HS Assay kit (Invitrogen) and the quality assessed using Nanodrop (Thermo Scientific).

The genomic DNA was sequenced by Pacific Biosciences using Single Molecule Real Time (SMRT) sequencing technology to detect DNA methylation sites ³³. DNA libraries were prepared using the SMRTbell® Express Template Prep Kit 2.0 and the Barcoded Overhang Adapter kit according to the manufacturer instructions. Sequencing was performed on a PacBio Sequel platform using Sequel sequencing kit 3.0. The data were analyzed using the Base Modification and Motif Analysis module of the SMRT-Link v7.0.1 software to detect methylation motifs.

Bisulfite library preparation and sequencing was done as previously described ¹⁶. Sequencing results were analyzed using Bismark v0.20.1 ³⁴ to identify methylated cytosines. The sequence neighborhood of methylated cytosines was analyzed using Weblogo ³⁵ to determine the methylation motif.

Raw sequencing data were deposited in Figshare for PacBio (doi.org/10.6084/m9.figshare.17295215) and for Bisulfite (doi.org/10.6084/m9.figshare.17295212) sequencing.

3.4.7. Construction of motif-containing conjugative plasmids

Synthetic constructs (Integrated DNA Technologies) of tetracycline-regulated YFP and a motif-adaptable module (MAM) (Sup. Table S4) were introduced into pSEVA_331 by restriction cloning. Putative DISARM motifs were removed from the backbone sequence of pSEVA_331 by replacing these regions with synthetic constructs (Sup. Table S4). For this, pSEVA_331 and the synthetic constructs were amplified using primers in Sup. Table S3 and Q5 DNA Polymerase (New England Biolabs), and cloned by restriction digest using the enzymes indicated in Sup. Tables S3 and S4, giving rise to plasmid pCONJ. Plasmids pCONJ_1 to pCONJ_8 (Sup. Table S2) were created by cloning synthetic constructs (Sup. Table S4) containing different motif combinations into the MAM regions of pCONJ using restriction digest as above. All plasmid constructs were confirmed by sequencing (Macrogen) and transformed into *E. coli* BL21-AI.

3.4.8. Conjugation efficiency

The *E. coli* strain S17-1 containing variants of the plasmid pCONJ (Sup. Table S2) was used as the donor strain. Cells were grown to early exponential phase ($OD_{600\text{ nm}}$ of 0.2-0.3) and, if necessary, induced with IPTG and L-arabinose, and grown for 90 min at 37 °C and 180 rpm. Approximately 5×10^8 cells of both donor strain and the DISARM or wild-type strain (recipient) were spun down and re-suspended in 5 ml of fresh LB media. The strains were combined in equal counts in a final concentration of 1×10^8 CFU/mL. After gently mixing, cells were pelleted at $2,000 \times g$ for 10 min at room temperature, and incubated at 37°C for 4 h without shaking. The cell mixture was plated onto LBA plates containing different antibiotics to determine the proportion of recipient cells that acquired the plasmid from the donor strain. Conjugation efficiency was estimated as the ratio of plasmid-acquisition events versus the total number of recipient strain cells.

3.4.9. Statistical analysis

The average values of three biological replicates were reported in the result and supplementary sections. Unpaired two-tailed t Test and one-way analysis of variance (ANOVA) with Dunnett's post-hoc multiple comparison test were used to compare the means between groups. Confidence intervals were set at 95% (* = $p < 0.05$; ** = $p < 0.001$; *** = $p < 0.0001$). Statistical analysis was performed using GraphPad Prism version 5.0 for Windows.

3

3.4.10. Data availability

Raw sequencing data were deposited in Figshare and can be accessed at doi.org/10.6084/m9.figshare.17295215 for PacBio data and doi.org/10.6084/m9.figshare.17295212 for Bisulfite data.

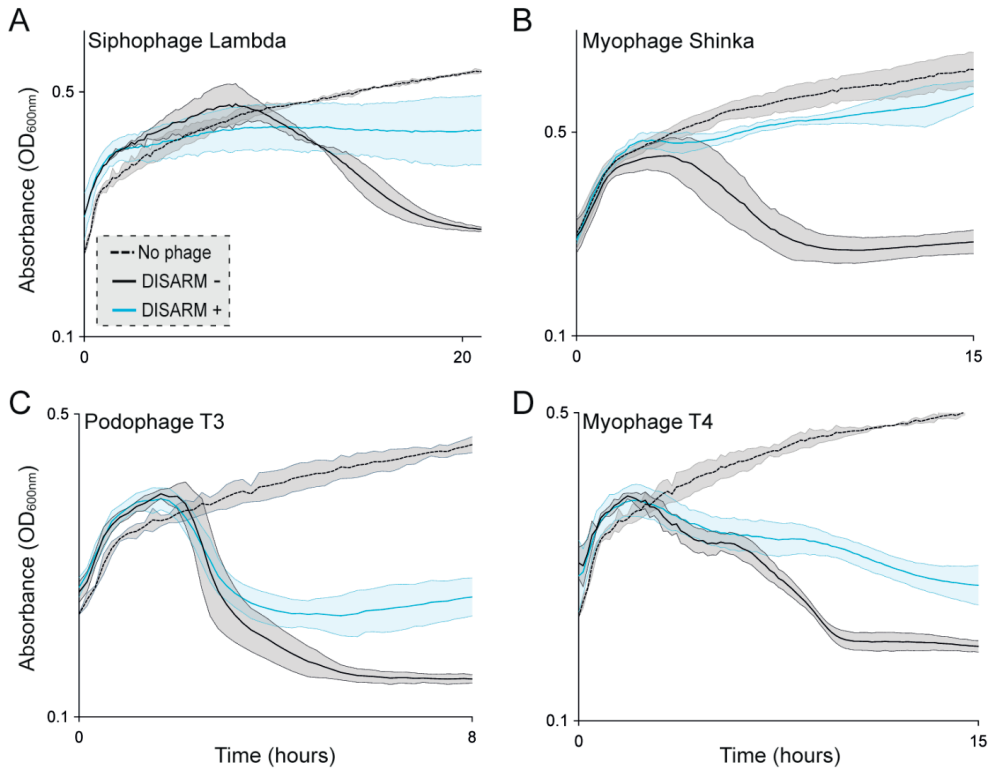
3.5. References

1. Labrie, S.J., Samson, J.E. and Moineau, S. (2010) Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.*, 8.
2. Bernheim, A. and Sorek, R. (2020) The pan-immune system of bacteria: antiviral defence as a community resource. *Nat. Rev. Microbiol.*, 18.
3. Tesson, F., Hervé, A., Touchon, M., Humières, C. and Cury, J. (2021) Systematic and quantitative view of the antiviral arsenal of prokaryotes. *bioRxiv*.
4. Tock, M.R. and Dryden, D.T.F. (2005) The biology of restriction and anti-restriction. *Curr. Opin. Microbiol.*, 8.
5. Stewart, F.J., Panne, D., Bickle, T.A. and Raleigh, E.A. (2000) Methyl-specific DNA binding by McrBC, a modification-dependent restriction enzyme. *J. Mol. Biol.*, 298.
6. Loenen, W.A.M. and Raleigh, E.A. (2014) The other face of restriction: Modification-dependent enzymes. *Nucleic Acids Res.*, 42.
7. Doron, S., Melamed, S., Ofir, G., Leavitt, A., Lopatina, A., Keren, M., Amitai, G. and Sorek, R. (2018) Systematic discovery of antiphage defense systems in the microbial pangenome. *Science*, 80.
8. Gao, L., Altae-Tran, H., Böhning, F., Makarova, K.S., Segel, M., Schmid-Burgk, J.L., Koob, J., Wolf, Y.I., Koonin, E. V. and Zhang, F. (2020) Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science*, 80.
9. Kronheim, S., Daniel-Ivad, M., Duan, Z., Hwang, S., Wong, A.I., Mantel, I., Nodwell, J.R. and Maxwell, K.L. (2018) A chemical defence against phage infection. *Nature*.
10. Swarts, D.C., Jore, M.M., Westra, E.R., Zhu, Y., Janssen, J.H., Snijders, A.P., Wang, Y., Patel, D.J., Berenguer, J., Brouns, S.J.J., et al. (2014) DNA-guided DNA interference by a prokaryotic Argonaute. *Nature*, 507.
11. Goldfarb, T., Sberro, H., Weinstock, E., Cohen, O., Doron, S., Charpak-Amikam, Y., Afik, S., Ofir, G., Sorek, R., Charpak-Amikam, Y., et al. (2015) BREX is a novel phage resistance system widespread in microbial genomes. *EMBO J.*, 34.
12. Cohen, D., Melamed, S., Millman, A., Shulman, G., Oppenheimer-Shaanan, Y., Kacen, A., Doron, S., Amitai, G. and Sorek, R. (2019) Cyclic GMP–AMP signalling protects bacteria against viral infection. *Nature*, 574.
13. Millman, A. (2020) Bacterial retrons function in anti-phage defense.
14. Rousset, F., Dowding, J., Bernheim, A., Rocha, E.P.C. and Bikard, D. (2021) Prophage-encoded hotspots of bacterial immune systems. *bioRxiv*.
15. Bernheim, A., Millman, A., Ofir, G., Meitav, G., Avraham, C., Shomar, H., Rosenberg, M.M., Tal, N., Melamed, S., Amitai, G., et al. (2021) Prokaryotic viperins produce diverse antiviral molecules. *Nature*, 589.
16. Ofir, G., Melamed, S., Sberro, H., Mukamel, Z., Silverman, S., Yaakov, G., Doron, S. and Sorek, R. (2018) DISARM is a widespread bacterial defence system with broad anti-phage activities. *Nat. Microbiol.*, 3.
17. Nobrega, F.L., Vlot, M., Jonge, P.A., Dreesens, L.L., Beaumont, H.J.E., Lavigne, R., Dutilh, B.E. and Brouns, S.J.J. (2018) Targeting mechanisms of tailed bacteriophages. *Nat. Rev. Microbiol.*, 16.

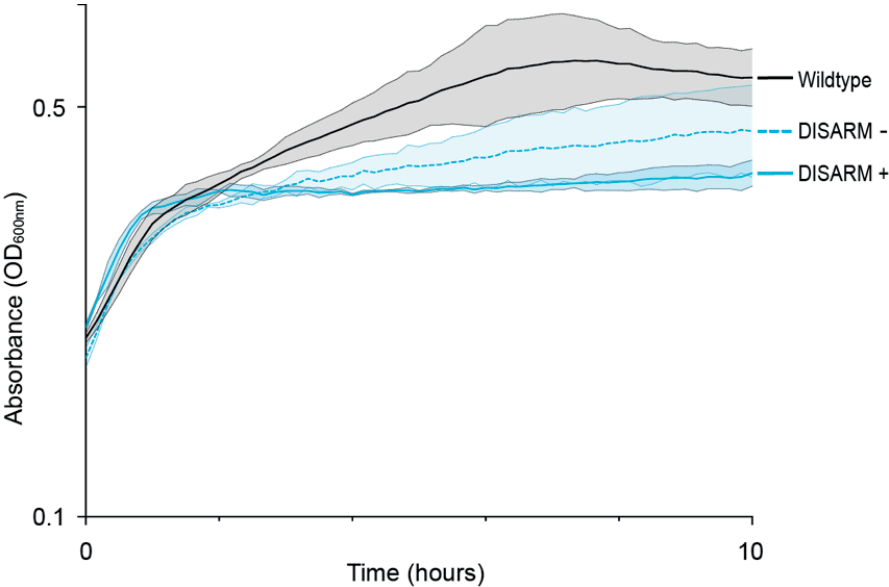
18. Howard-Varona,C., Hargreaves,K.R., Abedon,S.T. and Sullivan,M.B. (2017) Lysogeny in nature: Mechanisms, impact and ecology of temperate phages. ISME J., 11.
19. Venclovas,Č. (2019) Genome Replication of Bacterial and Archaeal Viruses. In.
20. Trun,N. and Trempey,J. (2009) Fundamental Bacterial Genetics Wiley.
21. Gordeeva,J., Morozova,N., Sierro,N., Isaev,A., Sinkunas,T., Tsvetkova,K., Matlashov,M., Truncaite,L., Morgan,R.D., Ivanov,N. V., et al. (2019) BREX system of *Escherichia coli* distinguishes self from non-self by methylation of a specific DNA site. Nucleic Acids Res., 47.
22. Wilson,G.G. and Murray,N.E. (1991) Restriction and Modification. Annu. Rev. Genet., 25.
23. Pleška,M. and Guet,C.C. (2017) Effects of mutations in phage restriction sites during escape from restriction–modification. Biol. Lett., 13.
24. Murphy,J., Mahony,J., Ainsworth,S., Nauta,A. and van Sinderen,D. (2013) Bacteriophage orphan DNA methyltransferases: Insights from their bacterial origin, function, and occurrence. Appl. Environ. Microbiol., 79.
25. Pleška,M., Qian,L., Okura,R., Bergmiller,T., Wakamoto,Y., Kussell,E. and Guet,C.C. (2016) Bacterial autoimmunity due to a restriction-modification system. Curr. Biol., 26.
26. Bravo,J.P.K., Aparicio-Maldonado,C., Nobrega,F.L., Brouns,S.J.J. and Taylor,D.W. (2022) Structural basis for broad anti-phage immunity by DISARM. Nat. Commun., 13.
27. Shen,B.W., Quispe,J.D., Luyten,Y., McGough,B.E., Morgan,R.D. and Stoddard,B.L. (2021) Coordination of phage genome degradation versus host genome protection by a bifunctional restriction-modification enzyme visualized by CryoEM. Structure, 29.
28. Fomenkov,A., Sun,Z., Murray,I.A., Ruse,C., Mcclung,C., Yamaichi,Y., Raleigh,E.A. and Roberts,R.J. (2020) Plasmid replication-associated single-strand-specific methyltransferases. Nucleic Acids Res., 48.
29. Bonilla,N., Rojas,M.I., Netto Flores Cruz,G., Hung,S.-H., Rohwer,F. and Barr,J.J. (2016) Phage on tap-a quick and efficient protocol for the preparation of bacteriophage laboratory stocks. PeerJ, 4.
30. Clokie,M.R.J. and Kropinski,A.M. (2009) Bacteriophages. Methods and Protocols Volume 1: Isolation, Characterization, and Interactions.
31. Sambrook,J. and Russell,D. (2011) Purification of Nucleic Acids by Extraction with Phenol:Chloroform. Cold Spring Harb. Protoc., 2006.
32. McAlister,V., Zou,C., Winslow,R.H. and Christie,G.E. (2003) Purification and *in vitro* characterization of the *Serratia marcescens* NucC protein, a zinc-binding transcription factor homologous to P2 Ogr. J. Bacteriol., 185.
33. Feng,Z., Fang,G., Korlach,J., Clark,T., Luong,K., Zhang,X., Wong,W. and Schadt,E. (2013) Detecting DNA Modifications from SMRT Sequencing Data by Modeling Sequence Context Dependence of Polymerase Kinetic. PLoS Comput. Biol., 9.
34. Felix Krueger and Simon R. Andrews. (2011) Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics. Vol 27.
35. Crooks, GE, Hon,G, Chandonia, JC, and Brenner SE. (2004). WebLogo: a sequence logo generator. Genome Res. 14

3.6. Supplementary material

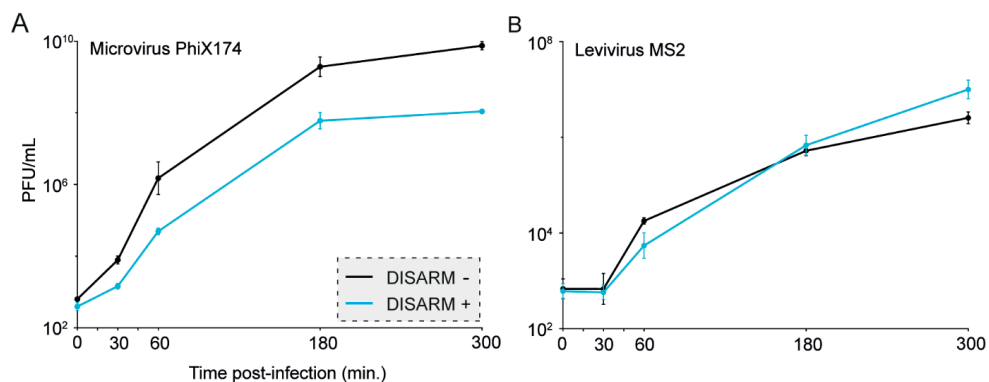
3.6.1. Supplementary figures



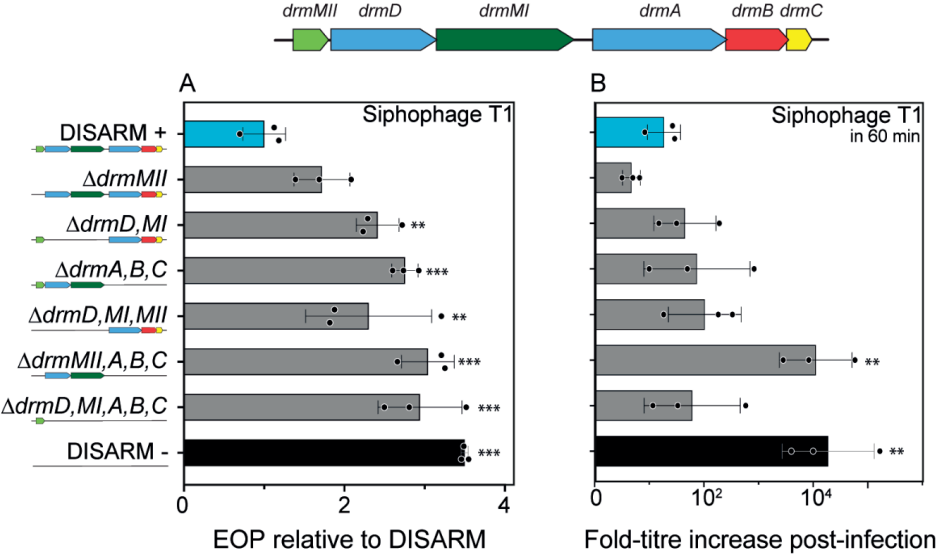
Supplementary Figure S1. Effect of phages on the growth curve of DISARM (+) or DISARM (-) strains. (A) Siphophage Lambda, (B) Myophage Shinka, (C) Podophage T3, and (D) Myophage T4. Lines and filled areas indicate average and standard deviation of three independent replicates, respectively.



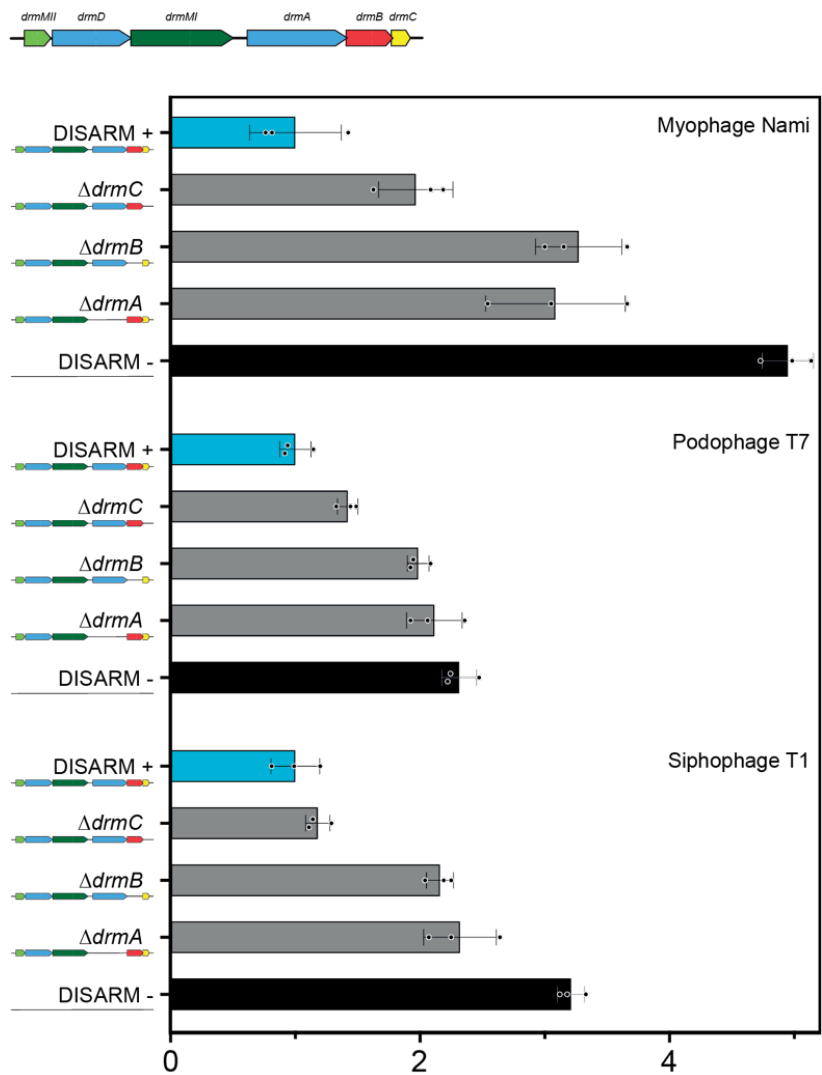
Supplementary Figure S2. Growth of DISARM induced (DISARM +) and non-induced (DISARM -) strains compared to the wild-type strains. Line and filled areas within dotted lines indicate average and standard deviations of three independent replicates, respectively.



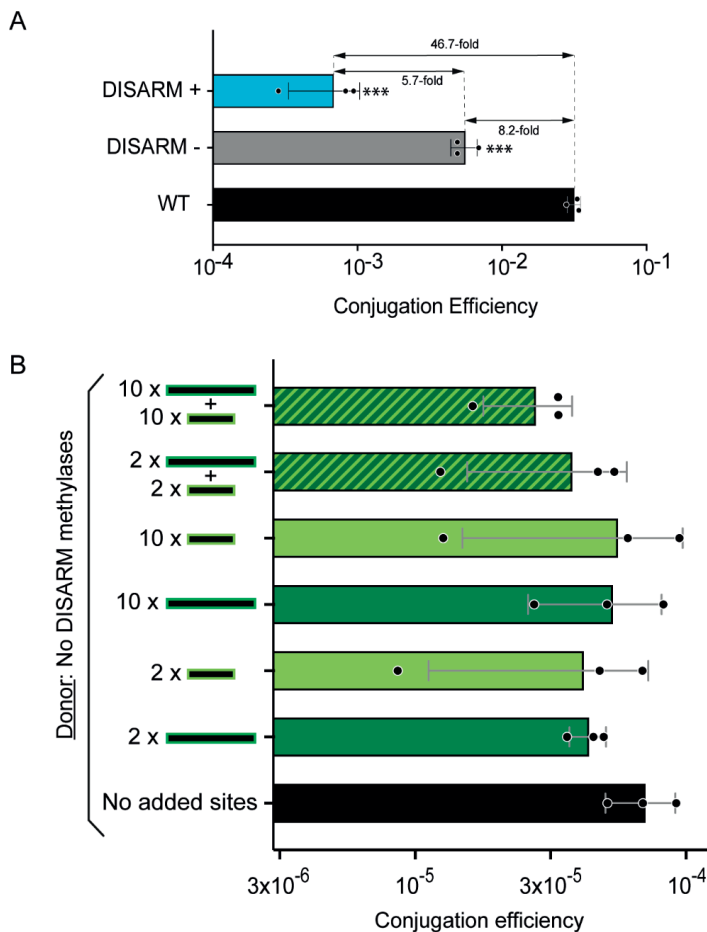
Supplementary Figure S3. Effect of Class 1 DISARM on phage replication over time. (A) Phage phiX174. (B) Phage MS2. Bacterial cultures of DISARM (+) or DISARM (-) strains were infected with phage at an MOI of 10^{-6} and the titer determined at selected time points.



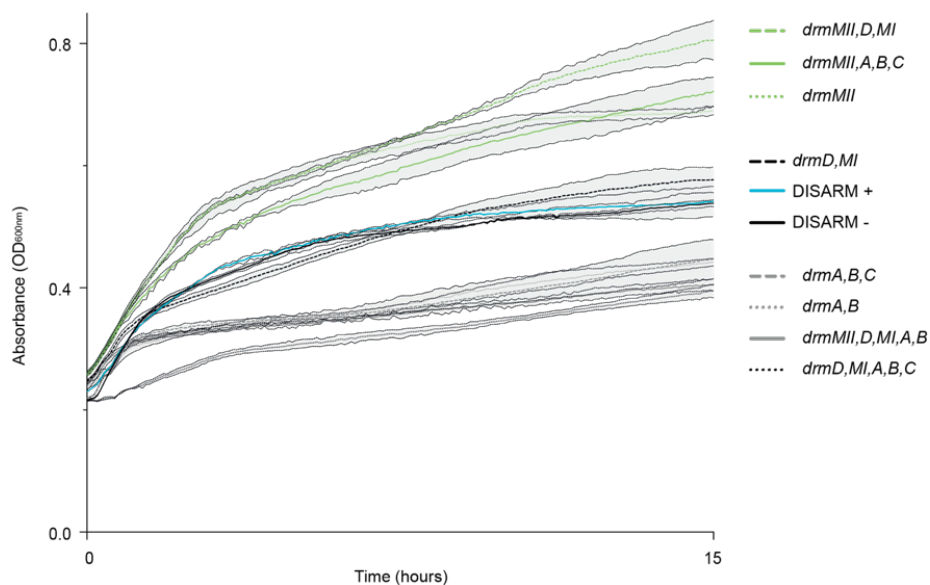
Supplementary Figure S4. Effect of Class 1 DISARM components on protection against T1 phage infection. (A) Efficiency of plating (EOP) of siphophage T1 on strains containing the modified DISARM system (DISARM (-) strain), normalized to the DISARM strain. (B) Titer-fold increase of siphophage T1 upon propagation in cultures of strains containing the complete or modified DISARM system. Graphic represents the time point at which maximum effect on phage replication was achieved. Bars depict the average and standard deviation of three independent replicates. Statistical significance was determined by one-way ANOVA+ with Tukey post-hoc test and is represented as *, **, or *** for $p < 0.05$, 0.01 or 0.001, respectively.



Supplementary Figure S5. Effect of Class 1 DISARM components on protection against phage infection. Efficiency of plating (EOP) of myophage Nami (top), podophage T7 (mid), and siphophage T1 (bottom) on strains containing the DISARM system with deletions of genes *drmA*, *drmb*, *drmc*; normalized to the DISARM (+) strain.



Supplementary Figure S6. Conjugation efficiency of plasmid pCONJ with variable amount of DISARM motifs. (A) Conjugation efficiency of plasmid pCONJ without motifs in DISARM induced (DISARM+), non-induced (DISARM -) and control strains. (B) Conjugation efficiency of plasmid pCONJ with variable amounts of unmethylated motifs ACACAG (dark green) and/or MTCGAK (light green) into the recipient control strain.



Supplementary Figure S7. Effect of DISARM on bacterial growth. Lines and filled areas within dotted lines indicate average and standard deviations of three independent replicates, respectively.

3.6.2. Supplementary tables

Supplementary Table S1. Bacteriophage features and host strains.

Phage	Replication strain	Infection type	Genetic material	Relevant genotype	Source
T1	BL21	Lytic cycle	dsDNA	dam+, dcm-	DSMZ
T3	BL21	Lytic cycle	dsDNA	dam-, dcm-, ocr+ (demethylase)	DSMZ
T4	BL21	Lytic cycle	dsDNA	dam+, dcm-, hmc+, glc-hmc+	Elizabeth Kutter's Lab
T7	BL21	Lytic cycle	dsDNA	ocr+ (demethylase)	Ian Molineux's Lab
Lambda-vir	BL21	Lytic cycle	dsDNA	dam-, dcm-	Elizabeth Kutter's Lab
Nami	Ci2	Lytic cycle	dsDNA	dam-, dcm-, hmc+	Brouns Lab collection
Shinka	Ci2	Lytic cycle	dsDNA	dam+, dcm-, hmc+	Brouns Lab collection
PhiX174	C3000	Lytic cycle	ssDNA	dam-, dcm-	DSMZ
M13	JM109	Chronic	ssDNA	dam-, dcm-	New England Biolabs
MS2	WG5	Lytic cycle	ssRNA	dam-, dcm-	DSMZ

Supplementary Table S2. List of plasmids used in this work.

Name in this study	Plasmid name	Insert	Backbone	Marker	Size (bp)
pDIS_1	pTU497	drmMII	pET52_duet	Ampicillin	6,165
pDIS_2	pTU496	drmD, drmMI	pCOLA_duet	Kanamycin	10,818
pDIS_3	pTU495	drmA, drmB, drmC	pCDF_duet	Streptomycin	10,276
pDIS_4	pTU501	drmMII, drmMI	pCOLA_duet	Kanamycin	8,777
pDIS_5	pTU498	drmA, drmB, drmC*	pCDF_duet	Streptomycin	10,285
pDIS_6	pTU499	drmA, drmB*, drmC	pCDF_duet	Streptomycin	10,285
pDIS_7	pTU500	drmA*, drmB, drmC	pCDF_duet	Streptomycin	10,285
pCONJ_0	pTU503	None	pSEVA331	Chloramphenicol	4,522
pCONJ_1	pTU504	2x ACACAG	pSEVA331	Chloramphenicol	4,517
pCONJ_2	pTU505	2x MTCGAK	pSEVA331	Chloramphenicol	4,522
pCONJ_3	pTU506	10x ACACAG	pSEVA331	Chloramphenicol	4,524
pCONJ_4	pTU507	10x MTCGAK	pSEVA331	Chloramphenicol	4,519
pCONJ_5	pTU508	2x ACACAG +2x MTCGAK	pSEVA331	Chloramphenicol	4,519
pCONJ_6	pTU511	10x ACACAG +10x MTCGAK	pSEVA331	Chloramphenicol	4,519

Supplementary Table S3. List of primers used in this work. Underlined the restriction site used for each primer. fw, forward primer; rv, reverse primer.

Name	Sequence (5'-3')	Description	fw/rv
BN861	TAAGGAGATATACCATGACTGATAACAACAAATCTAG	Clone <i>drmA,B,C</i> in pCDF-Duet	fw
BN1585	CTGCGCTAGTAGACGAGCTTAGCCAAGAATCAAAACGATCGC		rv
BN860	GGTAAGGAGATATACCATGGAAATTCTAATACACCTGATATCC	Clone <i>drmD,MI</i> in pCOLA-Duet	fw
BN2313	GCCTAGGTTAATTAAGCTGGCAGCAGCCTAGGTTAATTAAGCTGCGCTAG		rv
BN1219	ACTTTAAGAAGGAGATATACCATGAAGAAAGTATCTTGTGTCG	Clone <i>drmMII</i> in pET-Duet	fw
BN2082	GCAGCAGCCTAGGTTAATTAGTCATTCCACAGTGACCATATTCAC		rv
BN1052	ATTTTGTTTAACTTTAAGAAGGAGATATACCATGAAGAAAGTATCTTGTGTCG	Clone <i>drmMII</i> in pCOLA-Duet	fw
BN1053	TAGTTATTGCTCAGCGGTGGCAGCAGCCTAGGTTAATTAATTCATTCCACAGTGACCATATTC		rv
BN1056	TATATATACATATGATGAGTATGGCCAGCATAACG	Clone <i>drmMI</i> in pCOLA-Duet	fw
BN1057	ATATATATCTCGAGTTAGTTTCGCACCATCAGACAAATC		rv
BN1539	ACTTACATTTCGCGAACTGCTG	Insert early stop codon in <i>drmA</i> of pTU495 to create pTU498	fw
BN1540	TTATCATTATGATCCTCGATCTTTGGTATGTCTG		rv
BN1542	GCTGCTGTGCGAAAGGTATTAG	Insert early stop codon in <i>drmB</i> of pTU495 to create pTU499	fw
BN1543	TTATCATTAAGAAGCCGGCCTCTTGA		rv
BN1545	GAAAAGATACAAGCGATCGCTGGC	Insert early stop codon in <i>drmC</i> of pTU495 to create pTU500	fw
BN1546	TTATCATTAAGGAGAAACCAGACAACTAGCTC		rv
BN2208	ATATATATGGATCCGCTCTATCGGCATATGGTCAA	Amplify No-site module to clone with BamHI	fw
BN2209	ATATATATGGATCCATATATCCTTTGCCCCGTCC		rv
BN2218	ATATATATGGTACCGTCTATCGGCATATGGTCAA	Amplify No-site module to clone with KpnI and NotI	fw
BN2393	ATATATATGCGCCCGCATATATCCTTTGCCCCGTCC		rv
BN2228	ATATATATACTAGTCCGCATGGGCAATACGAGGT	Amplify 2 A-site module to clone with SpeI and BamHI	fw
BN2248	ATATATATGGATCCAGCGCATGGGTGACGGTTT		rv
BN2230	ATATATATACTAGTTAACCTAACGGTAAGAGGCT		fw

Characterization of Class 1 DISARM as bacterial defense system

BN2250	ATATATAT <u>GGATCC</u> TAAAGATTCATGTTAGTCCA	Amplify 10 A-site module to clone with SpeI and BamHI	rv
BN2228	ATATATATACTAGTCCGCATGGGCAATACGAGGT	Amplify 2 A-site module to clone with SpeI and SphI	fw
BN2396	ATATATATGCATGCAGCGGCATGGGTGACGGTTT		rv
BN2230	ATATATATACTAGTTAACCTAACGGTAAGAGGCT	Amplify 10 A-site module to clone with SpeI and SphI	fw
BN2397	ATATATATGCATGCTAAAGATTCATGTTAGTCCA		rv
BN2220	ATATATATGGTACCTTTGAATACCTTATATTATT	Amplify 2 C-site module to clone with KpnI and NotI	fw
BN2394	ATATATATGCGGCCGCCTGGTGACCGGTAGGTGTAC		rv
BN2222	ATATATATGGTACCGCGGATGGCGTTACAACCTTT	Amplify 10 C-site module to clone with KpnI and NotI	fw
BN2395	ATATATATGCGGCCGCACGGTCAGATTAGGCCCATG		rv
BN2214	ATATATATGGATCCTAACCGGCAATAATGCGC	Amplify pCONJ to insert modules using BamHI	rv
BN2215	ATATATATGGATCCTACCAAATGCGGGACAAC		fw
BN2226	ATATATATGGTACCGAACATTGCGTTGCCTTG	Amplify pCONJ to insert modules using KpnI and NotI	rv
BN2390	ATATATATGCGGCCGCAATGATGTCGTGGAACGTC		fw
BN2234	ATATATATACTAGTGAACATTGCGTTGCCTTG	Amplify pCONJ to insert modules using SpeI and SphI	rv
BN2391	ATATATATGCATGCAATGATGTCGTGGAACGTC		fw
BN2232	ATATATATACTAGTTAACCGGCAATAATGCGC	Amplify pCONJ to insert modules using SpeI and BamHI	rv
BN2215	ATATATATGGATCCTACCAAATGCGGGACAAC		fw
BN2241	TATATATATAAGGCGCGCCGAAGGCGTG	Amplify pSEVA331 backbone to clone Gblocks using AgeI and AscI	rv
BN2242	TATATATATAACCGGTTATACCACCGTTGATATATCCCAATGGC		fw
BN2243	ATAATATCGGCGCGCCTC	Amplify Gblock_1 to clone using AscI and HindIII	fw
BN2244	GACAGTCTGACAAGCTTCG		rv
BN2245	TATATATAGCGAAGCTTGGATAATTCC	Amplify Gblock_2 to clone using HindIII and AgeI	fw
BN2246	TGAAAATAACCGGTGATTTTTTCTC		rv

3

Supplementary Table S4. List of synthetic constructs used in this work, with cloning sites underlined and DISARM motifs in bold. Motifs in *italic* represent complementary sites for the non-palindromic site ACACAG.

Name	Description	Sequence
Gblock_1	Sequence modification to replace putative DISARM motifs in backbone sequence. Cloned using <i>AscI</i> and <i>HindIII</i>	ATAATATCGGCGCGCCTCCCCCTCCGGCAAAAAGTGGCCCTCCGGGGCTTG TTGATGGACTGCGCGGCCCTTCGGCCTTGCCCAAGGTGGCGCTGCCCCCTTGGA ACCCCCGCACTCGCCGCGCTGAGGCTCGGGGGGAGGGCGGGGGCTTCGCC TTGGACTGCCCCCACTCGCATAGGCTTGGGTCTGTTCCAGGCGCGTCAAGGCCA AGCCGCTGCGCGGTGCTGCGCGAGCCTTGACCCGCTTCCACTTGGTGTCCA ACCGGCAAGCGAAGCGCGCAGGCCCGAGGCCGAGGCTTTTCCCCAGAGAAA TTAAAAAAATTGATGGGGCAAGGCCGCGAGGCCGCGAGTTGGAGCCGGTGGGT ATGTGGTAGAAGGCAGGGTAGCCGGTGGGCAATCCCAGTGGTCAAGCTCGTGG GCAGGCGCAGCCTTCCATCAGCTTGTCCAGCAGGGTTGTCCACGGGCCGAGC GAAGCGAGCCAGCCGGTGGCGCTCGCGGCCATCGTCCACATATCCACGGGT GGCAAGGGAGCGCCGCGACCGCGCCGGCGAAGCTTGTGACACTGTCT
Gblock_2	Sequence to insert the YFP and MAM into the backbone sequence. Cloned using <i>HindIII</i> and <i>AgeI</i>	TATATATAGCGAAGCTTGGATAATTCCTAATTTTTGTTGACACTCTATCGTT GATAGAGTTATTTTACCCTCCCTATCAGTGATAGAGAAAAGAAATCAAAGA TCTAGGAGGAAAAAAATGGTGAGCAAGGGCGAGGAGTTGTTACCGGGGTGGT GCCCATCCTGGTTGAGCTGGACGGCGACGTAACCGGCCACAAATTTTCCGTGT CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATC TGCACCAACGGCAAGCTGCCCGTGCCTTGGCCACCCCTCGTGACCACCTTGGG CTACGGCCTGCAGTGCTTCGCCCGCTACCCCGACCATGAAGCAACACGACT TCTTCAAGTCCGCCATGCCGAAGGCTACGTCAGGAGCGCACCATCTTCTTC AAGGACGAGGCAACTACAAGACCCGCGCGAGGTGAAGTTTGAGGCGACAC CCTGGTGAACCGCATTTGAAGTGAAGGGCATTGACTTCAAGGAGGACGGCAACA TCTTGGGGGCCCAAGCTGGAAGTACAACATAACCGCCACCACGTCCATATCA CCGCCGACAAGCCGAAGAAGCGCATCAAGGCCAACTTCAAGATCCGCCACCAC CTCAAGGACGGCGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCAT CGGCGACGGCCCCGTGCTGCTGCGCGACCACTACCTAAGTAAGGATCTCC AGGCATCAATAAAACGAAAGGCTCAGTTGAAAGACTGGGCCTTTCCGTTTTAT GAGCAAGCCCGTAGGGGCCAGTCTAGCTTCAAGTATGACGGGCTGATACTGGG CCGGCAGGCGCTCCATTGCCAGTCGGCAGCGACATCCTTCGGGCGCATTATT GCCGGTTACTGCGCTGTACCAATGCGGGACACGTAAGCACTACATTTTCGCT CATCGCCAGCCAGTATCTGATGCACGGGCGCGAGTTCCATAGCGTTAAGGT TTCATTTAGCGCCTCAATAGATCCTGTTCAGGAACCGGATCAAGAGTTCTCT CCGCCGCTGGACCTACCAAGGCAACGCAATGTTCTCACGGAATGATGTCGTGG AACGTCACAATGGTGACTTCTACAGCGCGGAGAATCTCGCTCTCTCCAGGGG TTGCTTTTGTGAGCAAGAGCCAGCAGTTTAATTAAAGCGGATAACAATTTCT CTCGGGAGGCGCCTAGGCCGTCGTGACTGGGAAAACCTGGCGACGTGGGT CCCCAATAATTACGATTTAAATTGGCGAAAAATGAGACGTTGATCGGCACGTAA GAGGTTCCAACCTTTCACCATAATGAAATAAGATCACTACCGGGCGTATTTTT GAGTTATGGAGATTTTCAGGAGCTAAGGAAGCTAAATGGAGAAAAAATCAC CGTTATTTTCA
Gblock_3	Module containing No motifs (for control)	GTCTATCGGCATATGGTCAATTGAGTGTCGGGAGGCGGAAATCCGCCACGA ATGAGAATGTATTTCCCGACAATCATAATGGGGCGCTCCTAAGCTTTTCCAT TGTTGGGCGCGCTAGGCCATCATCTGCGCGAGTTTCGGCGCACTGCTGCCG ACATGCCGGGCATTGTTTTAGGGGCGTTATTCTTGAGGGCACTCGGAGCTAAT TTGTGCGGACCGCGGGGTAGTCATCGGGCTTATACACGAAAGGCCATGTA TCCGTTTCCAGTATGGAACGTCTTTAGCTCCGGCAAGCAATTAAAGAACACG CAAGCGGGACGGCGAAAGGATATAT
Gblock_4	Module containing 2 C motifs	TTTGAATACCTTATATTATCATCGCGGATATAACATGAGAAACGGCCGAA TACACCATGTTTCGTATCGTATCGGTAAATACCTCGCGTAGCCATGTGCCATAT GGTTGCGATCGATTGGTTATGCATATGGTCCACATGGACACTCTTCGCTTCCG GGTATGCGCTATATGTGACGGTCTTTAGGCGCACTGATGCTCATGCTCGAGTT AAACCAACCGACACCAGATCATGTAAGGTCCGCCACGCATGACGACATGCCCA CGGAGATCACCGACCGATCTATCTGATCGGCGACCATTTGTGTGTTACTGGGG CGGAGGTGTTGACCGGTAGGTGTAC

Characterization of Class 1 DISARM as bacterial defense system

Gblock_5	Module containing 10 C motifs	GCGGATGGCGTTACAACCTTCCTTAATGGCTCTTGGGCCGCGGTGCGTGACC TTGCAGGAATTGACTCGAGCCGTTAATTTCCCTTGCATACAATCGATGTTTTT TTGTCTTTTATCCGCTTCTCGAGATAAGAGTGACATACTTCTTAATCGATCG CCTCCGTACACATGTACGATCGCTCGAGCCATGAGAATAGGTATACCATGTAT CGATGTGAGCAACGAAAGCCTAAACGGGACTCGAGCGGCCAAAAGTCGGTGCG AATTCGAGTCATCGATATGTTGGTCTGGCTATGATCTACCTCGAGCAGCGGTT ACGACGGTCAGATTAGGCCCATG
Gblock_6	Module containing 2 A motifs	CCGCATGGGCAATACGAGGTCGTCTGCTCTGGTCAGCCTCTAATGGCTCGTT AGATAGTCTAGCCGCTGGTAATCACATGACCTCGTCTCCCCATTGGTGCTACG GCGATTCTTGAGACACAGGCTGCGATCGCTAATGTGAGGACATGTGTAATAT TAGCCTGTGTTAAGTCCCAATGGTTGTGGCCTTTTGAAAAGTGAACCTCATA ACATATGCATGTCTCACGCACATGGATGGTTTGGACAAATTGATTCAAGTCT GATCAACCTTCACACATGATCTAGAATCAAAAGCATGATCTCCCGGGTGCGAA ATAAGCGGCATGGGTGACGGTTT
Gblock_7	Module containing 10 A motifs	TAACCTAACGGTAAGAGGCTCAAAATACTACGTAACACACAGGACTGCGACG TTCTAAACTGTGTTCCGTCTGAACCGCCATCCATGGATCACACACAGCCGAAA AAAAGATATCAGGAATCTGTGTCTCACATCGGTATATGAAAATACATGGA CACAGCTTCCTGGCAACGGGGGGTGGTAATCCGCTGTGATGAGAAGGTATT TGCTCAATAATCAACACAGCAGGCATCTAACTTTTCCCATGCCCTGTGTCGGC TTGCCCATTTTGCCATGTAACACAGTTAGGACTCGCGCCAACGCGCACTGTGT ATTCGGTAAAGATTCATGTTAGTCCA

3

Supplementary Table S5. Methylated sites of the strains, determined by methylation-sensitive sequencing. m6A refers to ACACAG sites and 5mC to MTCGAK sites.

Strain	DISARM genes	Number of m6A sites detected	Ratio of m6A sites (%)	Number of 5mC sites detected	Ratio of 5mC sites (%)	Number of GmATC sites detected	Ratio of GmATC sites (%)
WT	-	0	0	0	0	37377	99.9
DISARM	drmMII, drmD, drmMI, drmA, drmB, drmC	929	84.3	1201	67	37359	99.9
DrmMII-DrmMI	drmMII, drmMI	815	73.9	794	36	37374	99.9
DrmMII	drmMII	0	0	234	34	37375	99.9
DrmMI	drmMI	619	56.1	0	0	37375	99.9

Supplementary Table S6. Methylation motifs found in phages and bacterial strains.

Phage/ bacterial strain	Genome size (bp)	Number of ACACA G sites	% ACAC AG sites per 100 kb	Expect ed ACAC AG sites	Number of MTCG AK sites	% MTCGAK sites per 100 kb	Expect ed MTCG AK sites	Numbe r of TCGA sites	% TCGA sites per 100 kb	Expect ed TCGA sites
T1	48,836	12	147	12	10	164	12	202	1,655	191
T3	38,209	12	188	9	8	167	9	134	1,403	149
T4	168,903	36	128	41	24	114	41	312	739	660
T7	39,937	18	270	10	8	160	10	111	1,112	156
Lambda-vir	48,502	19	214	12	10	165	12	121	998	189
Nami	165,679	33	120	40	26	126	40	311	751	647
Shinka	162,160	37	137	40	21	104	40	281	693	633
M13	7,222	1	83	2	2	222	2	11	1,218	28
PhiX174	5,386	1 (rev ^a)	111	1	1	149	1	10	743	21
<i>Serratia</i> sp. SCBI	5,034,688	717	85	1,229	1,399	222	1,229	24,114	1,916	19,667
<i>Serratia</i> plasmid	67,208	19	170	16	37	440	16	263	1,565	263
<i>Serratia</i> ATCC 39006	4,971,757	1783	221	1,214	946	152	1,214	13,376	1,076	19,421
<i>E. coli</i> Dh5 α	4,583,637	1,132	148	1,119	1,001	175	1,119	15,292	1,334	17,905
<i>E. coli</i> BL21- AI	4,530,564	1,106	146	1,106	993	176	1,106	14,937	1,319	17,698
<i>E. coli</i> K12 JM109	4,641,652	1,150	149	1,133	1,013	175	1,133	15,462	1,332	18,131
<i>E. coli</i> JM109(DE3)	4,661,885	1,159	149	1,138	108	19	1,138	15,409	1,322	18,210
<i>E. coli</i> C3000	4,617,024	1,134	147	1,127	1,010	175	1,127	15,575	1,345	18,035
<i>E. coli</i> WG5	4,592,887	1,127	147	1,121	1,002	175	1,121	15,449	1,345	17,941

^arev: site present in complementary strand of the ssDNA phage genome



4.

Structural characterization of Class 1 DISARM mechanism

Published as: *Structural basis for broad anti-phage immunity by DISARM.*
Cristian Aparicio-Maldonado*, Jack P.K. Bravo*, Franklin L. Nobrega, Stan J.J. Brouns,
David W. Taylor (2022). **Nature Communications**. doi: 10.1038/s41467-022-30673-1

Abstract

In the evolutionary arms race against phage, bacteria have assembled a diverse arsenal of antiviral immune strategies. While the recently discovered DISARM (Defense Island System Associated with Restriction-Modification) systems can provide protection against a wide range of phage, the molecular mechanisms that underpin broad antiviral targeting but avoiding autoimmunity remain enigmatic. Here, we report cryo-EM structures of the core DISARM complex, DrmAB, both alone and in complex with an unmethylated phage DNA mimetic. These structures reveal that DrmAB core complex is autoinhibited by a trigger loop (TL) within DrmA and binding to DNA substrates containing a 5' overhang dislodges the TL, initiating a long-range structural rearrangement for DrmAB activation. Together with structure-guided *in vivo* studies, our work provides insights into the mechanism of phage DNA recognition and specific activation of this widespread antiviral defense system.

4.1. Introduction

The threat of bacteriophage has driven bacteria to evolve a myriad of antiviral defense systems, capable of targeting phage at various stages of infection and replication ^{1,2}. The first lines of antiviral defense are typically rapid, broad, innate immune responses that include preventing phage adsorption ³, degrading foreign DNA using restriction-modification (RM) systems ⁴, or ultimately host cell destruction through abortive infection (Abi) systems ¹. Adaptive immune responses (including CRISPR-Cas systems) provide a highly-specific and long-term antiviral protection ^{5,6}. A number of recently discovered anti-phage systems have expanded the arsenal of defense tools used by microbes in the arms race against phage ⁷⁻⁹.

RM systems are the most highly abundant bacterial defense system, found in ~75% of all bacterial genomes ¹⁰. They often consist of DNA methyltransferase, restriction endonuclease, and target recognition modules. The vast majority rely on sequence-specific DNA methylation, which is recognized by and rapidly degraded by specific nucleases, allowing for self-versus-non-self discrimination to avoid autoimmunity. However, phage have adapted to evade such systems by either evolving to lack the restriction site sequences required for DNA cleavage, or carrying their own epigenetic modification to prevent being targeted ^{4,11}.

The largely uncharacterized DISARM (Defense Island System Associated with Restriction-Modification) systems are widespread in bacteria ⁷. The DISARM operon typically contains a DNA methyltransferase (DrmMI and/or DrmMII that methylate adenine and cytosine, respectively) along with a helicase (DrmA), a DUF(domain of unknown function)1998-containing protein (DrmB), a phospholipase D (PLD) domain nuclease (DrmC), amongst additional auxiliary genes ⁷ (**Fig. 1a**). DUF1998 domains are common in various defense systems including Druantia ⁸ and Dpd ¹², suggesting a common role in antiphage activity. While the presence of methyltransferase and nuclease genes within the DISARM operon hints that this system behaves akin to other RM systems, DrmC appears to be non-essential for DISARM activity, and DISARM can still restrict phage

that lack methylation target sequences¹³. Thus, the molecular mechanisms underlying phage targeting remain unclear.

Here, we report cryo-EM structures of DrmA:Drmb (DrmAB) in the presence or absence of target DNA, revealing the arrangement of conserved RecA helicase domains relative to the DUF1998 domain. *In vivo* studies of structure-guided DrmAB mutants demonstrate that ATP hydrolysis, DNA binding, and DrmAB heterodimer formation are essential for phage targeting by DISARM. We observe that DrmA contains an unstructured trigger loop that partially occludes the DNA-binding surface on the complex. This loop limits DNA binding for dsDNA and allows DrmAB to discriminate between targets based on DNA structure rather than DNA sequence as a mechanism of self-versus-non-self-discrimination. DNA loading induces long-range conformational changes likely associated with DISARM activation. Our results elucidate how DISARM systems can provide broad antiphage activity, while avoiding autoimmune activation and reveal key molecular mechanisms that underpin rapid DISARM activation upon phage infection.

4.2. Results

4.2.1. Cryo-EM of DrmAB complex

We purified the *Serratia* sp. SCBI (*Serratia*) DrmAB complex after co-expression of DrmA and DrmB in *E. coli*. SDS-PAGE analysis of the peak fraction from size-exclusion chromatography confirmed the presence of both DrmA and DrmB, in a ~1:1 stoichiometry (**Sup. Fig. 1**). Candidate nuclease DrmC did not co-purify with DrmAB when co-expressed, supporting the notion that it does not constitute part of the core DISARM complex ⁷. To capture the DrmAB complex bound to a phage single-stranded DNA mimic, we incubated DrmAB with an unmethylated forked DNA substrate in the presence of ADP. DrmAB:ADP:DNA complex formation was confirmed by native electrophoretic mobility shift assays (**Sup. Fig. 1**).

We used cryo-EM to investigate the architecture of the DrmAB:ADP:DNA complex. During data collection we observed that the complex adopts a preferred orientation in vitreous ice, which was ameliorated through collecting additional data at a -30° tilt. Multiple rounds of classification resulted in a 3D reconstruction of the complex at a nominal resolution of 2.8 Å and focused 3D classification improved densities corresponding to DNA and the N-terminus of DrmB. Notably, we were able to separate DNA-bound from DNA-free particles through focused 3D classification, ultimately yielding 3D reconstructions of DrmAB:ADP:DNA and DrmAB:ADP at nominal resolutions of 3.3 and 3.4 Å, respectively (**Sup. Fig. 2**). Between these two structures, we determined atomic models that account for >98% of the total residues of the ~220 kDa protein components, with an additional seven nucleotides of DNA in the DrmAB:ADP:DNA complex (**Sup. Fig. 3**).

4.2.2. Architecture of DrmAB nucleoprotein complex

The DrmAB complex exhibits a bi-lobed architecture, resembling a partially open clam shell (**Fig. 1b-c**). The top lobe contains the RecA1 domain of DrmA, while the bottom lobe contains the DrmA RecA2 domain and the entirety of DrmB.

As expected for a superfamily 2 (SF2) helicase, DrmA contains tandem RecA modules (residues 600-1150) that form an active site for ATP-binding and hydrolysis and DNA loading. The N-terminal half of DrmA acts as a structural chassis, holding these two helicase motor domains in place, and forming an interface with DrmB. While RecA1 and RecA2 have high structural similarity to many other SF2 helicase domains (Z-score >14 for more than 10 other helicase structures¹⁴), no other regions of DrmA were found to have significant structural homology to known protein domains. The solvent-exposed interface between the tandem RecA domains forms the nucleotide-binding site, while the bound DNA straddles these domains within the heart of the complex (**Fig. 1b-c**). Only four of the seven deoxynucleotides observed within our structure are solvent-exposed, with the remaining three bases deeply buried within the DrmA(RecA1,RecA2):DrmB interface.

DrmB contains a C-terminal DUF1998 module, a domain often enriched within bacterial genomic defense islands^{7,15}. DUF1998 domain sits towards the back of the complex, predominantly buried within DrmB, positioned towards the interface between the two RecA domains of DrmA. Like other proteins containing DUF1998 domains, DrmB contains the predicted zinc-coordinating four cysteine motif (C559,565,581,584, **Sup. Fig. 4**).^{16,17} Surprisingly, we also observed an additional four putative coordinated metal ions at highly conserved sites within DrmB (**Sup. Fig. 4**). While it is not possible to unambiguously determine the identity of these ions, based on the coordinating residues Zn^{2+} is a strong candidate¹⁸.

This raises the possibility that DrmAB may utilize these three highly conserved metal ion coordination sites to sense changes to cellular redox potential, triggered by environmental stresses such as phage infection, as has recently been demonstrated for the type III-A CRISPR system within *Serratia*¹⁹⁻²². Alternatively, these clusters may play a structural role, enabling proper folding of DrmB, as is the case for the DUF1998-containing StfH protein¹⁷.

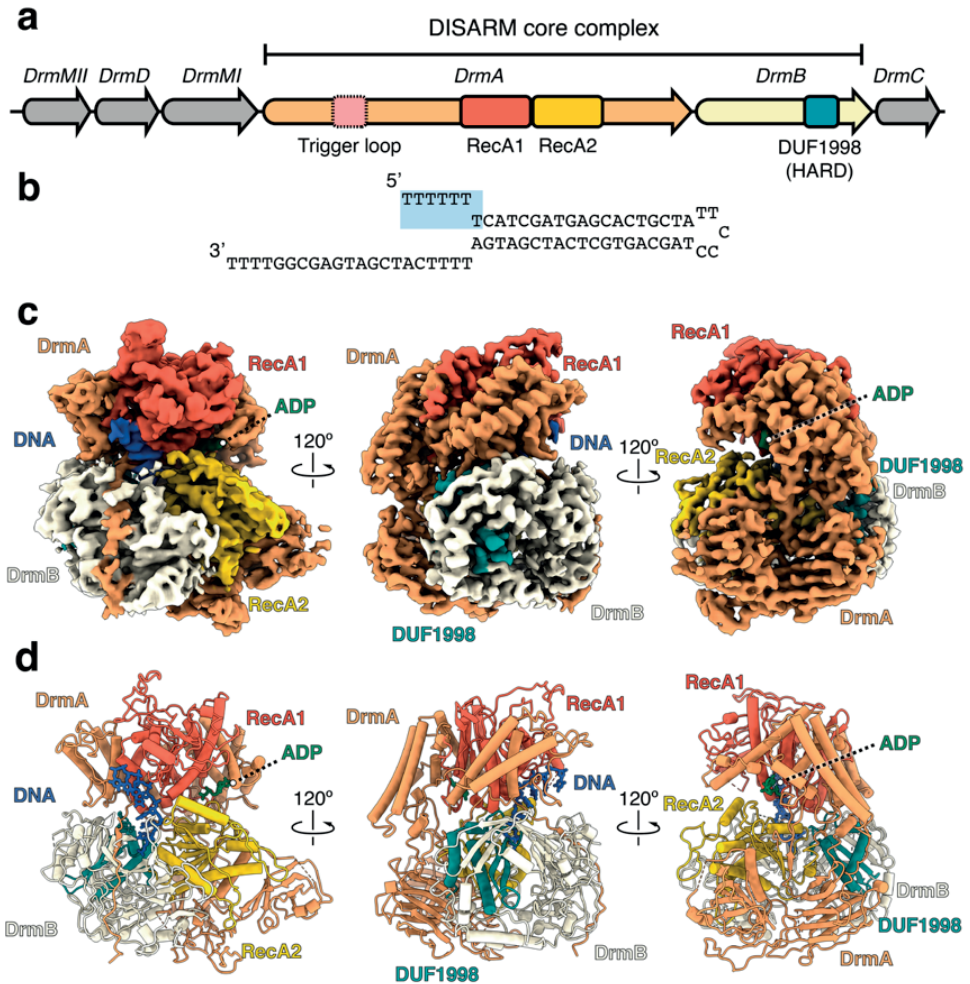


Figure 1. Architecture of the DrmAB:ADP:DNA complex. **a)** *Serratia* sp. DISARM operon. Conserved DISARM core components DrmA and DrmB are colored by structural domains. Presence of DrmA trigger loop (residues 176-232) is mutually exclusive with DNA-binding and is denoted by a dashed box. **b)** DNA substrate used for structure determination. Seven 5' bases that are present in our structure are highlighted blue. **c)** Cryo-EM reconstruction of DrmAB:ADP:DNA complex colored as in **(a)**. ssDNA is shown in blue and ADP in green. **d)** Atomic models built into the cryo-EM map **(b)**.

Multiple interaction surfaces contribute to the stabilization of the DrmAB complex, with a total buried surface area of $\sim 4,000 \text{ \AA}^2$. Most notably, the C-terminus of DrmA (residues 1296-1319) wraps around the entirety of the N-terminus of DrmB (**Fig. 1b-c**), with a total buried surface area of $\sim 1,850 \text{ \AA}^2$. Additionally, the DrmB DUF1998 domain mediates $\sim 1,230 \text{ \AA}^2$ of surface contacts with DrmA (**Fig. 1d**). DUF1998 is positioned towards the interface between the two RecA domains of DrmA but makes limited contacts with both (~ 130 and $\sim 110 \text{ \AA}^2$ buried surface area with RecA1 and RecA2, respectively). RecA2 makes additional surface contacts with the N-terminus of DrmB (482 \AA^2). Based on this network of inter-subunit interactions, we hypothesize that DUF1998-containing DrmB may act as a modulator of DrmA helicase activity. This is supported by the observation that DUF1998 domains are frequently associated with helicase domain-containing proteins, either as adjacent genes or as C-terminal fusions ¹⁵.

4.2.3. DNA binding by DrmAB

The DNA-bound DrmAB complex includes a 7-mer ssDNA segment (**Fig. 2a**). While this is only a fragment of the DNA used for complex assembly, comprising a 19-bp duplex with a 5-nucleotide hairpin on one end and a 7-nucleotide 5' tail and a 21-nucleotide 3' tail on the other end, we attributed this to the flexibility of DNA that does not make direct contact with the complex. We tested [the possibility of DrmAB requiring a 7-nt overhang to bind DNA] this by measuring the binding of DrmAB to DNA substrates with various 5' overhang lengths. We observed binding of DrmAB to a 7-nt 5' ssDNA overhang, and minimal binding to substrates with fewer than 7-nt overhangs (**Sup. Fig. 1e**).

DrmA has multiple interactions with the backbone of the DNA, mostly electrostatic contacts between positively charged side chains and DNA phosphate groups (**Fig. 2b-c**). Such non-specific electrostatic DrmAB:DNA contacts provide the molecular basis for the previously observed broad anti-phage targeting by DISARM ⁷. Since conventional RM systems require recognition of specific DNA sequence motifs for nuclease activity, they can be easily evaded through phage evolving escape mutations ^{11,23}. However, by lacking sequence preferences for

DNA binding, DISARM can provide effective anti-phage defense in the absence of a particular restriction site.

To address the functional relevance of these interactions, we mutated various DNA-interacting residues within DrmA (K803, R1294, R659, R810) to alanine. All four mutants showed reduced anti-phage protection *in vivo* (**Fig. 2d**). Together, these data underscore the functional role of DrmA as the DNA-targeting arm of the complex and provide the mechanism of targeting a wide range of phage by DISARM. Overall, the non-specific interactions between DrmA and DNA provide a structural basis for broad phage targeting by DISARM.

4.2.4. ATPase activity is critical for DISARM function

We observed strong density corresponding to ADP within our DNA-bound (**Fig. 2e**) and DNA-free maps, both of which were at the interface between the RecA1 and RecA2 domains, the canonical ATP-binding site of SF2 helicase domains. While the RecA domains of DrmA suggest that it is a helicase, the impact of its helicase activity on anti-phage targeting is poorly understood. We mutated four ADP-interacting residues to test their impact on DISARM activity. All four mutants showed reduced DISARM activity against a broad range of phage *in vivo*. This confirmed that in addition to DNA binding, ATP hydrolysis by DrmA is essential for DISARM activity. This is similar to type I RM systems, which encode a DEAD-box helicase domain protein that drives DNA translocation upon recognition of an unmethylated restriction site^{10,24}. Akin to DrmAB, ATP-dependent DNA translocation is essential for type I RM system activity²⁵. However, it is unlikely that these two antiphage systems share a common DNA degradation mechanism, since type I RM systems translocate and cleave dsDNA, while DISARM appears to bind exclusively ssDNA (**Fig. 3**).

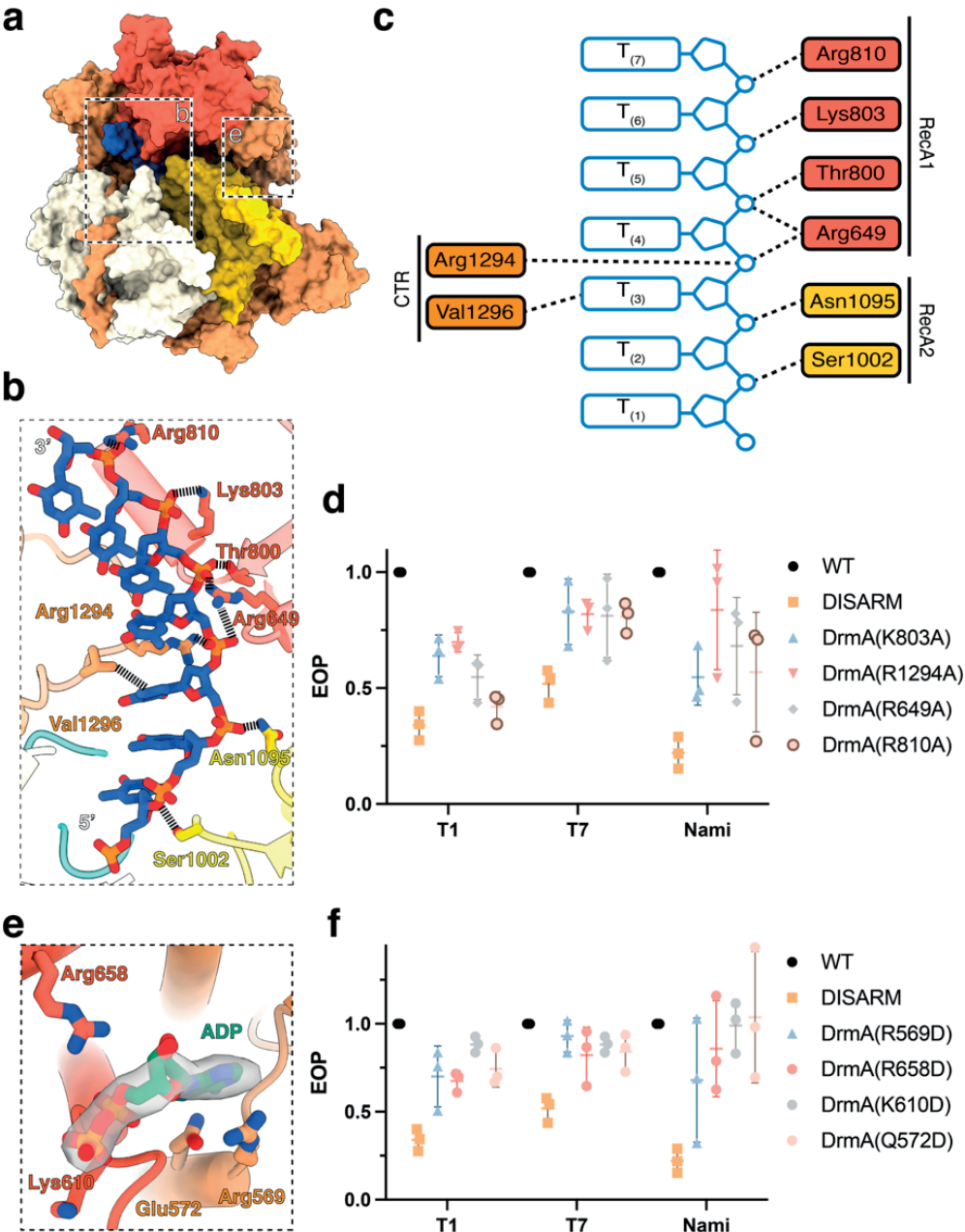


Figure 2. Both ssDNA and nucleotide cofactor binding are essential for DISARM function.
a) Overall structure of DrmAB:DNA:ADP. Boxes indicate regions featured in panels (**b**) and (**e**).
b) Close-up view of DrmA:DNA interactions. No interactions DrmB-DNA are observed.
c) Schematic of DNA-protein contacts colored by protein domain. **d)** Effect of mutagenesis of DrmA

4.2.5. DrmA contains an unstructured trigger loop that partially occludes the DNA-binding site

Two dominant, distinct populations (~120,000 particles each) emerged during focused 3D classification; one of which had strong DNA density (DrmAB:ADP:DNA) and one lacking observable DNA density (DrmAB:ADP). Comparison of our two maps revealed the presence of a 55-residue loop in DrmA that is otherwise absent in the DNA-bound structure (**Fig. 3a-b**). Due to the flexibility of this loop, we were unable to confidently assign sequence position. We modeled this region as poly-alanine. To further validate this model, we determined a structure of DrmAB in the absence of DNA and ADP at a resolution of 3.8 Å (**Sup. Fig. 5**). Rigid-body docking of our higher-resolution DrmAB:ADP model revealed density consistent with the presence of this loop (**Sup. Fig. 5**). Thus, the presence of this loop is mutually exclusive with DNA binding.

After superposition of our two models, we were surprised to observe a severe steric clash between DNA and this unstructured loop (**Fig. 3c**) that occurs at the interface between the 5' DNA end and RecA2. We hypothesized that this region of DrmA acts as a trigger loop (TL) (residues 176-232) that activates DrmAB after being dislodged by target DNA. Given that the presence of TL and DNA are mutually exclusive, we hypothesized that in the absence of single-stranded DNA, this loop may partially occlude the DNA binding site, preventing DrmAB from loading onto non-phage (i.e. host) DNA. To study its relevance in the DISARM mechanism, we created a mutant lacking this loop (DrmA(Δ loop)), removing residues 181-233. Since DrmA(Δ loop) maintained its ability to co-purify with DrmB, and the DrmA(Δ loop)B complex eluted at the expected volume according to size-exclusion chromatography (**Sup. Fig. 1a**), heterodimer assembly does not depend on this region of DrmA. We performed native electrophoretic mobility shift assays to compare binding of DrmAB and DrmA(Δ loop)B to fluores-

[Fig. 2]... DNA-, interacting residues on DISARM anti-phage activity. **e**) Close-up view of DrmA:ADP interactions **f**) Effect of mutagenesis of DrmA ADP-interacting residues on DISARM anti-phage activity. Points correspond to three biological replicates, with mean and standard deviation shown.

cently-labelled 75-nt single-stranded DNA (**Fig. 3d-e**). While DrmAB bound DNA at concentrations of 1 μ M or higher, the presence of smeared bands likely indicated dissociation of DNA from the complex. In contrast, DNA dissociation from DrmA(Δ loop)B bound DNA was not observed. Furthermore, a second super-shift occurred at protein concentrations of 1 μ M or higher, likely corresponding to multiple copies of the complex binding to the DNA concurrently, as observed for other nucleoprotein complexes ²⁶.

Since TL limits DNA binding, we performed additional binding assays to determine how this loop affects DNA substrate preferences of DrmAB (**Fig. 3f-g**). DrmAB bound to ssDNA and DNA with a 5'-overhang (ovh) but did not bind to dsDNA or DNA with a 3'ovh. DrmA(Δ loop)B displayed no DNA substrate preference, binding to ssDNA, dsDNA, and DNA with both a 5'ovh and 3'ovh. Thus, the TL limits structure-specific DNA binding by partially occluding the DNA-binding surface of DrmAB.

While the methylation modules of the DISARM operon recognize specific sequences, DISARM targets a diverse range of phage DNA sequences ⁷. We thus propose that TL functions as a specificity filter, preventing interactions with dsDNA and DNA containing 3'ovh. Since DNA would occupy a larger surface on DrmAB than TL, DNA with a 5'ovh would be an ideal substrate for efficiently competing with TL for loading onto the complex. This may function as a mechanism for conferring substrate specificity to DISARM and prevent interactions with bacterial chromosome and other self-DNA, which would likely result in deleterious autoimmune effects. By relying on DNA structural context rather than simply DNA sequence, DISARM appears to utilize an alternative mechanism for avoiding self-targeting to many other nucleic acid-based antiphage systems ^{5,24}.

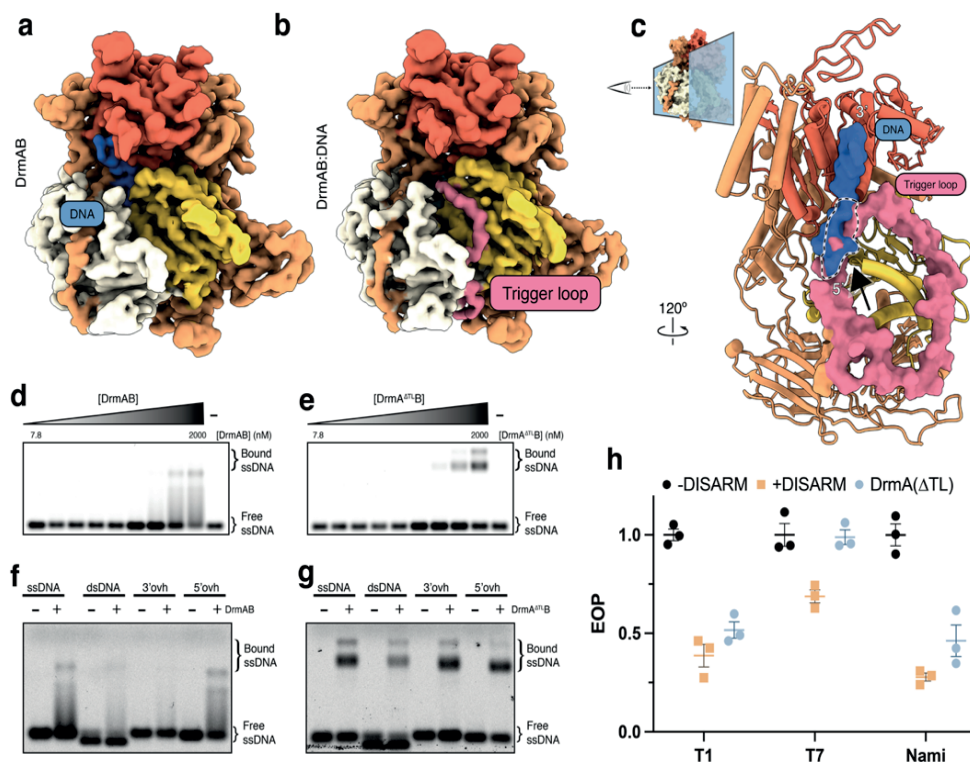


Figure 3. DrmA trigger loop (TL) partially occludes DNA-binding site. **a) & b)** Cryo-EM reconstruction of apo DrmAB (**a**) and DNA-bound DrmAB (**b**). Density corresponding to DrmA trigger loop is shown in pink. DrmAB are colored by structural domains as shown in **Fig. 1A**. **c**, Overlay of DrmAB-bound DNA (blue) and TL (pink) showing the steric clash (dashed black & white outline). TL partially occludes DNA-binding site. DrmB and parts of DrmA have been omitted for clarity. Graphic at the top right shows how the view in **c** is related to the structures in panels (**a** & **b**). **d) & e)** EMSA analysis of DrmAB and DrmA(Δ loop)B binding DNA. The smeared bands that occur at high concentrations of DrmAB correspond to bound DNA dissociating from the complex. This did not occur for DrmA(Δ loop)B, and an additional super-shift was present, corresponding to multiple copies of DrmA(Δ loop)B binding to the same 75-nt DNA concurrently. **f) & g)** DNA substrate preferences of DrmAB and DrmA(Δ loop)B. The 75-nt Cy5-labelled DNA (ssDNA) was annealed to complementary oligos corresponding to the full sequence (dsDNA), or 20 bases at the 5' or 3' ends (3'ovh and 5'ovh, respectively). DrmAB showed a preference for ssDNA and 5'ovh DNA, whereas DrmA(Δ loop)B bound to all DNA substrates tested. **h)** Effect of truncation of TL on DISARM anti-phage activity. Points correspond to three biological replicates, with mean and standard deviation shown.

4.2.6. Methylation sensing and DNA-mediated DrmAB activation

We sought to visualize additional conformational changes within DrmAB associated with DNA loading and complex activation. Within the core DNA-binding site of DrmAB, continuous cryo-EM density is observed between the DNA nucleobase at position 3 (T₍₃₎, **Fig. 2c**) and the side-chain of DrmA(V1296), likely corresponding to a van der Waals interaction (**Fig. 4a**). Since this residue is highly conserved (found in >98% of top 250 DrmA homologs) and is the sole interaction between DrmAB and a DNA base, we hypothesized that DrmA(V1296) may act as a sensor for DNA methylation status. *Serratia* DISARM operon contains both a 5-cytosine DNA methyltransferase and an N6-adenine DNA methyltransferase gene which target ACAC(mA)G and YMT(mC)GAKR motifs, respectively (**Fig. 1a**), which may constitutively methylate the host genome to allow distinguishing self (i.e. methylated) from non-self (i.e. unmethylated) phage DNA, akin to RM systems ^{7,11}.

Further analysis of our structure revealed that DrmA(V1296) is in close contact with the faces of the DNA nucleobases, tightly wedged between DNA positions 3 and 4. Modeling this position as a bulky tryptophan (Y) side-chain showed severe steric clashes, and this mutant rendered DISARM unable to protect against phage in our *in vivo* assay (**Fig. 4b**). We then tested the *in vivo* antiphage activity of DISARM with a glycine at this position (DrmA(V1296G)), which would no longer contact DNA nucleobases. This also resulted in severely reduced DISARM activity, supporting the notion that V1296 plays a critical role in DNA recognition (**Fig. 4b**). We propose that since this interaction is critical to DISARM function, differences in DNA methylation status may perturb this interaction, providing a structural mechanism for enabling DISARM to differentiate between self- and non-self DNA and preventing autoimmune targeting.

To further test this hypothesis, we performed ATPase assay of DrmAB in the presence and absence of methylated and unmethylated DNAs. In the absence of DNA, DrmAB did not exhibit ATPase activity (**Fig. 4b**). This is expected for a SF2 helicase, where DNA translocation is typically coupled to ATP hydrolysis ²⁷.

Robust ATPase activity is strongly stimulated by unmethylated DNA, but the ATPase rate is reduced in the presence of a methylated DNA substrate (containing three separate 5-methyl-cytosine bases, i.e. the modification provided by DrmMII⁷). This supports the model that this multi-protein complex detects phage DNA as it is injected into bacteria or is being replicated. Additionally, recent studies demonstrate that DISARM can protect against plasmid conjugation, and that DISARM targeting is enhanced by an order of magnitude in the absence of cognate methylation target sequences²⁸. This data indicates that DISARM is repressed by methylation rather than being activated by the presence of a specific signal sequence.

We then compared our DNA-bound and DNA-free DrmAB complexes by superposing our models and generating motion vectors (**Fig. 1a**). Upon loading of unmethylated DNA, we observed minor (mostly <5Å) conformational changes of the RecA1 and RecA2 domains (**Fig. 4a**). Surprisingly, we observed large conformational rearrangements in the N-terminal half (NTH) of DrmB, despite a lack of any contacts with DNA. DrmB NTH shifts ~10 Å towards the RecA2 domain of DrmA, tightly clamping the complex around DNA (**Sup. Fig. 5**).

This long-range allosteric communication appears to be mediated by the C-terminal region (CTR) of DrmA, which tightly wraps around the N-terminus of DrmB acting as a pivot arm (PA). While DrmA(V1296) undergoes a minor shift upon recognition of unmethylated DNA, this is propagated into a much larger conformation change in the end of the PA and DrmB NTH. Thus, conformational change upon loading unmethylated DNA activates DrmAB. Since the DUF1998 domain of DrmB acts as a hinge for the conformational changes induced by the PA, we propose that DUF1998 should be renamed as a Helicase Allosteric Relay (HAR) domain.

DrmA PA is tightly interwoven with DrmB NTH, forming a myriad of contacts. These include a network of electrostatic interactions, hydrophobic contacts and π - π stacking (**Fig. 4d**). We also observed that a five-residue segment of DrmA PA contributes to a 5-stranded β -sheet in *trans* with DrmB, forming an oligonucleotide/oligosaccharide-binding (OB) fold (**Fig. 4e**). The β -strand from

DrmA is sandwiched between strands from DrmB, further locking the PA of DrmA in place. These intimate contacts allow DrmA to allosterically communicate with DrmB, setting in motion major conformational changes upon loading of unmethylated DNA.

To test the importance of this conformational change for DISARM activity, we truncated DrmA so that it lacked the PA and tested the ability of DISARM to prevent phage replication *in vivo*. We found that DrmA-ΔPA had severely reduced DISARM activity against a broad range of phage, supporting the notion that the CTR-driven conformational changes of DrmAB result in DISARM activation (**Fig. 4f**). Based on our structural and functional data, we propose that upon displacement of the TL by ssDNA, DrmA(V1296) senses the DNA methylation status, allosterically activating DrmAB and triggering the DISARM antiphage response.

In summary, through coupling ATPase activity to DNA nucleobase methylation status, DrmAB may be able to achieve specific targeting of unmodified non-self (i.e. phage) DNA without relying on a given DNA sequence.

[Fig. 4]...Right: Motion vectors showing conformational changes of DrmAB upon DNA binding. DrmB N-terminal half (DrmB NTH, cyan) does not make direct contacts with DNA but undergoes a large conformational change upon DNA binding, mediated by the PA of DrmA (magenta). Distances of conformational changes are shown parenthetically. Motion vectors for shifts smaller than 3Å are omitted. **b**) EOP assay to determine the effects of V1296G or V1296W mutations to DISARM activity. Points correspond to three biological replicates, with mean and standard deviation shown. **c**) ATPase assay measuring the rate of ATP hydrolysis by DrmAB, as monitored by increase in absorbance at 360 nm. In the absence of DNA, no ATPase activity is monitored. In the presence of DNA stem loop lacking modification (blue), ATPase activity is observed. This is severely reduced for DNA containing three separate 5-methyl cytosine bases (orange). DNA sequences are listed in Sup. Table 2. Experiments were performed in triplicate, and the mean and standard deviation of the mean are depicted. **d**) Close-up views of DrmA(CTR):DrmB interactions. **e**) DrmA C-terminus forms a β-sheet in trans with DrmB. **f**) Effect of truncation of DrmA PA on DISARM anti-phage activity. Points correspond to three biological replicates, with mean and standard deviation shown.

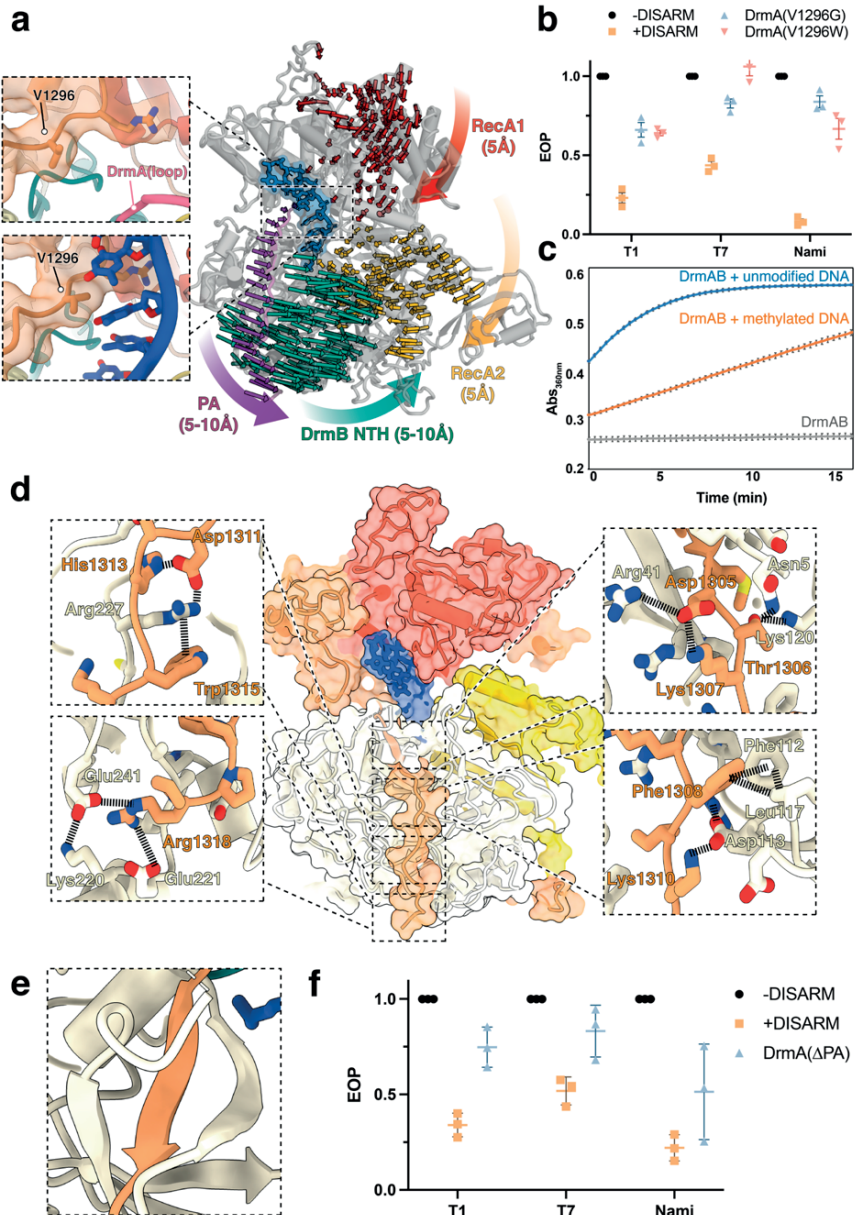


Figure 4. Binding of unmethylated ssDNA activates DrmAB through large conformational rearrangements. **a)** DrmA C-terminal pivot arm (PA) interface with substrate DNA. Left: close-up views of DrmAB map and model lacking DNA (top) and DrmAB:DNA map and model (bottom). In the DNA-bound map, a direct interaction between a DNA base and Val1296 is present (black arrow). In the DrmAB map lacking DNA, the DrmA autoinhibition loop is present (pink arrow).

4.3. Discussion

We propose a model whereby DrmA and DrmB are expressed by the host cell in an autoinhibited form in the absence of phage infection. Through the constitutive expression of the core DrmAB complex, a DISARM response can be rapidly activated upon recognition of phage stimulus without necessitating transcription of either component, in accordance with previous observations⁷. In the absence of DNA, the DNA-binding channel of DrmAB is occluded by an unstructured ~50 residue loop within TL. While truncation of this loop does not affect heterodimer formation, we observed that DrmA(Δ loop)B no longer exhibited discrimination for 5'ovh-containing DNA substrates and could bind both 3'ovh and dsDNA. We conclude that this loop may act as a selectivity filter or TL, ensuring the loading of specific substrates onto DrmAB. Since TL only partially occludes the DNA binding site, loading of 5'ovh DNA can effectively compete against and evict TL, which is likely intrinsically disordered. A similar mechanism of autoinhibition was recently reported for human Separase²⁹, and examples of substrate-alleviated autoinhibition are numerous amongst other SF2 helicase complexes^{30–33}. Interestingly, a recent study of the human DEAD-box helicase DHX37 revealed a highly similar autoinhibition mechanism, whereby an unstructured protein loop occupied the substrate-binding channel³⁴. It may emerge that similar mechanisms of autoinhibition are widespread amongst a diverse group of SF2 helicase proteins from all domains of life.

Based on our data, we propose that the activating substrate for DISARM is unmodified 5'ovh DNA. Unlike many RM systems, DISARM does appear to not rely on specific sequences, enabling broad anti-phage targeting that cannot be circumvented escape mutants. This is further supported by *in vivo* functional data that show that DISARM activity is enhanced by a lack of cognate methylation target sequences²⁸. Future studies are required to investigate the interplay between DrmAB and other less-conserved DISARM components (e.g. DrmD and DrmE) within DISARM, and how these subunits affect the specificity of DISARM activation.

The preference for 5'ovh ssDNA is particularly important since many dsDNA phage inject their genome in a linearized form with two sticky end 5'ovh, which subsequently become ligated during DNA circularization ^{35,36}, thus providing a stimulus for DISARM activation. This is supported by the previous observation that DISARM has no effect on phage adsorption but blocks phage DNA circularization ⁷. If a dsDNA phage does not have such 5'ovh, then it would still likely be targeted by DISARM since it would still replicate via rolling circle DNA replication, where the leading strand is replicated in a 5' to 3' direction and the lagging strand is synthesized as Okazaki fragments ³⁷. This model explains the previous observation that DISARM is activated by the onset of phage replication, likely resulting probably in the degradation of viral DNA ⁷. This data also explains how DrmAB target specificity is conferred by DNA structure rather than a specific sequence, providing an elegant mechanism for balancing broad DNA targeting with minimal autoimmune consequences. DISARM therefore represents a novel paradigm in bacterial antiviral defense mechanisms that target nucleic acids. Rather than achieving specific and efficient degradation of phage DNA through recognition of specific sequences, DISARM recognizes the structural context of DNA. The core DISARM complex DrmAB is able to confer specificity of DNA loading by targeting substrates with a 5'ovh, thereby avoiding degradation of the host chromosomal DNA. This specificity is further conveyed by a sensitivity to DNA methylation status, whereby methylated DNA limits the ATPase activity of DrmAB which is essential for antiphage defense. This amounts to a two-pronged mechanism of selectivity that does not rely on a given DNA sequence, and thereby circumvents evasion via mutational escape.

While this model is consistent with our data, the class 2 DISARM system has been demonstrated to confer protection against modified phage ⁷. While DISARM preferentially targets unmodified DNA, methylated DNA can still support ATP hydrolysis and thus defense activation. It may be the case that DISARM activation is significantly reduced by phage DNA methylation, but the abundance of phage 5'ovh during successive cycles of replication may provide

sufficient stimulus to activate DISARM. Additionally, other DISARM subunits may confer additional mechanisms to detect invading phage.

We hypothesize that once activated, DrmAB binds to, translocates, and unwinds phage DNA, blocking replication and transcription and disrupting the phage infection cycle. Additionally, DrmAB may recruit a nuclease to trigger DNA degradation. Since the PLD-containing candidate nuclease DrmC is dispensable for DISARM activity and is often missing from annotated DISARM operons, it has remained a mystery how this is achieved ⁷. However, since PLD-containing nucleases are highly abundant in bacterial genomes, we propose that DrmAB may recruit an alternative PLD-nuclease protein. While previous *in vivo* assays were performed in *Bacillus subtilis* (which encodes a 42.2% identity DrmC-like homolog), our *in vivo* assays were performed in *Escherichia coli* (which encodes a 93.4% identity DrmC-like homolog). The ubiquity and versatility of such DrmC-like non-specific PLD nucleases ^{38–41} may make DrmC redundant, allowing DISARM to ‘mix-and-match’ components from other innate immune systems encoded by the host bacteria in order to achieve rapid and broad anti-phage protection. While CRISPR-Cas systems rely on RecBCD and other host DNA double-strand break repair complexes for adaptation (i.e. acquiring spacers) ^{42–44}, it is unlikely that DISARM utilizes this mechanism, since it does not require spacer acquisition of phage target sequences.

Phage therapy represents a promising avenue to treat bacterial infection in an age where antibiotic resistance is widespread ⁴⁵. Since DISARM is an incredibly widespread mechanism of antiviral defense in bacteria, once we understand the fundamental mechanisms underlying DISARM activation it may be possible to develop small molecule treatments to inhibit DISARM, which could be delivered in conjunction with phage therapy to increase the effectiveness of such antibacterial treatment.

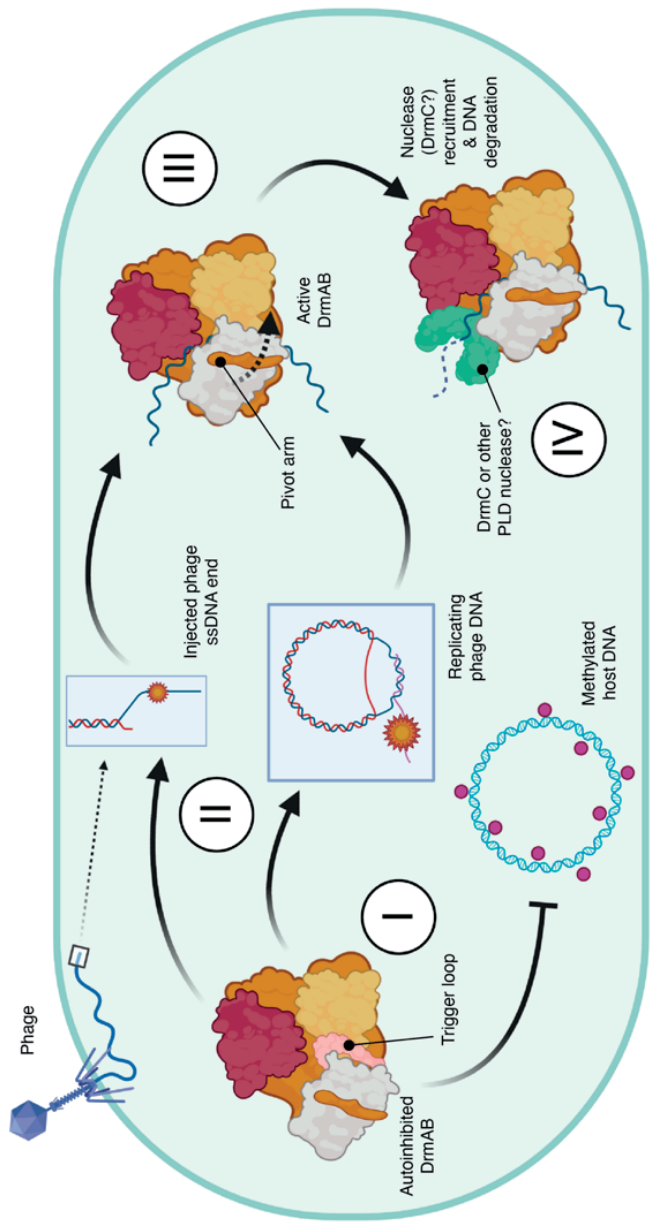


Figure 5. Model for anti-phage surveillance by DISARM. I) In the absence of phage infection, DrmAB complex is autoinhibited via TL. This allows the constitutive expression of the complex to enable rapid immune activation upon stimulus. II) DISARM is recruited to single-stranded 5'-OH DNA, which typically occurs during initial phage DNA injection prior to genome circularization, or during rolling circle DNA replication. III) Loading of ssDNA into DrmAB dislodges TL, resulting in a conformational change and DrmAB activation. IV) Once DrmAB is active, DrmC or other nucleases may be recruited to degrade foreign DNA. DISARM may act to defend against phage at step III by loading onto phage DNA ends and physically blocking replication, and at step IV through recruiting a nuclease and degrading phage DNA.

4.4. Material and methods

4.4.1. Cloning and protein expression

Genes *drmA* and *drmB* were amplified from genomic DNA of *Serratia* sp. SCBI using primers indicated in **Sup. Table 3**, with Q5 DNA Polymerase (NEB: M0491S) as indicated by the manufacturer. Amplified gene *drmB* was cloned in plasmid 13S-S (Addgene: 48329), resulting in a N-terminal 6X His tagged to the protein product. His-tag was separated from the gene with a tobacco etch virus (TEV) protease recognition sequence. Amplified gene *drmA* was cloned in plasmid pACYCDuet-1 (Sigma-Aldrich: 71147), resulting in a non-tagged protein product. Cloning products were transformed in Dh5 α using the heat-shock protocol ⁴⁶ and confirmed by sanger sequencing (Macrogen).

Both plasmids were transformed into *E. coli* BL21-AI™ (Thermo Fisher) to express the DrmAB complex. Briefly, cultures were grown in LB medium at 37°C with shaking until exponential growth phase (OD_{600 nm} of 0.5). After cooling down in ice for 20 min, they were induced by addition of 0.2 % (w/v) L-arabinose and 1 mM IPTG. After overnight incubation at 20°C with shaking, cells were pelleted by centrifugation, resuspended in Buffer A (25 mM HEPES, pH 8, 600 mM NaCl, 5% glycerol, 25 mM imidazole, cOmplete™ EDTA-free Protease Inhibitor), and sonicated (10 cycles of 30 sec separated by 30 sec breaks, 40% amplitude). Cell lysate was pelleted down (20,000 xg, 60 min, 4°C) to remove cell debris, and supernatant was filtered with a 0.45 μ m PES filter and placed on ice until proceeding with the purification.

All the mutants used here were generated by round-the-horn side-directed mutagenesis and purified as the wild-type (WT) proteins.

4.4.2. Protein purification

Purification was performed at 4°C using an ÄKTA™ pure (GE Healthcare) to control flow and column pressure. Lysates were loaded onto a 5 mL Ni-NTA Superflow Cartridges (Qiagen) equilibrated with Buffer B (25 mM HEPES, pH 8, 600 mM NaCl, 5% glycerol, 25 mM imidazole, 1 mM DTT).

Unbound lysate components were washed with 10 column volumes (CV) of Buffer B. Bound proteins were eluted with 5 CV of Buffer C (25 mM HEPES, pH 8, 600 mM NaCl, 5% glycerol, 150 mM imidazole, 1 mM DTT) and collected in fractions. After checking by SDS-PAGE electrophoresis using a 4–20% Mini-PROTEAN® TGX precast gel (Bio-Rad), fractions containing DrmAB were pooled together and buffer exchanged to Buffer D (25 mM HEPES, pH 8, 150 mM NaCl, 5% glycerol, 1 mM DTT). To separate the DrmAB complex from proteins that also bound the Ni-NTA column, sample was loaded onto a HiLoad 16/600 Superdex 200 (GE Healthcare). Column was washed with 2 CV of Buffer D and fractions collected. After SDS-PAGE electrophoresis check, fractions containing the complex DrmAB were pooled together and concentrated using 30 KDa NMWL Ultra-15 Amicon® (Merck). Protein concentration was estimated by the Bradford Assay (Thermo Scientific: 23246) as indicated by the manufacturer.

4.4.3. Bacterial strains used in phage assays

DISARM system genes from *Serratia* sp. SCBI were cloned in plasmids (Sup. Table 4), for both WT and mutant. Plasmids were transformed in *E. coli* BL21-AI. For assays, strains were cultured in LB media at 37°C, and induced with 0.1 % (w/v) L-arabinose and 0.5 mM IPTG at early exponential growth phase (OD_{600 nm} of 0.2-0.3). Cultures were incubated at 37°C for 90 min and then used for phage assays. In parallel, a control strain (BL21-AI WT) was grown for all assays. When required, LB media was supplemented with antibiotics at the following final concentrations: 100 µg/mL ampicillin, 50 µg/mL kanamycin, 50 µg/mL streptomycin, 50 µg/mL spectinomycin, 10 µg/mL chloramphenicol.

4.4.4. Phage strains

E. coli phage T1, T7, and Nami were used in this study. For their production, phage T1 and T7 were propagated in *E. coli* BL21-AI, as described previously⁴⁷, and bacteriophage Nami in its host *E. coli* isolate R10256 following same procedure. Briefly, bacterial cultures at exponential growth phase (aprox. 0.4 OD_{600 nm} in a 10 mm cuvette readers) were infected with a phage lysate and

incubated overnight. Then, cultures were spun down and the supernatant filtered through 0.2 μ m PES membranes. When required, phage were concentrated by addition of PEG-8000 and NaCl at final concentrations of 100 mg/mL and 1 M, respectively. Following, it was incubated overnight at 4°C, centrifuged at 11,000 xg at 4°C for 60 min, and the phage-containing pellet resuspended in the desired final volume of Saline Magnesium (SM) buffer. Phage stocks were stored at 4°C before their use and titer determined as indicated below.

4.4.5. Phage titering

Bacteria cultures in exponential growth phase were used to titer phage stocks. For this, 100 μ L of culture were mixed with 5 mL of 0.6 % LBA at 45 °C and poured to a LBA plate to form a bacterial layer. Ten-fold dilution of phage in SM buffer or LB media were plated on the top of the bacteria in 10 μ L drops and let dry for 20 min. Plates were incubated overnight up-side down at 37 °C. To determine the phage titer, the number of center of infections (plaques) were counted. *E. coli* BL21-AI was used for titering phage T1, T7, and Nami.

4.4.6. Efficiency of plating determination

To determine the efficiency of plating (EOP), phage were plated on the induced containing the DISARM system and compared to the plating on control strain BL21AI. For this, 200 μ L of the bacterial cultures were mixed a specific, countable number of infectious particles (50-150 PFU/plate) and 4.5 mL of 6 % LBA at 45 °C. The mixture was then poured on top of a LBA plate to form a bacterial layer containing the infectious particles. After overnight incubation at 37°C, the EOP was calculated by dividing the number of plaques counted on each plate by the number of plaques formed in the control strain.

4.4.7. CryoEM sample preparation, data collection and processing

To capture DrmAB in the act of unwinding a DNA substrate, a forked DNA construct consisting of a 19-bp stem-loop structure with a 6 base 5' overhang and a 20 base 3' overhang was designed. This substrate was chosen because

similar substrates have been used to capture unwinding intermediates of bacterial SF2 helicase complexes^{48,49}. DrmAB was buffer exchanged into cryoEM sample buffer (150 mM NaCl, 25 mM HEPES pH 7.4) through repeated application of buffer to sample within a 0.5 ml spin concentrator, with a 30 kDa size cutoff. 10 μ M DrmAB was incubated with 10-fold excess DNA stem loop (100 μ M) and 100-fold excess ADP (1 mM). Complex formation was monitored using electrophoretic mobility shift assay (EMSA), confirming DNA binding by DrmAB (**Sup. Fig. 1**). After incubation at room temperature ($\sim 25^{\circ}\text{C}$) for 30 minutes, DrmAB:ADP:DNA was applied to glow discharged holey carbon grids (C-flat 2/2, Protochips Inc.), blotted for 2 s with a blot force of 4 and rapidly plunged into liquid ethane using an FEI Vitrobot MarkIV.

Data was collected on an FEI Titan Krios cryo-electron microscope equipped with a K3 Summit direct electron detector (Gatan, Pleasanton, CA) operating in super-resolution counting mode. Images were recorded with SerialEM⁵⁰ with a pixel size of 1.1 Å over a defocus range of -1.5 to -2.5 μ m. During early stages of data collection, a preferred orientation was observed. To ameliorate this, a dataset of 6828 micrographs was collected at 30° tilt, in addition to the original 2548 micrographs collected without tilt. Movies were recorded at 13.3 electrons/pixel/second for 6 seconds (80 frames) to give a total dose of 80 electrons/pixel. CTF correction, motion correction and particle picking were performed in real-time using WARP⁵¹, resulting in 4,669,932 particles, which were uploaded to cryoSPARC v2⁵² (**Sup. Fig. 2**).

Multiple rounds of 3D classification within cryoSPARC yielded a final set of 144,989 particles that gave a 3D reconstruction at a global resolution of 2.84 Å using non-uniform refinement⁵³. However, since bound DNA had weak density, an alternative data processing strategy was implemented. Two rounds of ab initio reconstruction followed by heterogeneous refinement were performed on the 0° and 30° tilt datasets separately within cryoSPARC. The resulting subset of undamaged particles was combined and yielded a 3.1 Å reconstruction. These particles were then imported into Relion 3.1⁵⁴, where masks covering the core of the complex (to improve DNA density) and the bottom of the complex (to improve

the quality of DrmB density) were generated within ChimeraX⁵⁵ and Relion. These masks were then used for focused 3D classification within Relion (N = 6, T=25). Particles within classes corresponding to DrmAB:ADP:DNA and DrmAB:ADP were then selected for Relion 3D auto-refinement and post-processing, resulting in structures with global resolutions of 3.4 Å and 3.3 Å, respectively (**Sup. Fig. 2**).

To assist model building, the structures of multiple overlapping regions of DrmA and DrmB were predicted using trRosetta⁵⁶. These models were initially fitted into the map as rigid bodies using ChimeraX. Once suitable fits were found, bulk flexible fitting was performed using Isolde and Namdinator⁵⁷. Regions of the model that trRosetta failed to predict were either built de novo (in well-resolved regions of the map with local resolutions of up to 3.1Å), or omitted (in flexible, poorly-resolved regions). Between the two structures, 98% of the total DrmAB sequence was modelled.

Real-space model improvement was performed using Isolde⁵⁸, and final models were subjected to real-space refinement in Phenix⁵⁹. Structural figures were prepared using ChimeraX. Structural analysis of sequence conservation was performed using ConSurf⁶⁰, and visualized in ChimeraX. Motion vectors to visualize the conformational change from DrmAB:ADP to DrmAB:ADP:DNA were generated and visualized using PyMol.

4.4.8. Native electrophoretic mobility shift assays (EMSAs)

To determine the DNA substrate preferences of DrmAB, 5'Cy5-labelled 75-nt DNA was incubated with 1.5-fold excess complement, 5'block or 3'block oligos (**Sup. Table 2**) and heat annealed at 85°C for 10 minutes prior to cooling to 4°C. 5 nM DNA was incubated with 1 µM DrmAB or DrmA(Δ loop)B in cryoEM sample buffer for 30 minutes at room temperature in a total volume of 20 µl. A total of 10 µl of each sample was mixed with 2 µL 6× loading buffer (30% v/v glycerol, 5% Ficoll 400, 50 mM NaCl, 10 mM HEPES pH 8, 5 mM EDTA, 0.002% w/v bromophenol blue). Electrophoresis was carried out on a non-denaturing

1.5% agarose gel for 90 min at 100 V in $1 \times$ Tris-borate-EDTA buffer. Gels were imaged using a fluorescence scanner (Fujifilm FLA-5100) with 532 nm excitation.

For comparison of DNA binding between DrmAB and DrmA(Δ loop)B, serial 2-fold dilutions from either protein complex were incubated with 5 nM Cy5-labelled DNA. Samples were incubated and EMSAs were performed as described above.

For the overhang binding assay, a 3'-Cy5-labelled DNA substrate was hybridized to unlabeled DNAs of various lengths and binding was tested as described above.

4.5. References

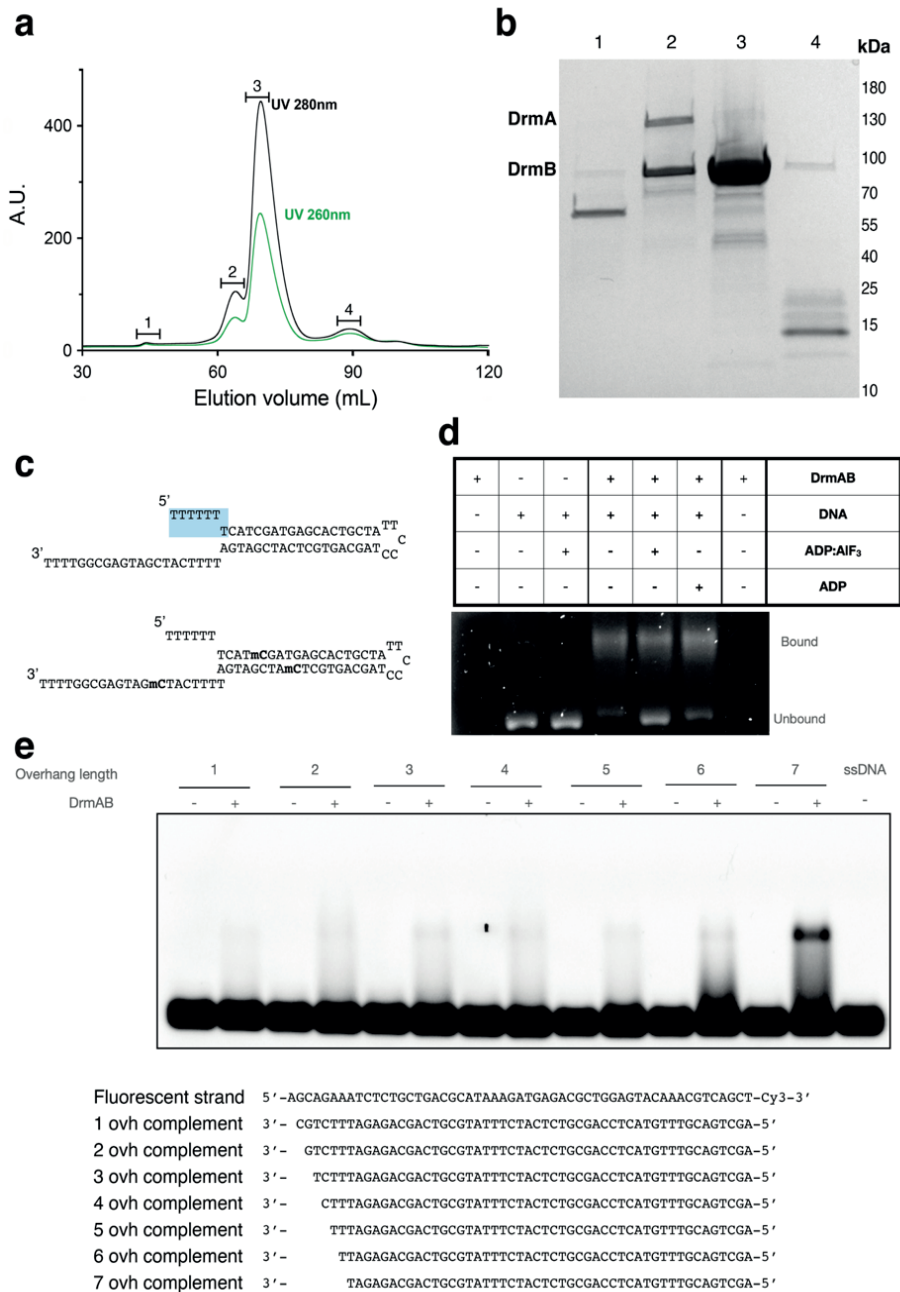
1. Hampton,H.G., Watson,B.N.J. and Fineran,P.C. (2020) The arms race between bacteria and their phage foes. *Nature*, 577.
2. Egido,J.E., Costa,A.R., Aparicio-Maldonado,C., Haas,P.-J. and Brouns,S.J.J. (2021) Mechanisms and clinical importance of bacteriophage resistance. *FEMS Microbiol. Rev.*
3. Labrie,S.J., Samson,J.E. and Moineau,S. (2010) Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.*, 8.
4. Samson,J.E., Magadán,A.H., Sabri,M. and Moineau,S. (2013) Revenge of the phages: Defeating bacterial defences. *Nat. Rev. Microbiol.*, 11.
5. Hille,F., Richter,H., Wong,S.P., Bratovič,M., Ressel,S. and Charpentier,E. (2018) The Biology of CRISPR-Cas: Backward and Forward. *Cell*, 172.
6. Bravo,J.P.K., Liu,M., McCool,R.S., Jung,K., Johnson,K.A. and Taylor,D.W. (2021) Structural basis for mismatch surveillance by CRISPR / Cas9. *bioRxiv*.
7. Ofir,G., Melamed,S., Sberro,H., Mukamel,Z., Silverman,S., Yaakov,G., Doron,S. and Sorek,R. (2018) DISARM is a widespread bacterial defence system with broad anti-phage activities. *Nat. Microbiol.*, 3.
8. Doron,S., Melamed,S., Ofir,G., Leavitt,A., Lopatina,A., Keren,M., Amitai,G. and Sorek,R. (2018) Systematic discovery of antiphage defense systems in the microbial pangenome. *Science*, 4120.
9. Xiong,X., Wu,G., Wei,Y., Liu,L., Zhang,Y.Y., Su,R., Jiang,X., Li,M., Gao,H., Tian,X., et al. (2020) SspABCD–SspE is a phosphorothioation-sensing bacterial defence system with broad anti-phage activities. *Nat. Microbiol.*, 6.
10. Gao,Y., Cao,D., Zhu,J., Feng,H., Luo,X., Liu,S., Yan,X.X., Zhang,X. and Gao,P. (2020) Structural insights into assembly, operation and inhibition of a type I restriction–modification system. *Nat. Microbiol.*, 5.
11. Rusinov,I.S., Ershova,A.S., Karyagina,A.S., Spirin,S.A. and Alexeevski,A. V. (2018) Avoidance of recognition sites of restriction-modification systems is a widespread but not universal anti-restriction strategy of prokaryotic viruses. *BMC Genomics*, 19.
12. Thiaville,J.J., Kellner,S.M., Yuan,Y., Hutinet,G., Thiaville,P.C., Jumpathong,W., Mohapatra,S., Brochier-Armanet,C., Letarov,A. V., Hillebrand,R., et al. (2016) Novel genomic island modifies DNA with 7-deazaguanine derivatives. *Proc. Natl. Acad. Sci. U. S. A.*, 113.
13. Ofir,G. and Sorek,R. (2018) Contemporary Phage Biology: From Classic Models to New Insights. *Cell*, 172.
14. Holm,L. (2020) Using Dali for Protein Structure Comparison. *Methods Mol. Biol.*, 2112.
15. Makarova,K.S., Wolf,Y.I., Snir,S. and Koonin,E. V. (2011) Defense Islands in Bacterial and Archaeal Genomes and Prediction of Novel Defense Systems. *J. Bacteriol.*, 193.
16. Roske,J.J., Liu,S., Loll,B., Neu,U. and Wahl,M.C. (2020) A skipping rope translocation mechanism in a widespread family of DNA repair helicases. *Nucleic Acids Res.*, 1.

17. Yakovleva,L. and Shuman,S. (2012) Mycobacterium smegmatis SftH exemplifies a distinctive clade of superfamily II DNA-dependent ATPases with 3' to 5' translocase and helicase activities. *Nucleic Acids Res.*, 40.
18. Pace,N.J. and Weerapana,E. (2014) Zinc-binding cysteines: diverse functions and structural motifs. *Biomolecules*, 4.
19. Vázquez-Torres,A. (2012) Redox active thiol sensors of oxidative and nitrosative stress. *Antioxidants Redox Signal.*, 17.
20. Fang,F.C., Frawley,E.R., Tapscott,T. and Vázquez-Torres,A. (2016) Discrimination and Integration of Stress Signals by Pathogenic Bacteria. *Cell Host Microbe*, 20.
21. Fang,F.C., Frawley,E.R., Tapscott,T. and Vázquez-Torres,A. (2016) Bacterial Stress Responses during Host Infection. *Cell Host Microbe*, 20
22. Smith,L.M., Jackson,S.A., Malone,L.M., Ussher,J.E., Gardner,P.P. and Fineran,P.C. (2021) The Rcs stress response inversely controls surface and CRISPR–Cas adaptive immunity to discriminate plasmids and phages. *Nat. Microbiol.*, 6.
23. Pleška,M. and Guet,C.C. (2017) Effects of mutations in phage restriction sites during escape from restriction–modification. *Biol. Lett.*, 13.
24. Murray,N.E. (2000) Type I Restriction Systems: Sophisticated Molecular Machines (a Legacy of Bertani and Weigle). *Microbiol. Mol. Biol. Rev.*, 64.
25. Davies,G.P., Powell,L.M., Webb,J.L., Cooper,L.P. and Murray,N.E. (1998) EcoKI with an amino acid substitution in any one of seven DEAD-box motifs has impaired ATPase and endonuclease activities. *Nucleic Acids Res.*, 26.
26. Bravo,J.P.K., Borodavka,A., Barth,A., Calabrese,A.N., Mojzes,P., Cockburn,J.J.B., Lamb,D.C. and Tuma,R. (2018) Stability of local secondary structure determines selectivity of viral RNA chaperones. *Nucleic Acids Res.*, 46.
27. Fairman-Williams,M.E., Guenther,U.P. and Jankowsky,E. (2010) SF1 and SF2 helicases: Family matters. *Curr. Opin. Struct. Biol.*, 20.
28. Aparicio-Maldonado,C., Ofir,G., Salini,A., Sorek,R., Nobrega,F.L. and Brouns,S.J.J. (2022) Class I DISARM provides anti-phage and anti-conjugation activity by unmethylated DNA recognition. *bioRxiv*.
29. Yu,J., Raia,P., Ghent,C.M., Raisch,T., Sadian,Y., Cavadini,S., Sabale,P.M., Barford,D., Raunser,S., Morgan,D.O., et al. (2021) Structural basis of human separase regulation by securin and CDK1–cyclin B1. *Nature*, 000.
30. Schmidt,C., Kowalinski,E., Shanmuganathan,V., Defenouillère,Q., Braunger,K., Heuer,A., Pech,M., Namane,A., Berninghausen,O., Fromont-Racine,M., et al. (2016) The cryo-EM structure of a ribosome-Ski2-Ski3-Ski8 helicase complex. *Science*, 354.
31. Richards,J.D., Johnson,K.A., Liu,H., McRobbie,A.M., McMahon,S., Oke,M., Carter,L., Naismith,J.H. and White,M.F. (2008) Structure of the DNA repair helicase Hel308 reveals DNA binding and autoinhibitory domains. *J. Biol. Chem.*, 283.
32. Lohman,T.M., Tomko,E.J. and Wu,C.G. (2008) Non-hexameric DNA helicases and translocases: Mechanisms and regulation. *Nat. Rev. Mol. Cell Biol.*, 9.
33. Sloan,K.E. and Bohnsack,M.T. (2018) Unravelling the Mechanisms of RNA Helicase Regulation. *Trends Biochem. Sci.*, 43.

34. Singh,S., Vanden Broeck,A., Miller,L., Chaker-Margot,M. and Klinge,S. (2021) Nucleolar maturation of the human small subunit processome. *Science*, 373.
35. Weigel,C. and Seitz,H. (2006) Bacteriophage replication modules. *FEMS Microbiol. Rev.*, 30.
36. Molineux,I.J. and Panja,D. (2013) Popping the cork: Mechanisms of phage genome ejection. *Nat. Rev. Microbiol.*, 11.
37. Taylor,K. and Węgrzyn,G. (1995) Replication of coliphage lambda DNA. *FEMS Microbiol. Rev.*, 17.
38. Yang,W. (2011) Nucleases: Diversity of structure, function and mechanism.
39. Schwardmann,L.S., Nölle,V. and Elleuche,S. (2020) Bacterial non-specific nucleases of the phospholipase D superfamily and their biotechnological potential. *Appl. Microbiol. Biotechnol.*, 104.
40. Sasnauskas,G., Zakrys,L., Zaremba,M., Cosstick,R., Gaynor,J.W., Halford,S.E. and Siksnys,V. (2010) A novel mechanism for the scission of double-stranded DNA: BfiI cuts both 3' -5' and 5' -3' strands by rotating a single active site. *Nucleic Acids Res.*, 38.
41. Grazulis,S., Manakova,E., Roessle,M., Bochtler,M., Tamulaitiene,G., Huber,R. and Siksnys,V. (2005) Structure of the metal-independent restriction enzyme BfiI reveals fusion of a specific DNA-binding domain with a nonspecific nuclease. *Proc. Natl. Acad. Sci. U. S. A.*, 102.
42. Levy,A., Goren,M.G., Yosef,I., Auster,O., Manor,M., Amitai,G., Edgar,R., Qimron,U. and Sorek,R. (2015) CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature*, 520.
43. Kurilovich,E., Shiriaeva,A., Metlitskaya,A., Morozova,N., Ivancic-bace,I., Severinov,K. and Savitskaya,E. (2019) Genome Maintenance Proteins Modulate Autoimmunity Mediated Primed Adaptation by the Escherichia coli Type I-E CRISPR-Cas System. *Genes (Basel)*.
44. Bernheim,A., Bikard,D., Touchon,M. and Rocha,E.P.C. (2019) A matter of background: DNA repair pathways as a possible cause for the sparse distribution of CRISPR-Cas systems in bacteria. *Philos. Trans. R. Soc. B Biol. Sci.*, 374.
45. Furfaro,L.L., Payne,M.S. and Chang,B.J. (2018) Bacteriophage Therapy: Clinical Trials and Regulatory Hurdles. *Front. Cell. Infect. Microbiol.*, 8.
46. Kostylev,M., Otwell,A.E., Richardson,R.E. and Suzuki,Y. (2015) Cloning should be simple: Escherichia coli DH5 α -mediated assembly of multiple DNA fragments with short end homologies. *PLoS One*, 10.
47. Bonilla,N., Rojas,M.I., Netto Flores Cruz,G., Hung,S.-H., Rohwer,F. and Barr,J.J. (2016) Phage on tap-a quick and efficient protocol for the preparation of bacteriophage laboratory stocks. *PeerJ*, 4.
48. Jia,N., Unciuleac,M.C., Xue,C., Greene,E.C., Patel,D.J. and Shuman,S. (2019) Structures and single-molecule analysis of bacterial motor nuclease AdnAB illuminate the mechanism of DNA double-strand break resection. *Proc. Natl. Acad. Sci. U. S. A.*, 116.

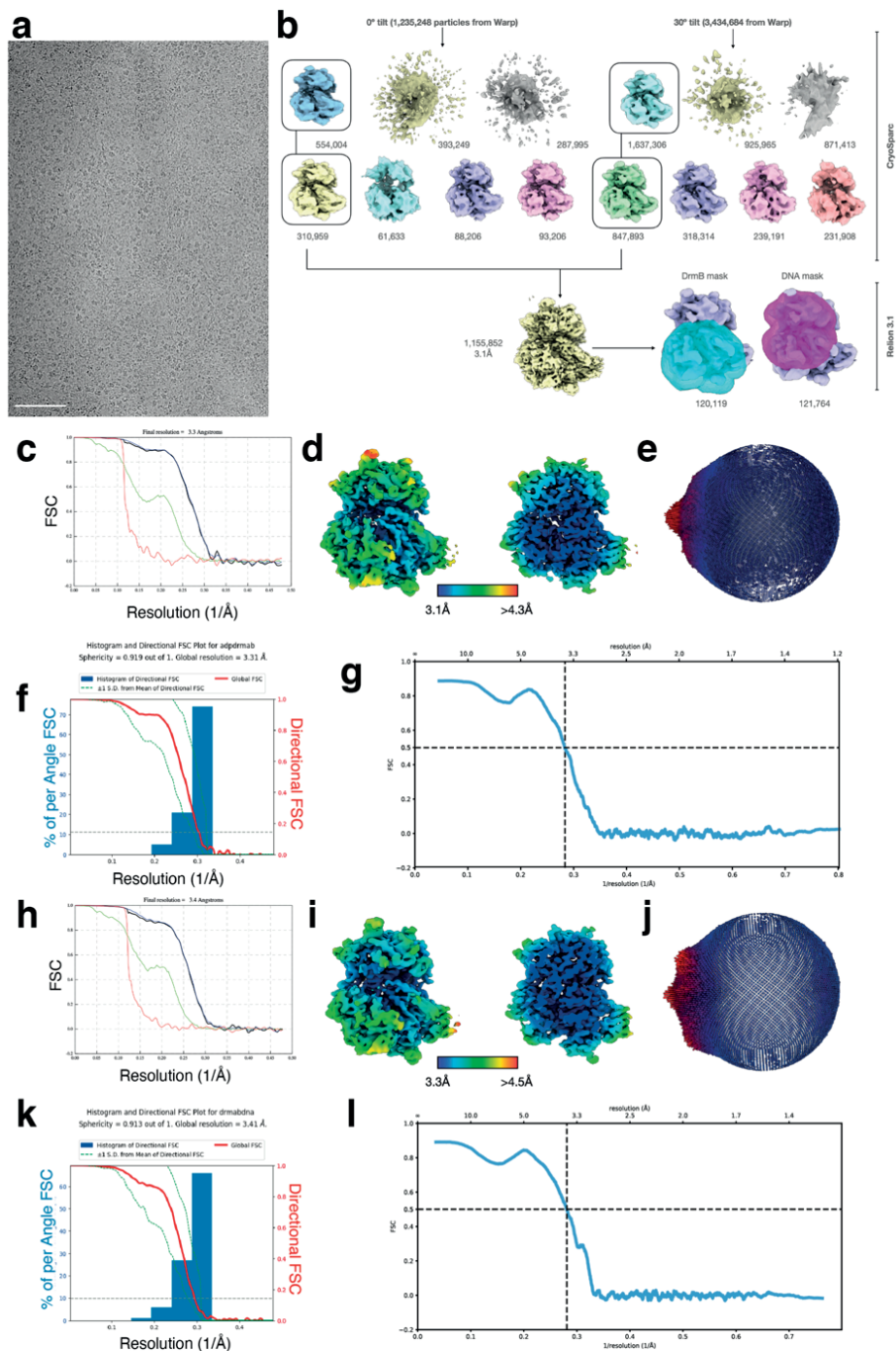
49. Cheng,K., Wilkinson,M., Chaban,Y. and Wigley,D.B. (2020) A conformational switch in response to Chi converts RecBCD from phage destruction to DNA repair. *Nat. Struct. Mol. Biol.*, 27.
50. Mastronarde,D.N. (2005) Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.*, 152.
51. Tegunov,D. and Cramer,P. (2019) Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods*, 16.
52. Punjani,A., Rubinstein,J.L., Fleet,D.J. and Brubaker,M.A. (2017) CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods*, 14.
53. Punjani,A., Zhang,H. and Fleet,D.J. (2020) Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods*, 17.
54. Zivanov,J., Nakane,T., Forsberg,B., Kimanius,D., Hagen,W.J.H.J., Lindahl,E., Scheres,S.H.W., Egelman,E.H. and rn Forsberg,B.O. (2018) RELION-3: new tools for automated high-resolution cryo-EM structure determination. *Elife*.
55. Pettersen,E.F., Goddard,T.D., Huang,C.C., Meng,E.C., Couch,G.S., Croll,T.I., Morris,J.H. and Ferrin,T.E. (2021) UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.*, 30.
56. Yang,J., Anishchenko,I., Park,H., Peng,Z., Ovchinnikov,S. and Baker,D. (2020) Improved protein structure prediction using predicted interresidue orientations. *Proc. Natl. Acad. Sci. U. S. A.*, 117.
57. Kidmose,R.T., Juhl,J., Nissen,P., Boesen,T., Karlsen,J.L. and Pedersen,B.P. (2019) Namdinator - Automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *IUCrJ*, 6.
58. Croll,T.I. (2018) ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. Sect. D*, 74.
59. Afonine,P. V., Poon,B.K., Read,R.J., Sobolev,O. V., Terwilliger,T.C., Urzhumtsev,A. and Adams,P.D. (2018) Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. D Struct. Biol.*, 74.
60. Ashkenazy,H., Abadi,S., Martz,E., Chay,O., Mayrose,I., Pupko,T. and Ben-Tal,N. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.*, 44.
61. Zi Tan,Y., Baldwin,P.R., Davis,J.H., Williamson,J.R., Potter,C.S., Carragher,B. and Lyumkis,D. (2017) Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods*, 14.
62. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Židek,A., Potapenko,A., et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*.

4.6. Supplementary material



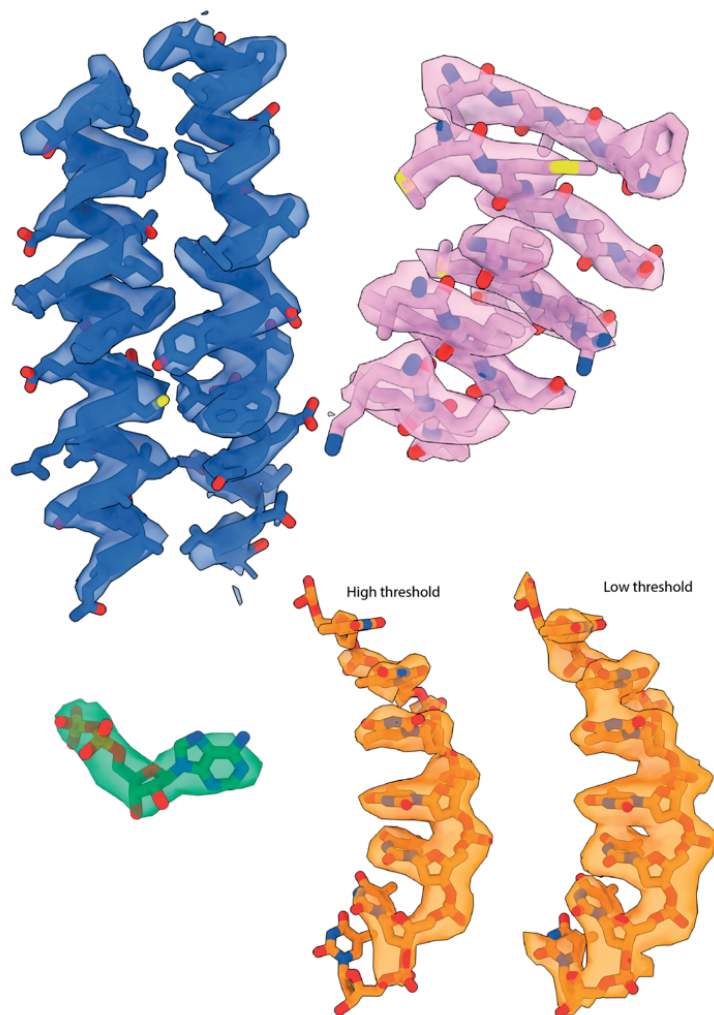
Supplementary Figure 1. DNA binding by DrmAB. a) Size-exclusion chromatogram of DrmAB complex. DrmB was tagged and used to pull down co-expressed DrmAB complex. Fractions 1 – 4

[Sup. Fig. 1]...are analyzed in panel B **b)** SDS PAGE analysis of DrmAB complex purified by size-exclusion chromatography. DrmAB assembles with a 1:1 DrmA:DrmB stoichiometry (fraction 2). Fraction 2 was used for subsequent structural and biochemical analysis. **c)** Top: Secondary structure diagram of DNA stem-loop used for structural analyses (**Sup. Table 2**). 5' region visible in cryo-EM reconstruction is denoted by orange box. Bottom: methylated DNA substrate used in ATPase assay. **d)** Native gel shift assay used to determine sample preparation conditions for cryo-EM. 4 μ M DNA hairpin (**Sup. Table 2**) was heat annealed and incubated with 10 μ M DrmAB in the absence or presence of 1 mM ADP or ADP:AlF₃. Bound and free DNA were separated using EMSA. **E:** EMSA of binding Cy5-labelled DNA with increasing overhang length by annealing complementary DNA of decreasing lengths. Binding is first observed when a length of 7-nt is present. Substrates used in successive overhang binding assay are shown below.

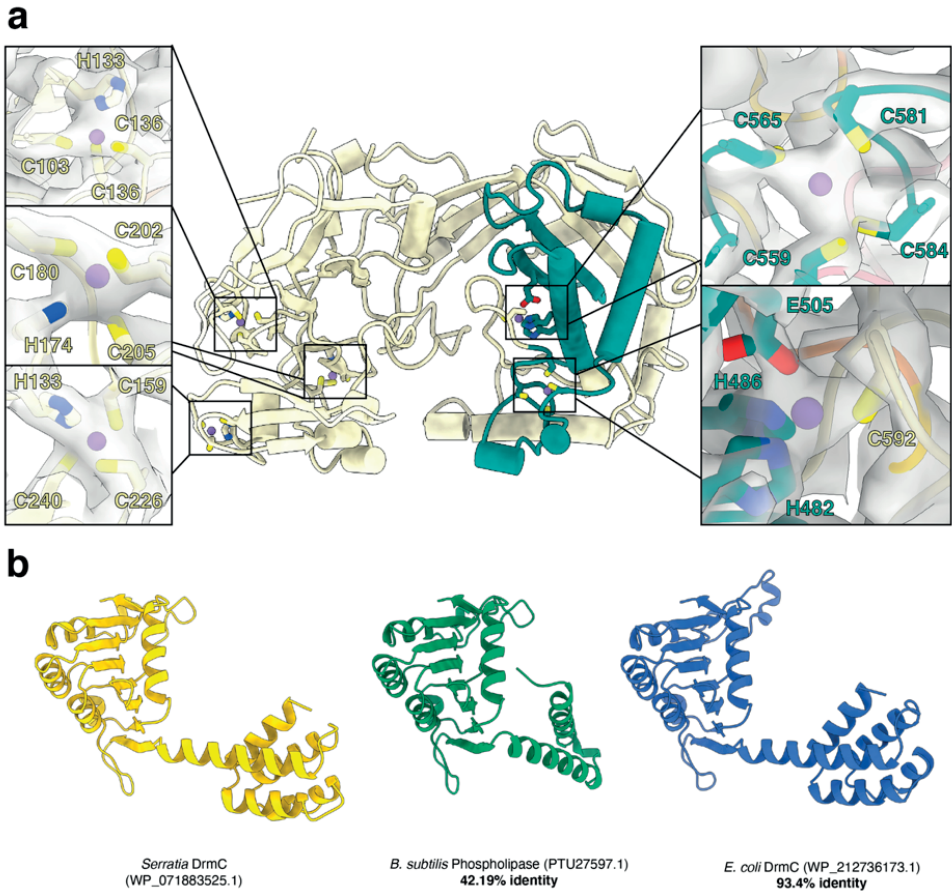


Supplementary Figure 2. Cryo-EM data processing of DrmAB-ADP-DNA dataset. a) Representative cryo-electron micrographs of DrmAB:ADP:DNA. Scale bar – 100 nm. b) Data proc-

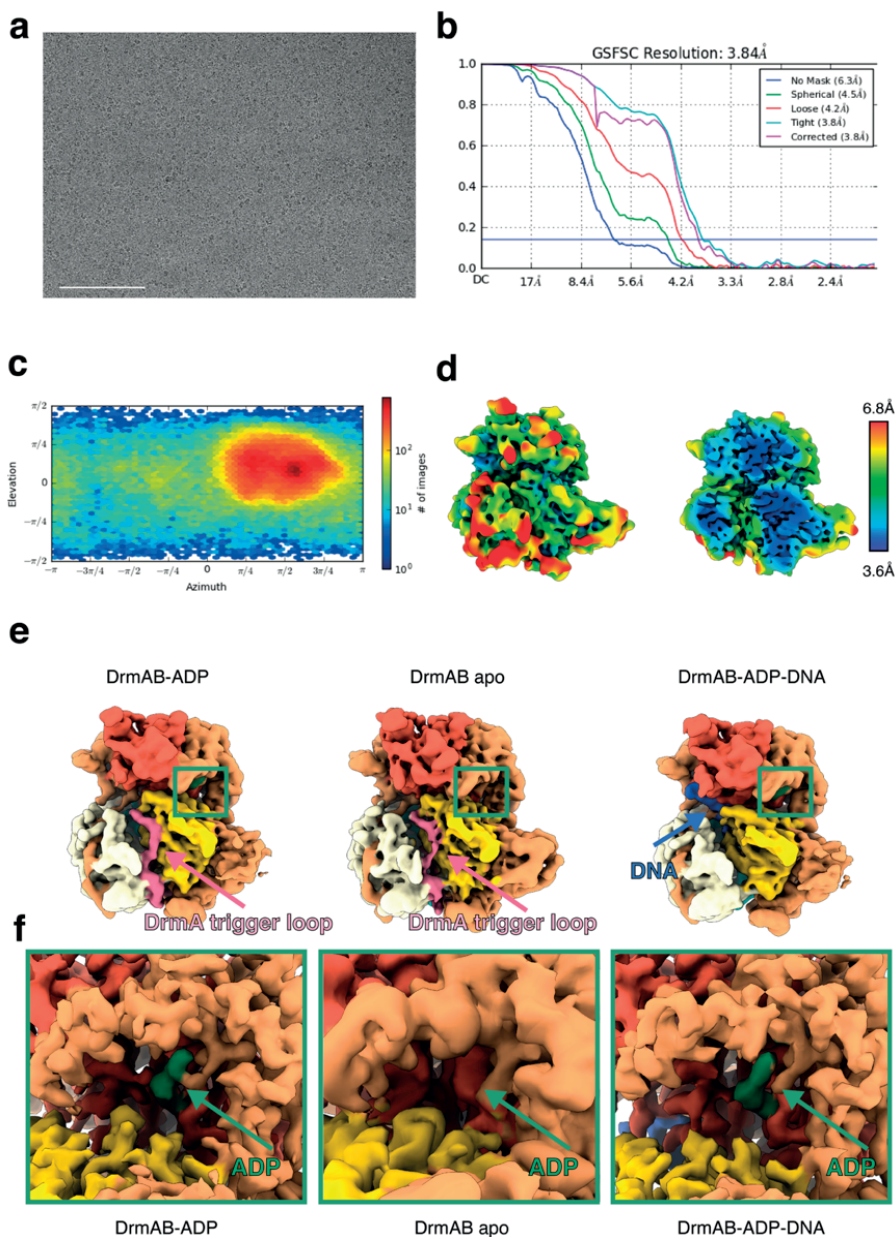
[Sup. Fig. 2]...essing workflow detailing how tilted and untilted particle stacks were combined to a consensus reconstruction. While further classification of said reconstruction yielded a 2.8Å-resolution map, the quality of density corresponding to both ssDNA and DrmB N-terminal half were poor. To ameliorate this, focused 3D classification using either a DrmB mask or a DNA mask was used. **c)** Gold-standard Fourier shell correlation (FSC) curve for DrmAB:ADP. **d)** EM density map of DrmAB:ADP color-coded according to local resolution. **e)** Angular distribution plot calculated in Relion of DrmAB:ADP. **f)** Directional 3D FSC for DrmAB:ADP calculated by 3DFSC⁶¹. **g)** Map-to-model FSC for DrmAB:ADP. **h)** Gold-standard Fourier shell correlation (FSC) curve for DrmAB:ADP:DNA. **i)** EM density map of DrmAB:ADP:DNA color-coded according to local resolution. **j)** Angular distribution plot calculated in Relion of DrmAB:ADP:DNA. **k)** Directional 3D FSC for DrmAB:ADP:DNA calculated by 3DFSC⁶¹. **l)** Map-to-model FSC for DrmAB:ADP:DNA.



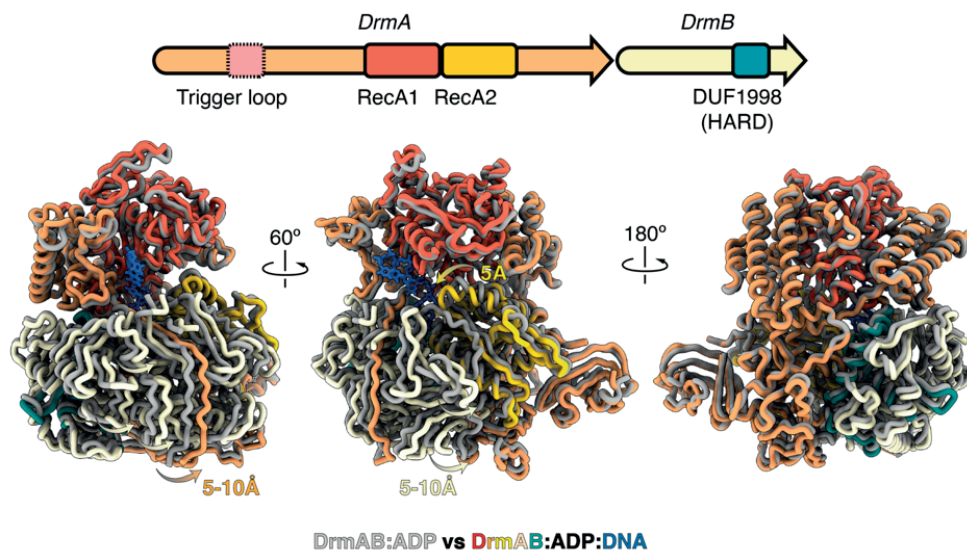
Supplementary Figure 3. Representative regions of DrmAB:ADP:DNA map with corresponding models. Blue: α -helices, pink: β -sheet, green: ADP, orange: DNA. DNA is shown at two different thresholds to illustrate that bases within the center of the molecule are well-resolved, but peripheral bases are less well-resolved, but can still be modelled with confidence.



Supplementary Figure 4. Structural analysis of DrmB and DrmC. **a)** Multiple putative metal coordination sites within DrmB, with representative cryo-EM densities. **b)** Structural models of DrmC from *Serratia* (yellow), a *B. subtilis* PLD-nuclease protein homologue of DrmC (green) and an *E. coli* PLD-nuclease protein homologue of DrmC (blue). The DISARM operon in this study originated in *Serratia*, and *in vivo* assays were performed in *E. coli*. *B. subtilis* was the model organism used for *in vivo* assays in a previous DISARM study ⁷. Structural models were generated using AlphaFold2 ⁶².



Supplementary Figure 5. Structural features of DrmAB complex and TL. **a)** Representative cryo-electron micrograph of DrmAB. Scale bar – 100 nm. **b)** Gold-standard FSC curves for DrmAB. **c)** Angular distribution plot for DrmAB calculated in cryoSPARC. **d)** EM density map of DrmAB color-coded according to local resolution. **e)** Presence of TL EM density features within DrmAB:ADP, DrmAB, and DrmAB:ADP:DNA maps. TL is pink, and DNA is blue. Green box



Supplementary Figure 6. Conformational changes upon DNA binding. Top: Domain architecture schematic of DrmA and DrmB. Bottom: DrmAB:ADP:DNA colored as in the above schematic, and the DrmAB:ADP complex colored in grey, showing conformational changes. DrmA trigger loop is absent in colored DrmAB:ADP:DNA model since it likely becomes disordered or flexible upon DNA binding.

[Sup. Fig. 5]... corresponds to ATP binding site. Φ ATP-binding site within DrmA RecA domains. DrmAB apo complex contains no nucleotide density, but still contains TL density.

Supplementary Table 1. Cryo-EM data collection, refinement and validation statistics.

	DrmAB-ADP (PDB 7S9V, EMD-24938)	DrmAB-ADP-DNA (PDB 7S9W, EMD-24939)	DrmAB
Data collection and processing			
Magnification	22,500		
Voltage (kV)	300 kV		
Electron exposure (e-/Å ²)	80		
Defocus range (μm)	-1.5 to -2.5		
Pixel size (Å)	1.1		
Symmetry imposed	C1		
Initial particle images (no.)	4,669,932		1,540,419
Final particle images (no.)	120,119	121,764	148,062
Map resolution (Å) at 0.143 FSC threshold	3.31	3.41	3.84
Map resolution range (Å)	3.1 to 4.3	3.3 to 4.5	3.6-6.8
Refinement			
Initial model used (PDB code)	De novo		
Model resolution (Å)	3.6	3.8	
FSC threshold	0.5		
Map sharpening <i>B</i> factor (Å ²)	-136.8	-151.1	
Model composition			
Non-hydrogen atoms	27680	28125	
Protein residues	1808	1751	
Nucleotides	0	7	
Ligands	1	1	
Mean <i>B</i> factors (Å ²)			
Protein	62.08	106.66	
Nucleotides	N/A	116.37	
R.m.s. deviations			
Bond lengths (Å)	0.002	0.002	
Bond angles (°)	0.617	0.631	
Validation			
MolProbity score	1.52	1.53	
Clashscore	3.27	3.34	
Poor rotamers (%)	0	0	
Ramachandran plot			
Favored (%)	93.93	93.90	
Allowed (%)	6.01	6.1	
Disallowed (%)	0.6	0	

Supplementary Table 2. DNA oligos used in study for *in vitro* assays.

Oligo name	Sequence (5'-3')
DNA stem loop	TTTTTTTCATCGATGAGCACTGCTATTCCCTAGCAGTGCTCATCGATGATTTTCATCGATGAGCGGTTTT
Methylated DNA stem loop*	TTTTTTTCATmCGATGAGCACTGCTATTCCCTAGCAGTGCTCATmCGATGATTTTCATmCGATGAGCGGTTTT
Fluorescent DNA	Cy5ATTTTGACAGCCACATGGCTTGATGAGTGGCGCACTCGCCAGCCTGAGCATGGCGAAAACCTCCAGTCTGCT
Unlabeled complement DNA	AGCAGACTGGAGGAGTTTTTCGCCATGCTCAGGCTGGCGAGTGCGCCACTCATCAAGCCATGTGGGCTGTCAAAT
5' block	GCCATGTGGGCTGTCAAAT
3' block	AGCAGACTGGAGGAGTTTTC

* mC corresponds to 5-methyl-Cytosine.

4

Supplementary Table 3. DNA oligos used for cloning.

Oligo name	Sequence (5' - 3')	Used for: (in fw/rv couples)	fw/rv
BN1750	TACTTCCAATCCAATGCAATGATCATC AATAACAAAACCTCCAG	Clone <i>drmB</i> into plasmid 13SS using LIC cloning	fw
BN1751	TTATCCACTTCCAATGTTATATTAA AACCCCTAAAAAATGCAGC	Clone <i>drmB</i> into plasmid 13SS using LIC cloning	rv
BN1060	TATATATACATATGACTGATAACAACA AATCTAG	Clone <i>drmA</i> into plasmid pACYC-Duet using RE cloning-NdeI	fw
BN1061	ATATATATCTCGAGTCAATCCTCGTCT TTCGTTG	Clone <i>drmA</i> into plasmid pACYC-Duet using RE cloning-XhoI	rv
BN172	AGATCTGCCATATGTATATCTCCTTC	Amplify pACYC-Duet backbone for cloning of <i>drmA</i> insert	rv
BN2290	GATCGCTGACGTCGGTACCCCTCGAGTC	Amplify pACYC-Duet backbone for cloning of <i>drmA</i> insert	fw
BN1064	TATATATACATATGATGGATGAACTCT TAGATGC	Clone <i>drmC</i> into plasmid pET-Duet using RE cloning-NdeI	fw
BN1065	ATATATATCTCGAGTTAAATCAGACTA ATGATATTGTG	Clone <i>drmC</i> into plasmid pET-Duet using RE cloning-XhoI	rv
BN172	AGATCTGCCATATGTATATCTCCTTC	Amplify pET-Duet backbone for cloning of <i>drmC</i> insert	rv
BN2290	GATCGCTGACGTCGGTACCCCTCGAGTC	Amplify pET-Duet backbone for cloning of <i>drmC</i> insert	fw
BN3096	CATACCCATCTCATTTGTTGAT	Create mutant DrmB-ADUF1998 in 13ss and in pCDF	fw
BN3095	TTATCATTACAGCATGGCATATCGGAT AC	Create mutant DrmB-ADUF1998 in 13ss and in pCDF	rv

BN3104	TTATCATTATCGCATTGACATAGGTAC AG	Create mutant <i>drmA</i> -ΔPA in pCDF	rv
BN3105	GAAGTTGAGTCTGGTGAC	Create mutant <i>drmA</i> -ΔPA in pCDF	fw
BN3436	TTGGTGCCAGTAGCCTC	Create mutant <i>drmA</i> -Δloop in pACYC and pCDF	rv
BN3437	TTCCGGCTGCTTCACGTGG	Create mutant <i>drmA</i> -Δloop in pACYC and pCDF	fw
BN3377	CAACACGACTGGTAGCCTGAATATAC	Amplify DNA from pCDF-Duet with <i>drmA</i> -ΔPA and clone into pACYC-Duet	rv
BN3378	GGACACCAAATTTATCAAGGACGATC	Amplify DNA from pCDF-Duet with <i>drmA</i> -ΔPA and clone into pACYC-Duet	fw
BN3375	GTATATTCAGGCTACCAGTCGTGTG	Amplify DNA from synthetic block to clone in 13SS and pCDF-Duet for <i>drmA</i> mutants R1294A, V1296G, V1296W	fw
BN3376	GATCGTCCTTGATAAATTTGGTGTC	Amplify DNA from synthetic block to clone in 13SS and pCDF-Duet for <i>drmA</i> mutants R1294A, V1296G, V1296W	rv
BN3377	CAACACGACTGGTAGCCTGAATATAC	Amplify vectors 13SS and pCDF-Duet DNA to insert amplified DNA of <i>drmA</i> mutants R1294A, V1296G, V1296W	rv
BN3378	GGACACCAAATTTATCAAGGACGATC	Amplify vectors 13SS and pCDF-Duet DNA to insert amplified DNA <i>drmA</i> mutants R1294A, V1296G, V1296W	fw
BN3379	GATGAACAGAACTTGAATGTGAGC	Amplify DNA from synthetic block to clone in 13SS and pCDF-Duet for <i>drmA</i> mutants R569D, Q572D, K610D, R658D, R649A, K803A, R810A	rv
BN3380	GATAATTAAATCAGGCGGTCGGATAG	Amplify DNA from synthetic block to clone in 13SS and pCDF-Duet for <i>drmA</i> mutants R569D, Q572D, K610D, R658D, R649A, K803A, R810A	fw
BN3381	GCTCACATTCAAGTTCTGTTTCATC	Amplify vectors 13SS and pCDF-Duet DNA to insert amplified DNA <i>drmA</i> mutants R569D, Q572D, K610D, R658D, R649A, K803A, R810A	rv
BN3382	CTATCCGACCGCCTGATTTAATTATC	Amplify vectors 13SS and pCDF-Duet DNA to insert amplified DNA <i>drmA</i> mutants R569D, Q572D, K610D, R658D, R649A, K803A, R810A	fw

Supplementary Table 4. Plasmids used in this study.

Plasmid	Description	Antibiotic resistance marker	Source
pTU515	13SS- <i>drmB</i> fused to His-tag	Spectinomycin	This paper
pTU518	13SS- <i>drmB</i> - Δ DUF1998 fused to His-tag	Spectinomycin	This paper
pTU516	pACYC-Duet- <i>drmA</i>	Chloramphenicol	This paper
pTU520	pACYC-Duet- <i>drmA</i> (Δ PA)	Chloramphenicol	This paper
pTU542	pACYC-Duet- <i>drmA</i> (Δ loop)	Chloramphenicol	This paper
pTU522	pACYC-Duet- <i>drmA</i> (V1296G)	Chloramphenicol	This paper
pTU523	pACYC-Duet- <i>drmA</i> (V1296W)	Chloramphenicol	This paper
pTU524	pACYC-Duet- <i>drmA</i> (R569D)	Chloramphenicol	This paper
pTU525	pACYC-Duet- <i>drmA</i> (Q572D)	Chloramphenicol	This paper
pTU526	pACYC-Duet- <i>drmA</i> (K610D)	Chloramphenicol	This paper
pTU527	pACYC-Duet- <i>drmA</i> (R658D)	Chloramphenicol	This paper
pTU528	pACYC-Duet- <i>drmA</i> (R649A)	Chloramphenicol	This paper
pTU529	pACYC-Duet- <i>drmA</i> (K803A)	Chloramphenicol	This paper
pTU530	pACYC-Duet- <i>drmA</i> (R810A)	Chloramphenicol	This paper
pTU531	pACYC-Duet- <i>drmA</i> (R1294A)	Chloramphenicol	This paper
pTU517	pET-Duet- <i>drmC</i>	Ampicillin	Aparicio-Maldonado et al.
pTU495	pCDF-Duet- <i>drmABC</i>	Streptomycin	Aparicio-Maldonado et al.
pTU519	pCDF-Duet- <i>drmABC</i> with <i>drmB</i> - Δ DUF1998	Streptomycin	This paper
pTU521	pCDF-Duet- <i>drmABC</i> with <i>drmA</i> (Δ PA)	Streptomycin	This paper
pTU543	pCDF-Duet- <i>drmABC</i> with <i>drmA</i> (Δ loop)	Streptomycin	This paper
pTU532	pCDF-Duet- <i>drmABC</i> with <i>drmA</i> (V1296G)	Streptomycin	This paper
pTU533	pCDF-Duet- <i>drmABC</i> with <i>drmA</i> (V1296W)	Streptomycin	This paper
pTU534	pCDF-Duet- <i>drmABC</i> with <i>drmA</i> (R569D)	Streptomycin	This paper
pTU535	pCDF-Duet- <i>drmABC</i> with <i>drmA</i> (Q572D)	Streptomycin	This paper
pTU536	pCDF-Duet- <i>drmABC</i> with <i>drmA</i> (K610D)	Streptomycin	This paper
pTU537	pCDF-Duet- <i>drmABC</i> with <i>drmA</i> (R658D)	Streptomycin	This paper
pTU538	pCDF-Duet- <i>drmABC</i> with <i>drmA</i> (R649A)	Streptomycin	This paper
pTU539	pCDF-Duet- <i>drmABC</i> with <i>drmA</i> (K803A)	Streptomycin	This paper
pTU540	pCDF-Duet- <i>drmABC</i> with <i>drmA</i> (R810A)	Streptomycin	This paper
pTU541	pCDF-Duet- <i>drmABC</i> with <i>drmA</i> (R1294A)	Streptomycin	This paper

5.

Short pAgos trigger cell death upon detection of invading DNA

Published as: *Short prokaryotic Argonaute systems trigger cell death upon detection of invading DNA*. Balwina Koopal, Ana Potocnik, Sumanth K. Mutte, Cristian Aparicio-Maldonado, Simon Lindhoud, Jacques J. M. Vervoort, Stan J. J. Brouns, Daan C. Swarts (2022). **Cell**. doi: 10.1016/j.cell.2022.03.012

Abstract

Argonaute proteins use single stranded RNA or DNA guides to target complementary nucleic acids. This allows eukaryotic Argonaute proteins to mediate RNA interference and long prokaryotic Argonaute proteins to interfere with invading nucleic acids. The function and mechanisms of the phylogenetically distinct short prokaryotic Argonaute proteins remain poorly understood. We demonstrate that short prokaryotic Argonaute and associated TIR-APAZ (SPARTA) proteins form heterodimeric complexes. Upon guide RNA-mediated target DNA binding four SPARTA heterodimers form oligomers in which TIR domain-mediated NAD(P)ase activity is unleashed. When expressed in *Escherichia coli*, SPARTA is activated in presence of highly-transcribed multi-copy plasmid DNA, which causes cell death through NAD(P)⁺ depletion. This results in removal of plasmid-invaded cells from bacterial cultures. Furthermore, we show that SPARTA can be repurposed for programmable detection of DNA sequences. In conclusion, our work identifies SPARTA as a prokaryotic immune system that reduces cell viability upon RNA-guided detection of invading DNA.

5.1. Introduction

Argonaute proteins recruit small (~15-30 nt) single stranded guide oligonucleotides to bind complementary single stranded target sequences. Eukaryotic Argonaute proteins (eAgos) use guide RNAs to bind and/or cleave target RNA, a process that underlies RNA interference (RNAi) ^{1,2}. Prokaryotes encode homologs of eAgos (pAgos) but lack other proteins required for RNAi ³. Most characterized pAgos share a fixed four-domain composition (N-PAZ-MID-PIWI) with eAgos. These “long” pAgos interfere with invading nucleic acids such as plasmids ^{4,5} and bacteriophages ⁶. Long pAgos have also been implicated in genome decatenation ⁷ and stimulation of homologous recombination ^{8,9}. In contrast to eAgos, certain long pAgos utilize single stranded (ss)DNA guides to target invading DNA sequences ^{5,6,10,11}. Other long pAgos use guide RNAs to act on RNA and/or DNA targets ^{4,12–16}. While our knowledge on the function and mechanisms of long pAgos is rapidly expanding, the majority of pAgos (58%) consist of poorly studied “short” pAgos ¹⁷.

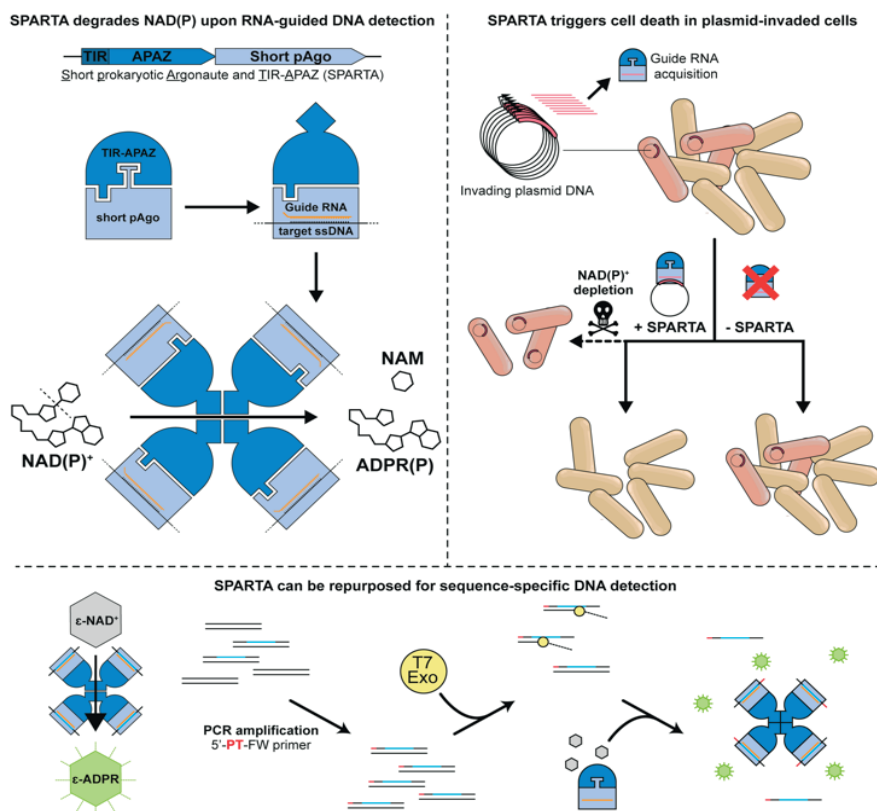
Short pAgos are comprised of the MID and PIWI domains only ^{17–19}. Akin to certain eAgos and long pAgos, short pAgos lack the catalytic tetrad required for target cleavage ¹⁹, which suggests they are catalytically inactive. However, MID domain residues required for guide 5' phosphate (5'-P) binding are conserved in short pAgos ¹⁷, and the MID and PIWI domains alone are sufficient to facilitate guide-mediated target binding in truncated long Agos ^{20–23}. This suggests small pAgos mediate 5'-P guide-mediated target binding.

Short pAgos are encoded in the same operon as (or are fused to) a domain that is named “Analog of PAZ” (APAZ) ¹⁸. However, APAZ domains are homologous to the N-domain of long pAgos rather than to their PAZ domain ^{24,25}. APAZ domains are commonly fused to SIR2, Toll/Interleukin-1 Receptor (TIR)-like, or DUF4365 domains ^{17,19}. It was predicted that these domains are nucleases that compensate for the loss of catalytic activity of the associated short pAgos ²⁶, and it was recently proposed that a short pAgo complexed with a DUF4365-APAZ domain mediates RNA-guided DNA cleavage ²⁷. However, most short pAgo-

associated domains are fused to a SIR2 or TIR domain. The function and mechanisms of these short pAgos remain unexplored.

Here, we describe the role and mechanism of short prokaryotic Argonaute TIR-APAZ (SPARTA) systems. We demonstrate that short pAgo and TIR-APAZ form a heterodimeric SPARTA complex. Upon small RNA-guided ssDNA target binding four SPARTA heterodimers form an oligomer in which the NAD(P)ase activity of the TIR domain is unleashed. *In vivo*, SPARTA is activated in presence of highly transcribed multi-copy plasmid DNA. This results in cellular NAD(P)⁺ depletion and consequentially reduces viability of cells harbouring invading plasmid DNA. As a result, uninvaded cells outcompete invaded cells, which drives the plasmid DNA to extinction. Finally, we show that SPARTA can be repurposed to detect DNA sequences of choice.

Graphic abstract



5.2. Results

5.2.1. Short pAgos cluster with APAZ-domain containing proteins only

To determine if short pAgos genetically cluster with proteins other than APAZ domain-containing proteins, we first identified 283 short pAgo homologs in the RefSeq database containing at least scaffold level assemblies. All short pAgos identified are comprised of the MID and PIWI domains and lack a catalytic DEDX tetrad. Short pAgos form four phylogenetic subclades (**Fig. 1A**) which we named S1A (SIR2-APAZ-pAgo fusions), S1B (operon with SIR2-APAZ and pAgo), S2A (operons with (Mrr)-TIR-APAZ and pAgo), and S2B (operons mainly consisting of pAgo and DUF4365-APAZ or DHS-like-APAZ). We coin the term ‘short pAgo system’ to indicate a short pAgo and its associated APAZ domain-containing protein (clades S1B, S2A, and S2B), or a fusion thereof (clade S1A). In all subclades, the APAZ-domain containing protein is the only protein that specifically associates with pAgo. This suggests pAgo systems function as stand-alone systems or rely on common host factors for their activity.

We name S2A systems ‘SPARTA systems’, for short prokaryotic Argonaute TIR-APAZ systems. In one subclade of SPARTA systems, a putative Mrr (Modification requiring restriction endonuclease) domain is fused to the TIR-APAZ protein (Mrr-TIR-APAZ; **Fig. 1A**). Short pAgo and associated TIR-APAZ proteins co-evolved in Bacteroidetes and in one lineage of Proteobacteria (**Fig. 1B** and **Sup. Fig. S1A**), which suggests they function in conjunction. In another lineage of Proteobacteria short pAgo and TIR-APAZ show distinct evolutionary patterns. This suggests there might be a functional difference between SPARTA systems found in distinct phylogenetic clades.

It has been suggested that APAZ-associated domains compensate for the lack of nuclease activity in short pAgos ²⁶. However, in both eukaryotic and prokaryotic immune systems, TIR domains degrade NAD(P)⁺ ^{28–32}. The catalytic residues required for NAD(P)ase activity are conserved in TIR-APAZ proteins (**Fig. S1B**), which suggests they might have a similar mode of action. Based on

the conservation of the 5'-phosphate (5'-P) binding residues in the pAgo MID domain (**Fig. S1C**), we hypothesized that in SPARTA systems short pAgos facilitate guide-mediated target binding. Consequently, TIR-APAZ could either mediate endonuclease activity (compensating for the lack of endonuclease activity of short pAgo) or NAD(P)⁺ nucleosidase activity (based on homology to other TIR domains).

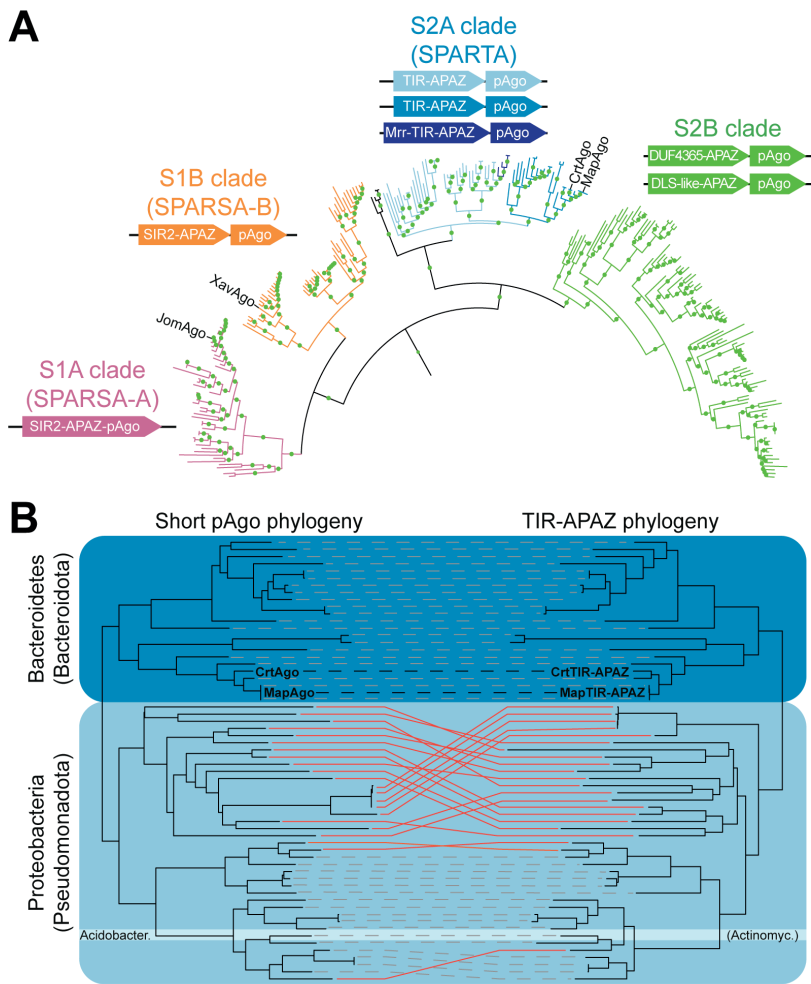


Figure 1. Phylogenetic analysis of short pAgo systems. **A)** Maximum-likelihood based phylogenetic analysis of 283 short pAgo proteins. Clade S1A: SIR2-APAZ-pAgo fusion (SPARSA-A). Clade S1B: operon with pAgo and SIR2-APAZ (SPARSA-B). Clade S2A: operon with pAgo and TIR-APAZ (SPARTA) or Mrr-TIR-APAZ. Clade S2B: operon with pAgo and APAZ fused to one of various domains, including DUF4365 or DHS-like domains. CrtAgo: *Crenotalea thermophila* pAgo. MapAgo: *Maribacter polysiphoniae* pAgo. Green circles

5.2.2. Short pAgo and TIR-APAZ proteins form a heterodimeric complex

To biochemically characterize SPARTA systems we co-expressed SPARTA proteins from *Maribacter polysiphoniae* (MapAgo/MapTIR-APAZ) or from *Crenotalea thermophila* (CrtAgo/CrtTIR-APAZ) in *Escherichia coli* (**Fig. 2A**). These pAgo and TIR-APAZ proteins share 81% and 73% sequence identity, respectively. For both systems N-terminally 6xHis-MBP-tagged TIR-APAZ could be used to pull-down untagged short pAgo. After removal of the 6xHis-MBP tag, the short pAgo and TIR-APAZ of both systems remained complexed throughout multiple chromatography steps (**Fig. 2B**). SEC–Multi-Angle Light Scattering (SEC-MALS) analysis of the purified complexes reveals that short pAgo and TIR-APAZ form a heterodimeric complex with a 1:1 stoichiometry (**Fig. 2C**).

5.2.3. TIR-APAZ converts NAD(P)⁺ to NAM and ADPR(P)

To determine if SPARTA shows nuclease activity, MapSPARTA and CrtSPARTA were incubated with 5' phosphorylated (5'-P) 21 nucleotide (nt) long single stranded (ss)RNA or ssDNA guides and complementary ssRNA or ssDNA targets. SPARTA does not facilitate guide-mediated target cleavage for any guide/target combination (**Fig. S2A**). Next, we investigated if TIR-APAZ shows NAD(P)ase activity. As MapTIR-APAZ is unstable when purified in absence of MapAgo, we focussed on CrtTIR-APAZ in these experiments.

LC-MS and ¹H NMR spectroscopy reveal that CrtTIR-APAZ is a nucleosidase that converts NAD(P)⁺ into non-cyclic ADPR(P) and NAM (**Fig. 2D-F** and **Sup. Fig. S2B-D**). Other TIR domains also convert NAD(P)⁺ to NAM and ADPR(P) or structural analogues thereof (cADPR or v-ADPR)^{28,29,31–33}. In contrast, CrtAgo and CrtSPARTA show no NADase activity (**Fig. 2D**), suggesting short pAgo controls the activity of TIR-APAZ.

[Fig. 1]... indicate bootstrap values >75%. SIR2: Interpro IPR039444. TIR: Interpro IPR000157. Mrr: Interpro IPR007560. DUF4365: InterPro IPR025375. DHS-like: InterPro IPR029035. B) Tanglegram of phylogenetic trees of 56 short pAgo proteins (left) and associated TIR-APAZ proteins (right). Dashed lines indicate co-evolved proteins, red lines indicate differential evolution. *Acidobacter*: Acidobacteria (*Actinomycetota*). CrtTIR-APAZ: *C. thermophila* TIR-APAZ. MapTIR-APAZ: *M. polysiphoniae* TIR-APAZ. See also **Sup. Fig. S1**.

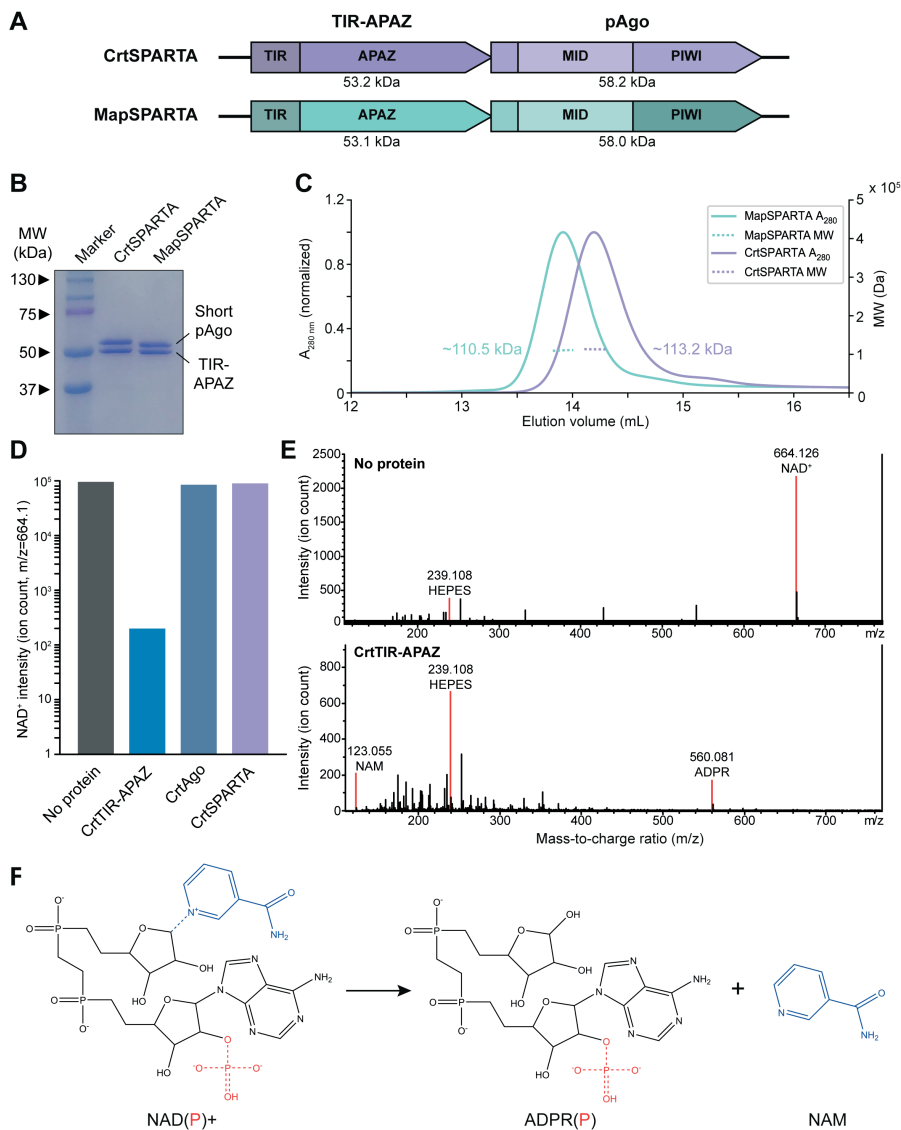


Figure 2. TIR-APAZ NAD(P)ase activity is controlled by short pAgo. **A)** Schematic diagram of CrtSPARTA and MapSPARTA operon structure and domain organization. **B)** TIR-APAZ binds to short pAgo. SDS-PAGE of co-purified TIR-APAZ and short pAgo proteins. **C)** TIR-APAZ and short pAgo form SPARTA complexes with a 1:1 stoichiometry. The molecular weight (MW) determined by size exclusion chromatography-multi-angle light scattering corresponds to theoretical MW of heterodimeric CrtAgo-CrtTIR-APAZ (111.4 kDa) and MapAgo-MapTIR-APAZ (111.1 kDa) complexes. **D-E)** TIR-APAZ is an NADase and short pAgo inhibits TIR-APAZ activity. LC-MS ion

5.2.4. Guide RNA-mediated target ssDNA binding activates SPARTA

We speculated that SPARTA could be catalytically activated upon guide-mediated target binding. MapSPARTA and CrtSPARTA were incubated with various guide/target combinations (**Fig. 3A**) and ϵ -NAD⁺, which is converted into NAM and fluorescent ϵ -ADPR by NADase activity. Apo-SPARTA, SPARTA with guide RNA or DNA alone, or with guide RNA and target RNA show no NADase activity (**Fig. 3B** and **Sup. Fig. S3A**). In contrast, both MapSPARTA and CrtSPARTA degrade NAD⁺ when provided with a guide RNA and a complementary target ssDNA. When the guide RNA and ssDNA are non-complementary SPARTA does not degrade NAD⁺, implying that activation of SPARTA is sequence-specific. Alanine substitution of Glu77 in the catalytic site of the TIR domain (CrtSPARTA^{TIR-E77A}) abolishes NADase activity (**Fig. S3B**). This confirms that the TIR domain confers the nucleosidase activity of SPARTA.

Albeit at lower efficiencies than guide RNA/target ssDNA, also guide ssDNA/target RNA combinations activate MapSPARTA (3.4-fold lower activity) and CrtSPARTA (1.5-fold lower activity), and guide ssDNA/target ssDNA combinations activate CrtSPARTA (3.4-fold lower activity) (**Fig. 3B** and **S3A**). However, the guide 5' group (5'-P or 5'-OH) has no effect for guide ssDNA/target RNA combinations (**Fig. 3C** and **Sup. Fig. S3C**), suggesting guide ssDNAs might not be properly bound by the MID domain. In contrast, NAD⁺ turnover is 8.5-fold (MapSPARTA) or 3.6-fold (CrtSPARTA) lower for 5'-OH guide RNAs compared to 5'-P guide RNAs. This implies that SPARTA is preferentially activated by 5'-P guide RNA-mediated target ssDNA binding.

[Fig. 2]... count (**D**) and spectra (**E**) for NAD⁺ after incubation with CrtSPARTA proteins. NAD⁺ is converted to nicotinamide (NAM) and adenosine diphosphate ribose (ADPR) upon incubation with CrtTIR-APAZ, but not when incubated with CrtAgo or CrtSPARTA. **F**) Reaction structural formulas of TIR-APAZ-mediated NAD(P)⁺ hydrolysis. Blue dotted line: scissile bond. See also **Sup. Fig. S2**.

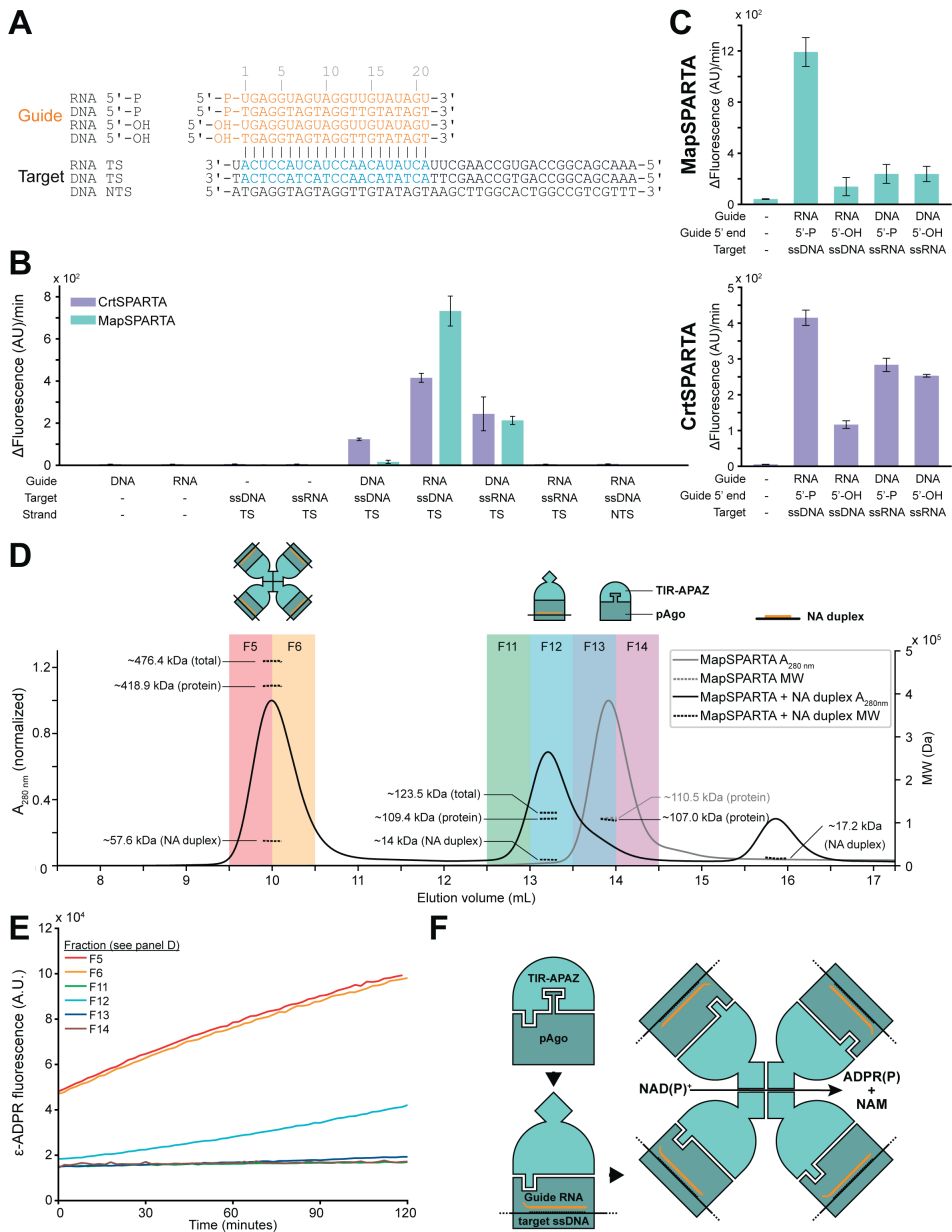


Figure 3. SPARTA degrades NAD⁺ upon RNA-guided ssDNA binding. A) Schematic representation of guide (orange) and target (grey, target sequence: blue) oligonucleotides used in panels (B) and (C). B-C) MapSPARTA or CrtSPARTA was mixed with guide and target oligonucleotides in a 1:1:2 molar ratio and incubated with ϵ -NAD⁺. Graphs indicate the maximum change in fluorescence over time. The average of three technical replicates is shown, error bars

Furthermore, guide RNAs that are 15-50nt in length facilitate target ssDNA-mediated SPARTA activation (**Fig. S3E** and **Sup. Fig. S3F**). This suggests guide RNA 3' end-binding is not required for SPARTA activation. MapSPARTA is also activated by longer RNA/DNA hybrids (**Sup. Fig. S3G**), possibly because they resemble guide RNA/target ssDNA ligands. dsRNA and dsDNA ligands do not trigger robust SPARTA activation, but when guide RNAs are provided, short (≤ 60 bp) but not long (≥ 152 bp) dsDNA targets trigger activation of SPARTA (**Fig. S3H**). Possibly their small size allows for unwinding which facilitates recognition by the SPARTA-guide RNA complex. Combined, these results show that TIR-mediated NAD(P)ase activity of SPARTA is preferentially activated by 5'-P guide RNA-mediated binding of ssDNA targets.

[Fig. 3]... indicate standard deviations. **B**) SPARTA is preferentially activated upon by guide RNA-mediated ssDNA target binding. **C**) SPARTA prefers 5'-P guides over 5'-OH guides. CrtSPARTA 5'-P guide samples are identical in **(B)** and **(C)**. **D**) SPARTA oligomerizes upon guide RNA-mediated target ssDNA binding. MapSPARTA was incubated with guide RNA and target ssDNA in a 1:1:1 molar ratio and subjected to size exclusion chromatography-multi-angle light scattering analysis. Coloured backgrounds indicate fractions that were tested for NADase activity **(E)**. Control without nucleic acids is identical to that in **Fig. 2C**. **E**) SPARTA is catalytically activated upon oligomerization. SEC-MALS fractions **(D)** were incubated with ϵ -NAD⁺ and change in fluorescence was determined over time. **F**) Proposed mechanism for SPARTA activation. TIR-APAZ and associated short pAgo form a catalytically inactive heterodimeric SPARTA complex. Guide RNA-mediated target ssDNA binding induces conformational changes that facilitate oligomerization. Guide/target-bound SPARTA heterodimers form a complex with a 4:4:4:4 (TIR-APAZ:short pAgo:guide:target) stoichiometry in which the NAD(P)ase activity of the TIR domain is unleashed. See also **Sup. Fig. S3a** 1:1 stoichiometry (MapSPARTA:NA duplex), while the second peak corresponds to an oligomeric complex with a 4:4 stoichiometry. Elution fractions containing 4:4 complexes show substantial NADase activity directly upon incubation with ϵ -NAD⁺ while the fraction containing 1:1 complexes only becomes active upon prolonged incubation (**Fig. 3E**). We speculate that the 1:1 complexes oligomerize during the assay, which renders them active. In conclusion, guide RNA-mediated ssDNA target binding triggers the formation of an oligomeric complex with a 4:4:4:4 stoichiometry (TIR-APAZ:short pAgo:guide:target) in which the NAD(P)ase activity of the TIR domain is unleashed (**Fig. 3F**).

5.2.5. SPARTA is catalytically activated upon oligomerization

As TIR-APAZ is catalytically active in absence of pAgo, we speculate it might be released from the SPARTA complex upon activation. However, SEC-MALS analysis of MapSPARTA pre-incubated with a guide RNA/ssDNA target nucleic acid (NA) duplex reveals two distinct elution peaks with molecular weights (MW) larger than that of apo-MapSPARTA (**Fig. 3D**). Conjugate analysis shows that the first peak corresponds to MapSPARTA bound to the NA duplex in a 1:1 stoichiometry (MapSPARTA:NA duplex), while the second peak corresponds to an oligomeric complex with a 4:4 stoichiometry. Elution fractions containing 4:4 complexes show substantial NADase activity directly upon incubation with ϵ -NAD⁺ while the fraction containing 1:1 complexes only becomes active upon prolonged incubation (**Fig. 3E**). We speculate that the 1:1 complexes oligomerize during the assay, which renders them active. In conclusion, guide RNA-mediated ssDNA target binding triggers the formation of an oligomeric complex with a 4:4:4:4 stoichiometry (TIR-APAZ:short pAgo:guide:target) in which the NAD(P)ase activity of the TIR domain is unleashed (**Fig. 3F**).

5.2.6. SPARTA recruits guide RNAs to target highly expressed genes

To shed light on the biological role of SPARTA we examined the nucleic acids it associates with *in vivo*. 5' group-dependent ³²P nucleic acid labelling reveals that MapSPARTA co-purifies with 5'-P RNAs that are predominantly ~15-25nt or ~40nt in length (**Fig. 4A**). MapSPARTA also co-purifies with 5'-OH RNAs of various lengths and with DNAs that are >50nt in length (**Fig. 4C** and **Sup. Fig. S4A**).

A preference for a specific base at the 5'-end of the guide is a general feature of Argonaute proteins^{4,5,7,34}. Indeed, small RNAseq reveals that MapSPARTA-associated small RNAs have a bias toward an adenine base at the 1st (73%) and 2nd (64%) position at the 5' end, which is not observed for *E. coli* extracted small RNAs (**Fig. 4E**). This implies that these small RNAs are specifically bound by MapSPARTA. No such bias is observed for long RNAs

associated with MapSPARTA, which suggests these RNAs co-purify non-specifically.

Abundance of specific sequences in the transcriptome and corresponding MapSPARTA-associated small RNA pools are strongly correlated (**Fig. 4F**; Pearson Correlation coefficient $r > 0.99$, $p < 10^{-99}$). This implies that the abundance of RNA sequences in the transcriptome determines the chance of SPARTA being loaded with corresponding small RNAs. Of note, the MapSPARTA-associated 39-41 nt long RNAs are tRNA fragments (tRFs). tRFs also associate with RsAgo⁴ and eAgos^{24,35,36}, but the biological relevance is unclear.

The correlation between the abundance of mRNAs and MapSPARTA-associated small RNAs suggests that MapSPARTA recruits small RNAs to target highly expressed genes. Interestingly, plasmid-encoded mRNA is more abundant than genome-encoded mRNA in the transcriptome (**Fig. 4D**), which is probably caused by a combination of strong promoters and the multi-copy nature of the MapSPARTA expression plasmid (15-20 copies per cell). Of note, the high number of reads assigned to *E. coli* genes *lacI* and *malE* are likely transcribed from the plasmid-encoded *lacI* and the MBP-tag of *MBP-MapTIR-APAZ*, which share respectively 100% and 99.8% sequence identity to the *lacI* and *malE* on the genome of *E. coli*. Results of CrtSPARTA-associated RNA sequencing are almost identical to that of MapSPARTA-associated RNA (**Sup. Fig. S4**), but the length of associated small RNAs is less-well defined and no tRF peak is observed for CrtSPARTA.

Sequencing of MapSPARTA-associated DNA confirms that it associates with both plasmid and genomic DNA, and shows a relatively weak but highly significant correlation between specific DNA and small RNA sequence abundance (**Fig. 4G**; Pearson Correlation coefficient $r > 0.32$, $p < 10^{-99}$). In line with the observation that SPARTA is preferentially activated by 5'-P guide RNAs-mediated ssDNA *in vitro*, these results suggest that *in vivo* SPARTA recruits 5'-P guide RNAs with 5' adenosine bases from the transcriptome to target highly transcribed DNA.

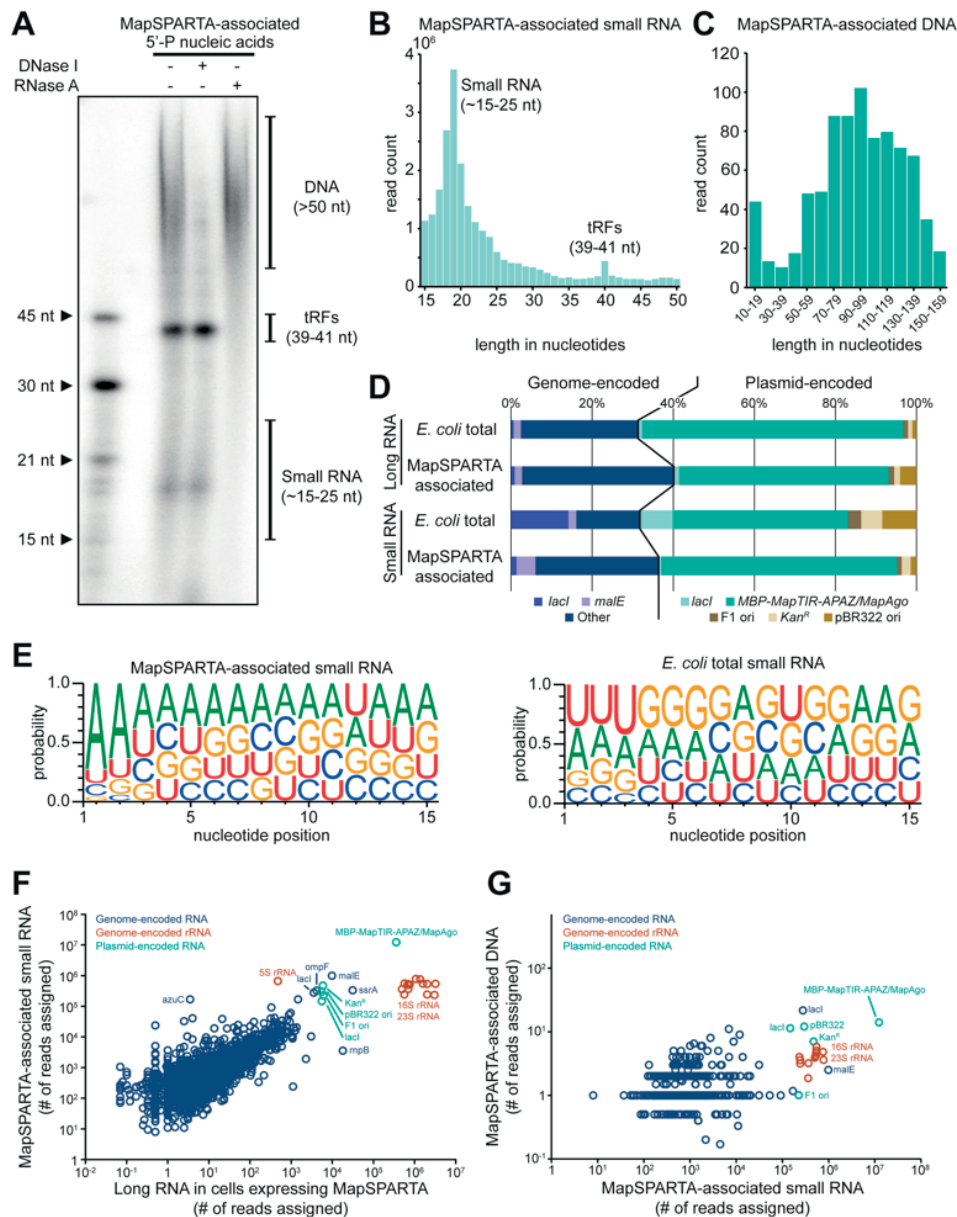


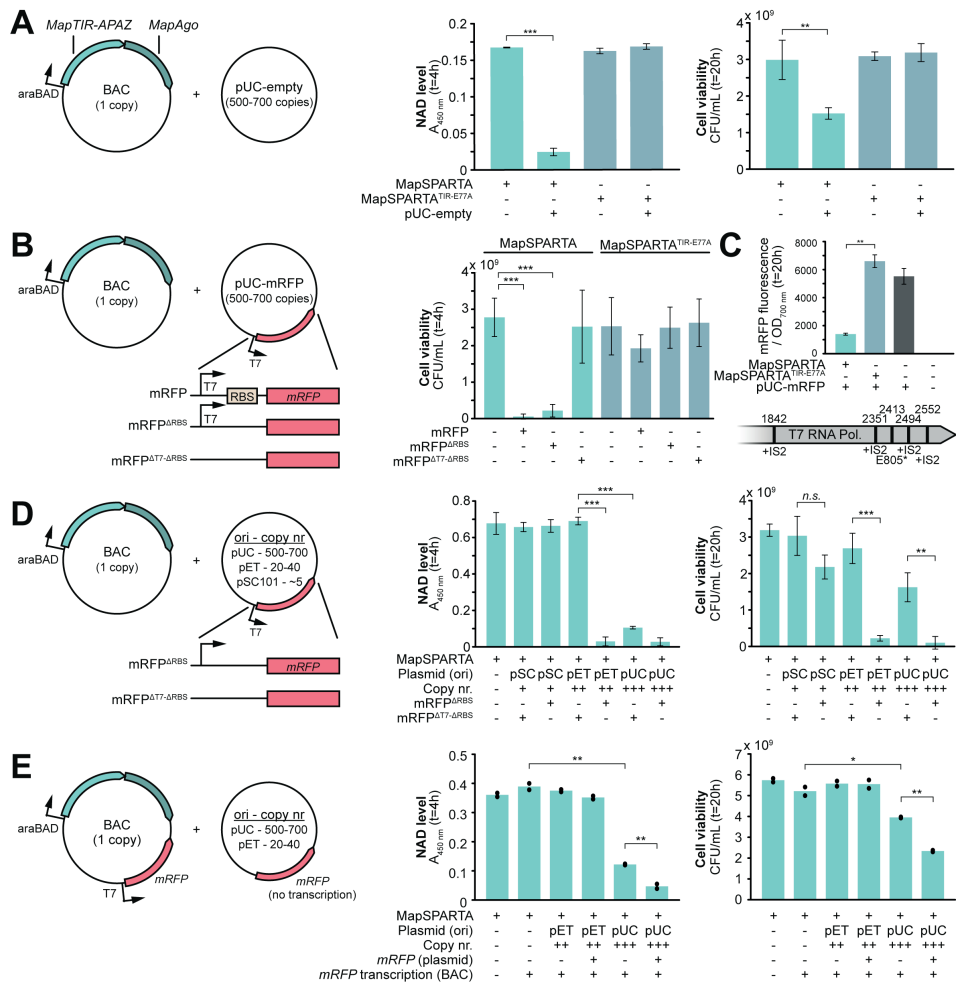
Figure 4. MapSPARTA associates with guide RNAs targeting highly transcribed genes. A) MapSPARTA associates with 5' phosphorylated (5'-P) small RNAs. Nucleic acids that co-purified with MapSPARTA were [γ - 32 P]-ATP labelled, treated with DNase I or RNase A, and resolved on a denaturing polyacrylamide gel. nt: nucleotides; tRFs: tRNA fragments. **B-C)** Length distribution of small RNA (**B**) and DNA (**C**) co-purified with MapSPARTA as determined by sequencing. **D)** Percentages of long or small RNA reads that align to specific genomic or plasmid

5.2.7. SPARTA induces cell death in the presence of plasmid DNA

Based on its association with plasmid-targeting small RNA-guides *in vivo*, we speculated that SPARTA could detect invading plasmid DNA and subsequently deplete NAD⁺ and NAD(P)⁺. NAD⁺ and NADP⁺ are central in various metabolic pathways and their depletion might induce lower fitness, cell dormancy, or programmed cell death, as observed in other eukaryotic and prokaryotic immune systems ^{29–32}.

MapSPARTA was selected for *in vivo* characterization in *E. coli* as, in contrast to CrtSPARTA, it shows robust activity at 37 °C *in vitro* (**Sup. Fig. S3B**). To mimic genomic expression levels MapSPARTA was expressed from a bacterial artificial chromosome (BAC) with a copy number of 1 ³⁷. NAD (NAD⁺ + NADH) and NADP (NADP⁺ + NADPH) levels, cell density (OD_{600 nm}) and viability (defined here as the number of colony forming units (CFU)) were determined at 4h and/or 20h after induction of MapSPARTA expression. After 20h, MapSPARTA expression resulted in a modest reduction of NAD (-30%, p<.01) and an increase of NADP (+18%, p<.001) compared to MapSPARTA^{TIR-E77A} expression, but no change in OD_{600 nm} or CFU was observed (S5A-D). This implies that MapSPARTA expression does not lower cell viability under these conditions. In contrast, in presence of the high copy number plasmid pUC-empty (500-700 copies per cell ³⁸) expression of BAC-encoded MapSPARTA resulted in a strong reduction of NAD (4h: -85%, 20h: -89%, p<0.001), NADP (4h: -79%, 20h: -91%, p<0.001), and cell viability (4h: -27%, n.s., 20h: -49%, p<0.01) (**Fig. 5A** and **Sup. Fig. S5A**). This shows that presence of a high copy number plasmid triggers SPARTA-mediated NAD(P) degradation and induces cell death.

[Fig. 4]... sequences. RNA was extracted from *E. coli* expressing MapSPARTA (*E. coli* total) or from purified MapSPARTA (MapSPARTA-associated). rRNA-derived reads are excluded. **E**) Small RNA MapSPARTA-associated have a bias for 5' adenosine bases. Nucleotide bias of small RNAs MapSPARTA-associated (left) or *E. coli* total (right). **F**) Correlation between MapSPARTA-associated small RNA and *E. coli*-extracted long RNA sequences. **G**) Correlation between MapSPARTA-associated small RNA and DNA sequences. See also **Sup. Fig. S4** and **Sup. Table 1**.



5.2.8. High transcription of plasmid DNA enhances SPARTA activity

As SPARTA-associated guide RNA abundance is correlated with transcript abundance (**Fig. 4F**), we speculated that high transcription of plasmid-encoded genes might trigger a stronger SPARTA-mediated response. To test this, BAC-encoded MapSPARTA was expressed in presence of one of three pUC-derivates encoding monomeric Red Fluorescent Protein (mRFP), and viability was determined 4h after induction of MapSPARTA expression (**Fig. 5B**). Compared to cells expressing MapSPARTA without plasmid, no significant loss of viability is observed in presence of pUC when *mRFP* lacks a promoter and ribosomal binding site (pUC-mRFP^{ΔT7-ARBS}). In contrast, viability is drastically lowered when mRFP transcription is under control of a T7 promoter (pUC-mRFP; -98%, $p < 0.001$), even when an RBS is absent (pUC-mRFP^{ARBS}; -92%, $p < 0.001$). Also when mRFP transcription is under control of the XylS/Pm promoter³⁹ which relies on native RNA polymerases for transcription, cell viability is reduced compared to cells in which mRFP is not transcribed (pUC-mRFP^{XylS/pm}, -59%, $p < 0.001$; **Sup. Fig. S5B**). This excludes the possibility that MapSPARTA activation relies specifically on transcription by phage-derived T7 RNA polymerase. In conclusion, high transcription of plasmid-encoded genes results in stronger SPARTA-induced cell death.

5

[Fig. 5]...expressing MapSPARTA have point mutations or insertions of transposable elements (IS2) in the gene encoding T7 RNA polymerase. **D**) SPARTA activation relies on plasmid copy number and transcription. NAD levels and cell viability were determined in *E. coli* cultures containing a BAC expressing MapSPARTA in the absence or presence plasmids with different origin of replications (ori) from which mRFP was transcribed (mRFP^{ARBS}) or encoded but not transcribed (mRFP^{ΔT7-ARBS}). **E**) Transcription of RNA complementary to plasmid DNA enhances SPARTA-mediated cell death. NAD levels and cell viability were determined in *E. coli* cultures containing a BAC expressing MapSPARTA or MapSPARTA and mRFP, in absence or presence of plasmids lacking or containing or a non-transcribed *mRFP* gene. Graphs show the average of three (**A-D**) or two (**E**) biological replicates. Error bars indicate standard deviation. N.s. = not significant; * = $p < 0.05$, ** = $p < 0.01$; *** = $p < 0.001$. See also **Sup. Fig. S5**

Overnight cultures of mRFP-expressing strains show bright mRFP fluorescence in presence of MapSPARTA^{TIR-E77A}, but not in presence of MapSPARTA (-80%, $p < 0.01$; **Fig. 5C**). Plating and subsequent sequencing revealed that in all viable MapSPARTA-expressing clones the gene encoding T7 RNA polymerase was disrupted either by a point mutation or by insertion of an IS2 transposable element ⁴⁰ (**Fig. 5C**). As loss of the T7 RNA polymerase results in loss of *mRFP* transcription, these ‘escape mutants’ elude SPARTA-mediated cell death.

5.2.9. SPARTA activation relies on plasmid copy number and transcription

SPARTA associates with genome-targeting guide RNAs (**Fig. 4**), but does not lower cell viability in absence of plasmids (**Fig. 5A**). We therefore hypothesized that the copy number of the target DNA plays a role in formation of activated SPARTA complexes. To test this, MapSPARTA was expressed in presence of a plasmid with one of three different origin of replications (ori) resulting in a low (pSC101 ori; ~5 copies), medium (pET ori, 15-20 copies), or high (pUC ori, 500-700 copies) copy number ^{38,41}. Each of these plasmids encodes an *mRFP* gene without promoter (no transcription; mRFP^{AT7-ARBS}) or with T7 promoter (transcription; mRFP^{ARBS}).

NAD levels and cell viability were determined respectively 4h and 20h after induction of MapSPARTA expression, but are not significantly reduced in cultures with the low copy plasmid, even when *mRFP* is transcribed (**Fig. 5D**). This could explain why MapSPARTA does not lower cell viability in absence of a plasmid DNA despite acquiring guides that target the single-copy genome of *E. coli*. Presence of the medium copy plasmid results in lowered NAD (-96%, $p < 0.001$) and cell viability (-93%, $p < 0.001$) only when *mRFP* is transcribed. Presence of the high copy number plasmid without *mRFP* transcription lowers both NAD levels (84%, $p < 0.001$) and cell viability (49%, $p < 0.001$), and *mRFP* transcription further reduces cell viability (-96%, $p < 0.01$). These results imply

that both plasmid copy number and transcription of plasmid-encoded genes play a role in SPARTA activation.

However, higher plasmid copy numbers also result in more transcripts of plasmid-encoded genes. To determine if SPARTA is activated by increased availability of target DNA independent of higher transcript abundance, T7 promoter-controlled *mRFP* was placed on the MapSPARTA expression BAC (copy number = 1). Transcription of *mRFP* from the BAC does not result in reduction of NAD levels or viability in absence of plasmid DNA or in presence of a medium copy pET plasmid, even if it encodes a non-transcribed *mRFP* gene (**Fig. 5E**). While the presence of a high copy pUC plasmid itself results in lower NAD levels (-69%, $p < 0.001$) and cell viability (-24%, $p < 0.05$), inclusion of a non-transcribed *mRFP* gene on that plasmid further reduces NAD levels (-88% $p < 0.001$) and cell viability (-55%, $p < 0.001$). This implies that BAC-encoded transcripts facilitate SPARTA activation in presence of high copy number DNA containing a sequence complementary to that transcript. We conclude that both high transcription levels *and* plasmid DNA abundance stimulate the formation of a higher number of activated SPARTA complexes, resulting in lower NAD levels and consequential cell death.

5.2.10. SPARTA confers selective pressure against plasmid DNA

Given that SPARTA induces cell death in the presence of high copy plasmids, we investigated whether SPARTA interferes with plasmid transformation or propagation. MapSPARTA does not significantly reduce transformation efficiency of pUC-empty or pUC-mRFP^{XylS/pm} (**Sup. Fig. S6A**). Possibly SPARTA is not activated as the plasmid copy number directly after transformation is presumably low. Furthermore, SPARTA does not induce plasmid loss in *E. coli* cells harboring pUC-empty (**Sup. Fig. S6B**). This suggests that SPARTA is unable to protect individual cells against invading DNA: it cannot prevent it from entering or remove it from individual cells after entry.

We hypothesized that SPARTA provides immunity on the population-level instead. To test this, 'mixed cultures' consisting of plasmid-free (4%) and

pUC-empty-containing *E. coli* cells (96%) were cultivated (**Fig. 6A**). In presence of an empty BAC or a BAC expressing MapSPARTA^{TIR-E77A}, the percentage of pUC-empty-containing cells remains stable around 96% (**Fig. 6B**). In contrast, cultures expressing MapSPARTA are dominated by plasmid-free cells after three days (**Fig. 6B**, $p < 0.001$). This implies that SPARTA provides a selective advantage for plasmid-free cells by driving the plasmid-invaded cells to extinction.

5.2.11. SPARTA does not provide robust defense against selected bacteriophages

Infection by lytic bacteriophages (from here: phages) leads to rapid replication (and thus high copy number) and high transcription of phage genomes⁴². However, no defence is observed for *E. coli* strains expressing MapSPARTA or MapSPARTA^{TIR-E77A} in plaque assays with phage T1, T4, T7, Nami, and Lambda-vir (**Sup. Fig. S6C**). Also no protection is observed in culture collapse studies with phage T4 and T7 (**Sup. Fig. S6D**). Culture collapse studies with phage T1 and Nami show a small but significant ($p < 0.005$) reduction in collapse for strains expressing MapSPARTA or MapSPARTA^{TIR-E77A}, and for phage Lambda-vir in the strain expressing MapSPARTA ($p < 0.03$; **Sup. Fig. S6D**). Furthermore, both MapSPARTA and MapSPARTA^{TIR-E77A} show a small but significant ($p < 0.001$) reduction in release of infectious virions from the chronic-infecting phage M13 (**Sup. Fig. S6E**). In all cases the observed defence is limited compared to known anti-phage systems^{43,44}. The observation that the SPARTA-mediated phage defence does not rely on NAD(P)ase activity in most cases suggests that MapSPARTA is not properly activated in response to the phages used.

5.2.12. SPARTA can be repurposed for sequence-specific nucleic acid detection

To determine if SPARTA could be programmed to detect nucleic acids akin to various CRISPR-Cas effector enzymes^{45–49}, the specificity and sensitivity of CrtSPARTA and MapSPARTA were determined. CrtSPARTA activity is attenuated by single mismatches between the guide RNA and target ssDNA at

guide positions 5-10, and is further lowered by double mismatches (**Fig. 7A**). Double mismatches at middle-region positions 5-6 to 9-10 and 14-15 result in low to no detectable CrtSPARTA activity, while mismatches toward the 5' end or 3' end of the guide are well-tolerated. MapSPARTA is less sensitive to mismatches under the tested conditions, but also for MapSPARTA double mismatches in the middle region lower activity most (**Fig. 7B**). Both CrtSPARTA and MapSPARTA detect target ssDNA in the low nM range (**Fig. 7C-D**).

Although sensitivity of SPARTA is comparable to the detection limit of Cas12a and Cas12b^{45,48}, it is not sensitive enough to achieve clinically relevant attomolar detection thresholds^{50–53}. CRISPR-based detection tools are often combined with DNA amplification techniques to improve their detection limit. As proof-of-concept, we enhanced SPARTA sensitivity and enabled dsDNA detection based on a previously developed method⁵⁴ (**Fig. 7E**): To-be-detected DNA was PCR amplified using forward primers with a 5'-phosphorothioate (PT) group. Subsequent or simultaneous incubation with T7 exonuclease removes the reverse strand lacking 5'-PT groups, facilitating MapSPARTA-mediated dsDNA detection at aM sensitivity (**Fig. 7F** and **Sup. Fig. S7**). This shows that SPARTA can be combined with nucleic acid amplification techniques and can be used to detect both ssDNA and dsDNA sequences in a sequence-specific manner.

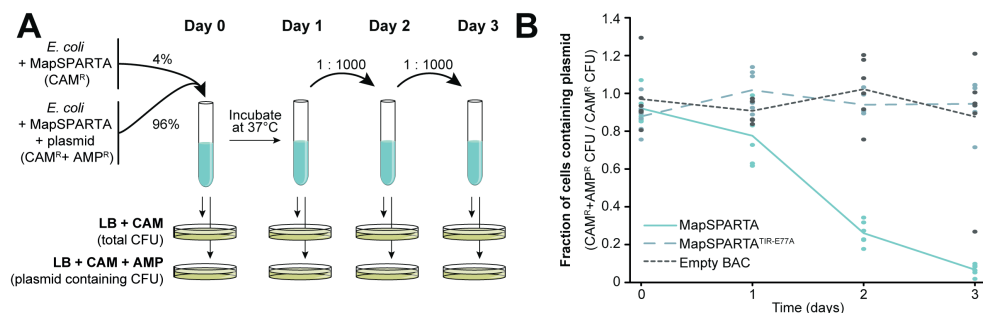


Figure 6. SPARTA depletes plasmid-harboring cells from bacterial cultures. **A)** Graphical schematic representation of the mixed-culture competition experiment (**B**). **B)** MapSPARTA-expressing *E. coli* cells that contain pUC-empty or that were plasmid free were mixed in a 24:1 ratio, cultivated in LB, and the percentage of plasmid-containing cells was determined daily. *E. coli* with an empty BAC or a BAC expressing MapSPARTA^{TIR-E77A} were used as control. Graph shows the average of six biological replicates. See also **Sup. Fig. S6**.

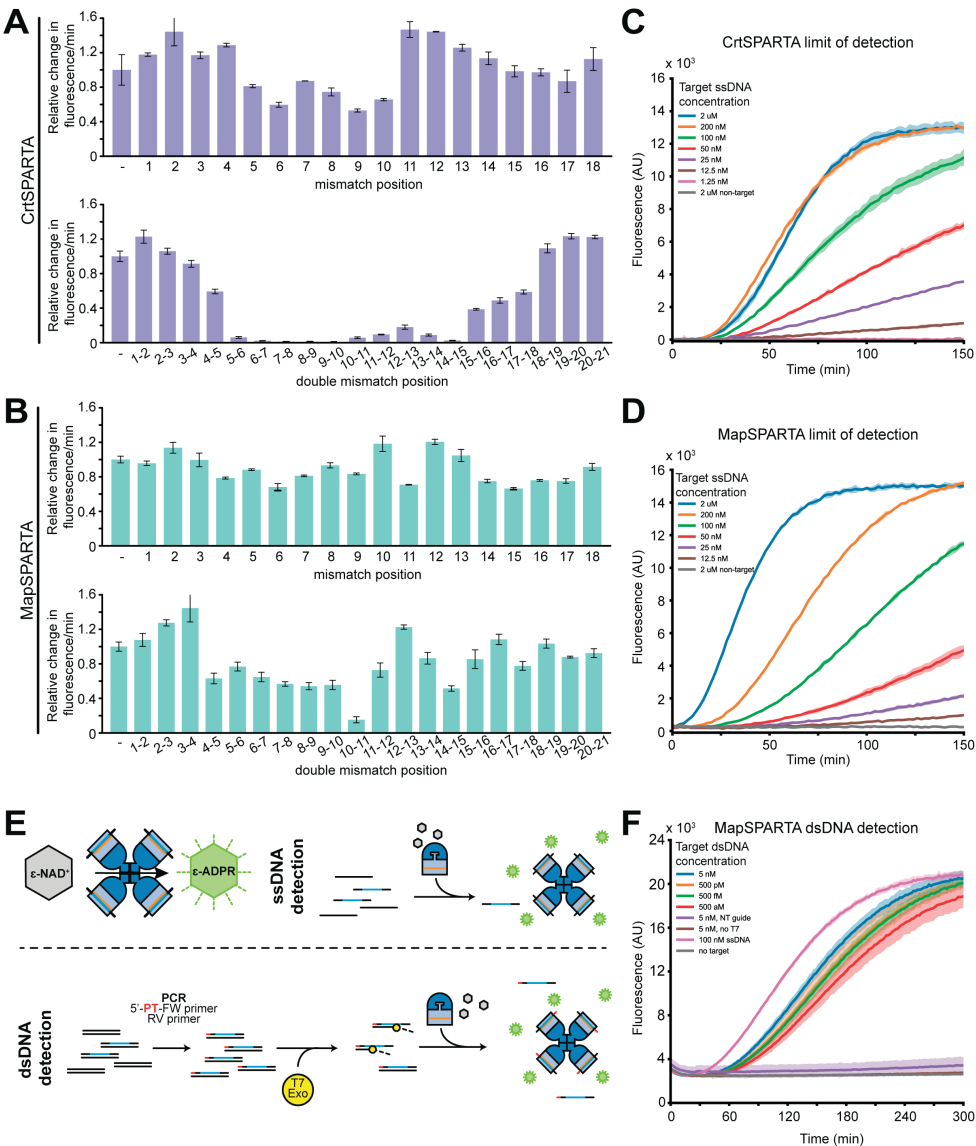


Figure 7. SPARTA systems can be used for nucleic acid detection. A-B) Effect of mismatches on SPARTA activation. CrtSPARTA (A) or MapSPARTA (B) were mixed with guide RNA and target ssDNA containing single or double mismatches and incubated with ϵ -NAD⁺. Graphs indicate the change in fluorescence over time relative to a control target ssDNA without mismatches. **C-D)** SPARTA detects target ssDNA at nM concentrations. CrtSPARTA (C) or MapSPARTA (D) were mixed with guide RNA in a 1:1 molar ratio and target ssDNA was added at different concentrations. After addition of ϵ -NAD⁺, total fluorescence was measured over time (corrected

5.3. Discussion

In this study we describe the mechanism and function of SPARTA, a prokaryotic immune system that belongs to the short pAgo family. Short pAgo and its associated TIR-APAZ protein form a catalytically inactive heterodimeric SPARTA complex. Guide RNA-mediated binding of a target ssDNA facilitates formation of an oligomeric SPARTA-NA complex with a 4:4:4:4 stoichiometry (TIR-APAZ:short pAgo:guide:target) with NAD(P)ase activity (**Fig. 3**). Further structure-function studies are required to unravel the molecular mechanism for guide-mediated target binding and subsequent activation of SPARTA.

TIR domain activation by oligomerization has also been observed in the prokaryotic immune systems Pycsar and STING^{29,55} and in plant receptor proteins^{56,57}, and other prokaryotic immune systems also rely on TIR-mediated NAD⁺ depletion^{29,31,58}. Some of these systems are activated by small molecules^{29,31,55}, but SPARTA is not activated by NAD⁺, NADP⁺, NAM, ADPR(P), c-ADPR, or other cyclic NMPs, and these small molecules do not enhance guide RNA/target ssDNA-mediated SPARTA activation (**Sup. Fig. S7C**). While different immune systems might rely on distinct mechanisms for invader recognition, TIR-mediated NAD(P)⁺ depletion is an immune strategy found in both prokaryotes and eukaryotes^{28,32,59–61}.

In vivo, SPARTA acquires 5'-P guide RNAs from the transcriptome (**Fig. 4**). Guide RNAs and mRNA abundance are correlated, implying that SPARTA does not specifically recruit invading DNA-targeting guides. Similar observations

[Fig. 7]... with no target control). **E**) Schematic representation of SPARTA-mediated DNA detection. Activated SPARTA converts ϵ -NAD⁺ to fluorescent ϵ -ADPR. DNA is PCR amplified using a phosphorothioate (PT) modified forward (FW) primer and a non-modified reverse (RV) primer. PCR products are treated with T7 exonuclease (T7 exo) yielding ssDNA products detectable by SPARTA. **F**) SPARTA combined with PCR and T7 exonuclease facilitates detection of dsDNA at aM levels. MapSPARTA was mixed with guide RNA in a 1:1 molar ratio and PCR-amplified T7 exonuclease-digested target DNA was added. After addition of ϵ -NAD⁺, total fluorescence was measured over time. All graphs show the average of three technical replicates. Error bars (**A & B**) or shadings (**C, D & F**) indicate standard deviations. See also **Sup. Fig. S7**.

were made for RsAgo, a catalytically inactive long pAgo ⁴. This suggests RNA-guided pAgos from different phylogenetic clades rely on similar mechanisms for guide acquisition, which remain to be determined.

Multi-copy plasmid DNA trigger SPARTA-mediated NAD(P)⁺ depletion and consequential cell death (**Fig. 5**), while single or low copy number plasmids do not trigger significant levels of SPARTA-mediated NAD⁺ degradation. This explains why SPARTA does not cause cell death when associating with genome-targeting guide RNAs. High transcription of plasmid-encoded genes or of sequences homologous to plasmid DNA enhance SPARTA-mediated NAD⁺ degradation and consequential cell death, presumably because SPARTA can acquire more plasmid-targeting guides from the transcriptome. In theory this allows for genomic transcription of RNA to program SPARTA to target complementary multi-copy invader DNA. It remains to be determined whether enhanced SPARTA activation by multi-copy target DNA is a consequence of higher target sequence abundance only, or if also their replication rate plays a role, as this might increase accessibility of the target DNA.

Highly transcribed multi-copy elements are not only a metabolic burden to the cell, but may also indicate plasmid or viral replication and impending spread to neighbouring cells. Unless high transcription is prevented through mutational escape (which might neutralize the invader), SPARTA lowers cell viability. When SPARTA triggers cell death, this could prevent spread of the invader to protect the remainder of the bacterial population. While this does not protect the infected cell itself, it provides population-level immunity akin to abortive infection systems ⁶². Reduced cell viability could also induce a state of dormancy in which also the invader cannot replicate efficiently ^{63,64}. Bacteriostasis does not only allow other immune systems to remove the invader ⁶⁵, but also lets uninvaded bacterial cells outcompete infected cells (**Fig. 6**).

SPARTA provides minor protection against phages which, strikingly, does not rely on its NADase activity (**Sup. Fig. S6**). Other catalytically inactive pAgos also interfere with invading DNA ^{4,6}. This suggests that pAgos do not critically rely on catalytic activity for DNA interference, but might provide a

minimal level of defence by binding invading DNA which could hamper replication and/or transcription. In nature, many phage replication cycles are (partially) lysogenic and/or phages (temporarily) exist extrachromosomal, akin to plasmids ⁶⁶. Such infection strategies might provide a larger window of opportunity for SPARTA to become activated and remove the invader DNA from the population.

We hypothesize that short pAgos from other phylogenetic clades also control the activity of their associated APAZ-fused catalytic domains. S1A and S1B SPARSA (Short prokaryotic Argonaute SIR2-APAZ) systems contain a SIR2 domain which is a predicted NADase, akin to SPARTA TIR domains. Expression of SPARSA systems in *E. coli* also triggers NAD depletion (**Sup. Fig. S7D**), suggesting a similar mode of action. Certain clade S2B APAZ proteins are fused to a DHS-like domain which is a predicted NAD/FAD-binding domain (interpro IPR029035) ⁶⁷, and various long pAgos are encoded in operons with SIR2-like proteins ^{17,19}. This suggests that NADase activity is not restricted to SPARTA and SPARSA systems, but might have independently evolved several times in both short and long pAgo systems.

Finally, we provide proof-of-concept for SPARTA-mediated detection of DNA sequences. Upon DNA detection, SPARTA NADase activity can be monitored using fluorescence or by colorimetric cycling assays. Its limit of detection is comparable to that of CRISPR-Cas-based detection tools, but SPARTA makes use of shorter RNA guides and does not require the presence of a protospacer adjacent motif in the target sequence. Combined with pre-amplification techniques, DNA can be detected at aM detection levels. We predict that DNA can also be amplified using isothermal amplification techniques, and that RNA could be detected using reverse transcriptase-based amplification techniques. As such, SPARTA-based tools can potentially be used for the detection of pathogens or specific mutations in DNA. In conclusion, our study elucidates the biological role and mechanism of SPARTA systems and facilitates its implementation as a nucleic acid detection tool.

5.3.1. Limitations of study

While SPARTA is preferentially activated with 5'-P guide RNAs and target ssDNA *in vitro* and associates with 5'-P guide RNAs *in vivo*, it cannot be ruled out that guide ssDNA/target RNA combinations also play a role in activation of SPARTA.

Compared to other bacterial immune systems, SPARTA systems provide relatively little (and TIR-independent) protection against some of the phages in our panel. However, many prokaryotic immune systems are effective against certain types of phages only ^{31,43}. Furthermore, the laboratory experiments performed are optimal for phage propagation and might be inappropriate to observe SPARTA-mediated immunity.

SPARTA-mediated immunity might further benefit from host-specific factors for guide generation and invader targeting. Furthermore, SPARTA-generated ADPR(P) and/or NAM might function as secondary messenger molecules to trigger auxiliary host-specific proteins.

5.4. Material and methods

5.4.1. Bacterial strains and cultivation method

For cloning purposes, NEB 5-alpha *E. coli* (New England Biolabs) was used.

For protein expression *E. coli* BL21 Star (DE3) (ThermoFisher Scientific) transformed with plasmids encoding the protein(s) of interest was used.

For *in vivo* investigation of the role of SPARTA systems *E. coli* BW25113 (Coli Genetics Stock Center) transformed with bacterial artificial chromosomes (BACs) encoding the protein(s) of interest, and pUC-empty when applicable, were used. For *in vivo* experiments that required T7 RNA polymerase-mediated expression of RFP from pUC-mRFP or derivatives thereof, *E. coli* BW25113(DE3) was used (kind gift from Prof. dr. Ruud Weusthuis (Wageningen University)).

For phage M13 propagation and M13 plaque experiments *E. coli* ER2738 and *E. coli* K12 JM109(DE3) were used, respectively.

All *E. coli* strains were cultivated in LB Broth Low Salt (Duchefa Biochemie) or on LB agar plates (LB Broth Low Salt (Duchefa Biochemie) supplemented with 1% agar) at 37 °C unless stated otherwise. Antibiotics corresponding to the plasmids used in each experiment were added to the following concentrations: ampicillin (AMP): 100 µg/mL; kanamycin (KAN): 34 µg/mL, chloramphenicol (CAM): 25 µg/mL. When required, expression was induced by adding IPTG (for protein expression and purification and for expression of mRFP during *in vivo* assays), *m*-toluic acid (for expression of mRFP during *in vivo* assays) or arabinose (for *in vivo* assays) to final concentrations of 0.25 mM, 1 mM and 0.2% w/v, respectively.

5.4.2. Plasmid construction

Genes encoding short pAgo systems or individual proteins thereof were amplified from genomic DNA obtained from DSMZ (*Crenatolea thermophila*: DSM 14807; *Maribacter polysiphoniae* DSM 23514; *Joostella marina*: DSM 19592, *Xanthomonas vesicatoria*: DSM 22252) or from synthetic DNA constructs

ordered from Twist Bioscience. The genes were cloned in the pET-His6-MBP-TEV-LIC cloning vector (1M), which was a kind gift of Prof. Dr. Martin Jinek (University of Zurich) (Addgene #29656). In this plasmid the genes are placed under control of a T7 promoter which allowed for expression of the proteins that are linked to an N-terminal 6x-His-Maltose Binding Protein (MBP) tag by a Tobacco Etch Virus (TEV) cleavage site-containing linker.

For *in vivo* experiments genes were cloned in the Bacterial Artificial Chromosome (BAC) pBeloBAC11 (New England Biolabs) under control of an araBAD promoter. For M13 experiments, the low-copy number vector pBbS5k-RFP-based vector pBK133 was used. For combined expression of short Ago and TIR-APAZ proteins, an artificial operon was constructed by inserting a 45 nt spacer containing an additional ribosomal binding site (RBS) between the two genes. All cloning procedures, including minor modifications of RBSs and making point mutations, were performed using the NEBuilder HiFi DNA Assembly kit (New England Biolabs), subsequent DNA cleanup (Zymo) and electroporation or heat-shock transformation into NEB 5-alpha *E. coli* cells (New England Biolabs).

Plasmids and primers for cloning used in this study are given in **Sup. Table 2**.

5.4.3. Protein purification

Individual SPARTA proteins and SPARTA complexes with N-terminal 6x-his-MBP-tags were heterologously expressed in *E. coli* BL21 Star (DE3) and purified by a combination of Ni²⁺-affinity, Amylose-affinity, heparin cation exchange, and size exclusion chromatography. Plasmids used for protein expression were transformed into *E. coli* BL21 Star (DE3). A single colony was used to inoculate a 10 mL LB culture which was incubated o/n at 37 °C at 180 RPM. The o/n culture was used to inoculate 8x750 mL LB medium which was incubated at 37 °C at 150 RPM. When an OD_{600 nm} of ~0.25 was reached, the temperature of the incubator was switched to 18 °C and after 1 h (when OD_{600 nm} reached ~0.6), protein expression was induced by addition of IPTG. Protein expression took place at 18 °C for 16 hours. Cells were harvested by

centrifugation at 5000 x g at 4 °C for 15 min. Cells were resuspended in NiNTA lysis buffer (500 mM NaCl, 20 mM Imidazole, protease inhibitors (200 µg/mL AEBSF, 1 µg/mL pepstatin), 20 mM HEPES pH 7.5) and stored at -80 °C until used for protein purification. After thawing, cells were lysed by sonication (QSONICA Q700A-220 sonicator with ½" tip; amp 50%, 1s ON/2s OFF, 5 minutes total ON time) and the lysate was centrifuged for 30 min at 30,000 x g at 4 °C. The clarified lysate was applied to a 5 mL HisTrap HP column (Cytiva Life Sciences). The column was washed with ~6 column volumes (CV) NiNTA buffer A (500 mM NaCl, 5 mM Imidazole, 20 mM Tris pH 8) or until $A_{280\text{ nm}}$ of the eluate was below 250 a.u. The protein was eluted in NiNTA buffer B (500 mM NaCl, 350 mM Imidazole, 20 mM Tris pH 8). Fractions containing the protein of interest were pooled, and loaded on a 20 mL Amylose resin (New England Biolabs) column. The column was washed with ~3x CV Amylose wash buffer (500 mM NaCl, 1 mM DTT, 20 mM Tris pH 8) or until the $A_{280\text{ nm}}$ was ~0 a.u.. The protein was eluted in Amylose elution buffer (500 mM NaCl, 1 mM DTT, 10 mM maltose, 20 mM Tris pH 8). Fractions containing the protein of interest were pooled, EDTA was added to a final concentration of 2 mM, and per TEV protease was added in a 1:50 (w/w) ratio. The sample was dialyzed in SnakeSkin dialysis tubing (10kDa MWCO, Thermo Scientific) against dialysis buffer (250 mM KCl, 1 mM DTT, 2 mM EDTA, 20 mM HEPES pH 7.5) for 16 h. Cleavage of the MBP tag by TEV was confirmed by SDS-PAGE and Coomassie Brilliant Blue (CBB) staining. The sample was diluted 1:1 with 20 mM HEPES pH 7.5 and loaded on a HiTrap Heparin column (Cytiva Life Sciences). The column was washed with 98.7% ion-exchange buffer I (100 mM KCl, 1 mM DTT, 8 mM HEPES pH 7.5) and 1.3% ion-exchange buffer II (2.5 M KCl, 1 mM DTT, 20 mM HEPES pH 7.5) until the $A_{280\text{ nm}}$ was ~0. To protein was eluted with ion-exchange buffer II by applying a gradient from 1.3% to 50% over a total volume of 60 mL. Peak fractions were analyzed by SDS-PAGE and CBB staining and fractions containing the protein of interest were combined. The sample was concentrated to ~1 mL using centrifugal filter units (Amicon) according to the protocol of the manufacturer. The sample was centrifuged for 10 min at 16,000 x g at 4°C and the protein was

fractionated on a custom 200 mL Superdex 200 Prep Grade resin (Cytiva Life Sciences) column, which was equilibrated in and eluted with SEC buffer (500 mM KCl, 1 mM DTT, 20 mM HEPES pH 7.5). Peak fractions were analyzed by SDS-PAGE and CBB staining and fractions containing the protein of interest were combined. The sample was concentrated to the desired concentration using centrifugal filter units (Amicon) according to the protocol of the manufacturer. Samples were aliquoted, flash frozen in liquid nitrogen, and stored at -80 °C until further use.

5.4.4. Nucleic acid extraction

CrtSPARTA and MapSPARTA were purified as described above, but expression cultures (750 mL for CrtSPARTA, 8x750mL for MapSPARTA) were incubated at 37 °C throughout the procedure, and cells were harvested 3 h after induction of expression. SPARTA complexes were purified using the NiNTA and Amylose resin purification steps only. All buffer solutions used were as described above but instead contained 125 mM NaCl and they were supplemented with 2 mM MnCl₂ and 2 mM MgCl₂ each. After elution from the Amylose resin column and pooling, the SPARTA complex was concentrated to 10 μM using centrifugal filter units (Amicon) according to the protocol of the manufacturer. Samples were flash frozen in liquid nitrogen and stored at -80 °C until further use.

Further centrifugation steps were performed in a tabletop centrifuged cooled to 4 °C, at 16000 x g. To extract nucleic acids co-purified with the complexes, 360 μL protein solution was supplemented with EDTA (10 mM final concentration) and 40 uL proteinase K 20 mg/ml (New England Biolabs). The sample was incubated for two hours at RT. 400 μL of Roti phenol/chloroform/isoamyl alcohol was added to the mixture, followed by brief vortexing and centrifugation for 15 minutes. The upper aqueous phase was isolated and 0.1 volumes of 1M sodium acetate, 3 volumes 100 % ethanol and 10 μL linear polyacrylamide (Alfa Aesar) were added. This mixture was vortexed briefly and incubated at -20 °C for 16 hours. Samples were centrifuged for 30 min and the supernatant was removed from the pellet. The pellet was washed with

500 μ L 70% ethanol (pre-cooled to -20°C). The pellets containing the copurified nucleic acids were dried for 10 min at 50°C , and pellets were resuspended in 50 μ L RNase free H_2O .

Purified nucleic acids were [γ - ^{32}P]-ATP labelled with T4 polynucleotide kinase (PNK; Thermo Scientific) in exchange- or forward-labelling reactions. After stopping the reaction by incubation at 75°C for 10 min, the labelled oligonucleotides were separated from free [γ - ^{32}P] ATP using a Sephadex G-25 column (GE Healthcare). Labelled nucleic acids were incubated with nucleases (RNase A, DNase and protease-free (Thermo Scientific), or DNase I, RNase-free (Thermo Scientific) for 30 min at 37°C . After nuclease treatment, samples were mixed with Loading Buffer (95% (deionized) formamide, 5 mM EDTA, 0.025% SDS, 0.025% bromophenol blue and 0.025% xylene cyanol), heated for 5 min at 95°C and resolved on 15% denaturing (7M Urea) polyacrylamide gels. Radioactivity was captured from gels using phosphor screens and imaged using a Typhoon FLA 7000 laser-scanner, GE Healthcare).

5.4.5. RNA and DNA sequencing and analysis

RNA was sequenced by GenomeScan (Leiden, The Netherlands). Total RNA was isolated from *E. coli* after 3h of MapSPARTA or CrtSPARTA expression at 37°C , or was extracted from purified MapSPARTA or CrtSPARTA complexes as described above. All small RNA and long RNA libraries were generated by GenomeScan. DNA was extracted from purified MapSPARTA complexes as described above, and pre-treated with RNase A, DNase and protease free (Thermo Scientific). The remaining DNA was purified using the MagMAX mirVana Total RNA Isolation Kit (Thermo Fisher), using the protocol of the manufacturer to isolate RNA from serum and plasma samples but excluding both DNase treatment steps. Subsequently the samples were treated with T4 PNK (Thermo Fisher) according to the protocol of the manufacturer, and the DNA was purified again using the MagMAX mirVana Total RNA Isolation Kit. A library of each sample was generated using the NEB Next small RNA library prep set for illumina (NEB) according to the protocol of the manufacturer, but using a DNA

oligo with same sequence instead of the 5' SR adapter. The reverse transcriptase treatment was omitted and during the PCR amplification treatment samples were first incubated for 15 minutes at 70°C without addition of SR and index primers to facilitate to facilitate dsDNA formation.

All libraries were sequenced by GenomeScan using Illumina NovaSeq6000 sequencing with paired-end reads and 150bp read length. Paired-end small RNA and DNA reads were merged and adapter sequences were trimmed using BBtools ⁶⁸. Processed reads of all sequencing libraries were aligned to the genome of *E. coli* BL21 (GenBank: CP053602.1) and to the expression plasmid (pBK086 or pBK094) using HISAT2 ⁶⁹. Length and sequence distribution of smallz RNAs were analysed using FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc) using a library with determined minimum and maximum lengths of respectively 15 and 60 nucleotides. FeatureCounts ⁷⁰ was used to assign reads to genomic features.

5

5.4.6. Size Exclusions Chromatography – Multi-Angle Light Scattering

Purified MapSPARTA complex was diluted to a 1 mg/ml concentration incubated for 15 min at 37°C either without nucleic acids, or in 1:1 molar ratio with a 5'-phosphorylated guide (oBK084) in SEC-MALS buffer (125 mM KCl, 2 mM MgCl₂, 20 mM HEPES pH7.5). Subsequently, target DNA (oBK382) was added in a 1:1 molar ratio and samples were incubated for 30 min at 37°C. CrtSPARTA was diluted to a 1 mg/ml concentration in SEC-MALS buffer without nucleic acids and incubated for 30 min at 55°C. After incubation samples were resolved at RT on a Superdex 200 Increase 10/300 GL column (Cytiva Life Sciences) connected to a 1260 Infinity II HPLC system (Agilent) using SEC-MALS buffer. The eluate was subjected to Multi-Angle Light Scattering using an Optilab 1090 Differential Refractive Index detector (Wyatt Technology) and a miniDawn 1065 Multi-Angle Light Scattering system (Wyatt Technology). The data were analyzed using Astra 8.0. The conjugate analysis method in ASTRA was used to determine the contribution of protein and nucleic acid in the eluting

species, using dn/dc values of 0.185 mL/g and 0.17 mL/g for protein and nucleic acid, respectively, and using theoretical extinction coefficients at 280 nm of 1.5655 mL/(mg*cm) and 10.766 mL/(mg*cm) for MapSPARTA and RNA/DNA duplex, respectively. Subsequently, fractions containing guide/target-complexed MapSPARTA were incubated with 25 μ M ϵ -NAD⁺ to determine NADase activity, as described below.

5.4.7. *In vitro* nucleic acid cleavage assays

Duplicate or triplicate reactions were made of CrtSPARTA, CrtSPARTA^{TIR-E77A} or MapSPARTA mixed with various guide oligonucleotides and Cy5-labeled target oligonucleotides in reaction buffer on ice resulting in a total reaction volume of 15 μ L with the following final concentrations: 2 μ M CrtSPARTA, CrtSPARTA^{TIR-E77A} or MapSPARTA, 10 mM HEPES pH 7.5, 125 mM KCl, 2 mM MnCl₂ and 2 mM MgCl₂. Final guide and target concentrations were 2 μ M and 0.2 μ M in the same reaction mixture. First, the guide nucleic acid (ogDS01/ogDS02, **Sup. Table 2**) was added followed by 10 min incubation at 37 °C. Next, the target nucleic acid (oDS401/oDS403, **Sup. Table 2**) was added followed by 30 min incubation at 37 °C. 15 μ L 2X gel loading dye (250 mM EDTA, 5% v/v glycerol, 95% v/v formamide, Bromophenol Blue) was added. Samples were incubated at 95 °C for 3 min and nucleic acids were resolved on 20% denaturing (7M Urea) polyacrylamide gels. Gels were imaged on an Ettan DIGE Imager (GE Healthcare).

5.4.8. *In vitro* NADase assays

- **LC-MS**

Reaction mixtures with a volume of 50 μ L were prepared with the following final concentrations: 0.5 μ M CrtTIR-APAZ, CrtAgo or CrtSPARTA complex, 500 μ M NAD⁺, 10 mM HEPES pH 7.5 and 125 mM KCl. The mixture was incubated for 1h at 37 °C, after which they were transferred to ice. 150 μ L MilliQ H₂O was added to the mixture and the mixture was subsequently filtered using centrifugal filters (Pall, 3 kDa MWCO) according to the instructions

provided by the manufacturer. The flow through was collected and analysed by LC-MS using an G1311A HPLC pump (Agilent) with G1367B WP autosampler (Agilent) and an Alltima HP C18 3 μ m reverse phase column (Avantor) connected to a microTOF benchtop ESI-TOF MS (Bruker Daltonics).

- **NMR**

CrtTIR-APAZ activity was measured in a total reaction volume of 600 μ L with the following final concentrations: 0.5 μ M 6xHis-MBP-CrtTIR-APAZ was incubated with 800 μ M substrate (NAD⁺, NADH, or NADP) in 15 mM phosphate buffer pH 7.5, 125 mM KCl and 10% D₂O at 37 °C. The reaction with NADH was performed in an anaerobic bottle flushed with nitrogen gas to minimize spontaneous degradation of NADH. NMR measurements were performed on a standard-bore 600 MHz NMR spectrometer (Bruker BioSpin GmbH). 1D Nuclear Overhauser Effect Spectroscopy (NOESY) ⁷¹ experiments were conducted at RT. The ¹H-¹H mixing was 10 ms. A saturation of 72 Hz was applied on water during relaxation delay and mixing. NMR spectra were analysed using TopSpin 4.1.3 (Bruker BioSpin GmbH).

- **ϵ -NAD assays**

A reaction mixture of purified SPARTA complex in SEC buffer, ϵ -NAD⁺, RNA guide (oBK084, ogDS02

or a variant thereof, **Sup. Table 2**), and 5X reaction buffer (50 mM MES pH 6.5, 375 mM KCl, and 10 mM MgCl₂) was prepared on ice in 96-well or 384-well plates. The mixture was incubated at room temperature for 15 min, after which DNA target (oBK090, oDS400 or a variant thereof, **Sup. Table 2**) was added to a final concentration of 200 nM unless otherwise indicated. For experiments with double stranded ligands, ligands were prepared by mixing forward and reverse oligonucleotides in a 1.5:1 ratio and annealed by incubation at 95°C for 2 min and gradually cooled to room temperature. Final concentrations of each component were 1 μ M SPARTA complex, 25 μ M ϵ -NAD⁺ (Nicotinamide 1,N6-ethenoadenine dinucleotide), 10 mM MES pH 6.5, 125 mM KCl, and 2 mM MgCl₂ in a final volume of 60 μ L or 20 μ L. After addition of the target DNA, the plate was transferred to a preheated Synergy Neo2 or SH1M2FG plate reader (Biotek).

Kinetic measurements took place at 55 °C for CrtSPARTA and 37 °C for MapSPARTA unless specified otherwise. Fluorescence intensity was measured in kinetic mode using an excitation wavelength of 310 nm and emission wavelength of 410 nm. All experiments were performed in triplicates and error bars indicate standard deviations.

For pH-range experiments, reaction mixtures were prepared as described above, but instead of 10 mM MES pH 6.5, each sample contained a final concentration of 10 mM buffer mix (BTP, MES, and citric acid in a 1:1:1 molar ratio) with a pH in the range of 4.0 to 9.5. For optimum temperature determination, reaction mixtures were incubated at the indicated temperatures for 60 min, and end-point fluorescence intensity was measured using excitation of wavelength of 310 nm and emission wavelength of 410 nm using a Synergy Neo2 plate reader (Biotek). All experiments were performed in triplicates and error bars indicate standard deviations.

To determine the limit of detection for dsDNA targets, initial PCR amplification was followed by detection with MapSPARTA. 35-cycle PCRs were performed either on colonies transformed with the plasmid or on a purified plasmid in MilliQ H₂O. The concentration of the purified plasmid ranged from nM to aM and was determined using the Qubit™ dsDNA HS Assay Kit (ThermoFisher Scientific). Colonies were suspended in 20 µL MilliQ H₂O and incubated at 95 °C for 5 min. 50 µL PCR mixture contained 1 µL template, 25 µL Phusion High-Fidelity PCR Master Mix (ThermoFisher Scientific) and 0.5 µM of each primer (oAP030/oAP039, **Sup. Table 2**). The first five 5' nucleotides of oAP039 contain phosphorothioate groups to protect that strand from degradation by T7 exonuclease. After PCR cleanup (New England Biolabs) dsDNA products were either melted at 95 °C for 10 min and mixed with gRNA (oBK084, **Sup. Table 2**; 1:2) to form RNA:DNA hybrids or were used directly as targets in MapSPARTA detection assays containing T7 exonuclease. SPARTA reaction mixtures were prepared as described above. Reactions with T7 exonuclease were performed in the presence of 5U T7 exonuclease (New England Biolabs). All dsDNA detection reactions were incubated at 37 °C. Fluorescence intensity was

measured in kinetic mode every 2 minutes, using excitation wavelength of 310 nm and emission wavelength of 410 nm in a Synergy Neo2 or SH1M2FG plate reader (Biotek).

5.4.9. Effect of MapSPARTA on total NAD(P) levels and cell viability

Precultures of *E. coli* BW25113 or BW25113 containing a BAC encoding MapSPARTA (pBK116) or MapSPARTA^{TIR-E77A} (pBK124) or an empty BAC (pBeloBAC11), as well as a pUC21-based high copy number plasmid pUC-empty (pBK117) were grown o/n in LB supplemented with CAM and AMP in a shaker incubator at 37 °C. For experiments which includes multicopy plasmids encoding mRFP and derivatives thereof (pBK144-pBK146, pBK157, pBK159-pBK161, pBK164-pBK166), homologous regions between the expression BAC and pUC-mRFP were first removed from the BAC. In that case pBK138 and pBK139 were used for expression of respectively MapSPARTA and MapSPARTA^{TIR-E77A}. For experiments to determine the effect of increased plasmid encoded target concentration (pBK179-pBK184) without increasing transcription pBK175 and pBK176 were used for expression of MapSPARTA and mRFP, or MapSPARTA only, respectively. The precultures were diluted 1:100 in 5 mL fresh LB containing CAM and AMP and were incubated in a shaker incubator at 37 °C for 3 h. When applicable, the cultures were used as input for *in vivo* NAD/NADH quantification using the NAD/NADH cycling assay (see below). To normalize the amount of cells, the amount of cells corresponding to 2 mL of OD_{600 nm} = 1 was pelleted by centrifugation for 10 min at 3,500 x g. The supernatant was removed and the cell pellets were resuspended in 2 mL fresh LB supplemented with CAM, AMP, and arabinose. 10 µL of the normalized cultures was used to make a 1:2*10⁵ dilution in LB and 50 of this dilution µL was plated on LB agar plates supplemented with CAM and AMP. The remainder of the normalized culture was divided into two 13 mL tubes, which were incubated at 37 °C in a shaker incubator. When applicable mRFP expression was induced after one hour by adding IPTG or m-toluic acid. After 4h and 20h one of the two tubes was taken

for analysis. A 1×10^6 dilution was made and 60 μL was plated on LB agar plates supplemented with CAM and AMP. When applicable, the $\text{OD}_{600 \text{ nm}}$ was again measured and samples were taken as input for *in vivo* total NAD quantification using an NAD/NADH cycling assay (see below) and/or *in vivo* total NADP quantification using an NADP/NADPH cycling assay (see below). The remaining tube after 20h was also used to measure RFP fluorescence as described below. All experiments were performed as biological triplicates, and error bars indicate standard deviations.

5.4.10. Effect of SPARSA on total NAD levels

Precultures of *E. coli* BL21 Star(DE3) containing a plasmid encoding CrtTIR-APAZ (pBK026), XavSIR2-APAZ (pBK012), JomSIR2-APAZ-pAgo (pBK014) or an empty vector (pML-1M) were grown o/n in LB supplemented with KAN in a shaker incubator at 37 °C. The precultures were diluted 1:100 in 5 mL fresh LB containing KAN and were incubated in a shaker incubator at 37 °C. When an $\text{OD}_{600 \text{ nm}}$ of ~ 0.4 was reached, protein expression was induced by addition of IPTG. The cultures were grown for 1.5 hours and used as input for *in vivo* NAD/NADH quantification using the NAD/NADH cycling assay (see below). All experiments were performed as biological triplicates, error bars indicate standard deviations.

5.4.11. NAD/NADH cycling assays

For NAD/NADH cycling assays after *in vivo* experiments (for experimental conditions see above), the $\text{OD}_{600 \text{ nm}}$ of cultures were measured and normalized to 0.5. 0.5 mL cells were harvested by centrifuging 10 min at 3,000 x g at 4 °C. Cells were washed by adding 150 μL assay buffer of the NAD/NADH Cell-Based Assay Kit (Cayman Chemical) or PBS (when using the NAD/NADH Quantification Kit from Sigma) to the pellet, resuspending the cells, and repeating the centrifugation procedure. 150 μL assay buffer (Cayman Chemical) or NADH/NAD Extraction Buffer (Sigma) was added to the pellet and samples were stored at -80 °C until further use. To analyze the total protein amount,

17 μL of each sample was resolved by SDS-PAGE and subsequent CBB staining. The remainder of the samples were in a sonicator bath (QSONICA Q700A-220 with cup horn; amp 70%, 1s ON/4s OFF, 3 minutes total ON time) at 4 °C. After sonication, samples were filtered using centrifugal filters (Pall, 3 kDa MWCO or Merck, 10 kDa MWCO) according to the instructions provided by the manufacturer. 20 μL flowthrough was mixed with + 80 μL assay buffer (Cayman Chemical, and that mixture was used as input for the NAD/NADH Cell-Based Assay Kit (Cayman Chemical) according to the instructions provided by the manufacturer. Alternatively, 20 μL flowthrough was mixed with 30 μL NADH/NAD Extraction Buffer (Sigma) and used as input for the NAD/NADH Quantification Kit (Sigma) according to the instructions provided by the manufacturer.

5.4.12. NADP/NADPH cycling assays

The same samples that were prepared for NAD cycling assays were used for NADP cycling assays. 50 μL of each sample was used as input for the NADP/NADPH Quantitation Kit (Sigma-Aldrich) according to the instructions provided by the manufacturer.

5.4.13. RFP fluorescence

The effect of MapSPARTA on plasmid encoded RFP fluorescence was determined as described for the effect of MapSPARTA system on NAD/NADH and cell viability (described above). Cell densities of cells harboring plasmid pBK144 and therefore producing mRFP were determined by measuring the $\text{OD}_{700\text{ nm}}$ instead of $\text{OD}_{600\text{ nm}}$, as mRFP interferes with $\text{OD}_{600\text{ nm}}$ determination⁷². To measure mRFP fluorescence, cultures from the cell viability assay (described above) were used. 20h after induction of mRFP expression a 4x dilution of the culture was made and was used to determine both the mRFP fluorescence (Excitation: 584/10, Emission: 607/10) and the $\text{OD}_{700\text{ nm}}$ in a Synergy Neo2 plate reader (Biotek). Additionally a 1×10^6 dilution of the 20 h culture was made in LB and 30 μL was plated on LB agar plates supplemented with CAM and AMP. To

assess whether the T7 RNA polymerase-encoding gene in the genome of the survivor colonies was intact, a colony PCR was performed using primers pBK405 and pBK406. PCR products were analyzed by Sanger sequencing using primers pBK406-410 (Sup. Table 2).

5.4.14. Effect of MapSPARTA on plasmid transformation

E. coli BW25113 containing a BAC encoding MapSPARTA (pBK116) or MapSPARTA^{TIR-E77A} (pBK124), or an empty BAC (pBeloBAC11), were made competent using the TSS buffer protocol. Triplicate cultures of each strain were grown o/n at 37 °C in a shaker incubator. The next day the cultures were diluted 1:100 in 50 mL fresh LB and incubated for 1 h at 37 °C in a shaker incubator. At that point, arabinose was added induce SPARTA expression and cultures were further incubated until an OD_{600 nm} of 0.6 was reached. The cultures were cooled on ice for 20 min and cells were harvested by centrifugation at 3350 x g at 4 °C for 10 min. The cells were resuspended in 3 mL TSS buffer (10% w/v PEG8000, 30 mM MgCl₂, 5% v/v DMSO in LB), divided in 100 µL aliquots and stored at -80 °C. For transformation duplicates of each triplicate batch of competent cells were thawed on ice, 3 ng pBK117 was added, and the cells were incubated on ice for another 30 min. Subsequently, the cells were incubated at 42 °C for 45 sec and incubated on ice for two min. Cells were diluted in 1 mL LB and 20 µL of the cell suspension was plated on LB plates supplemented with AMP, CAM and arabinose. To determine the total number of colony forming units, every sample was further diluted 1:5*10⁴ in LB and 30 µL of the diluted cells was plated on LB agar plates supplemented with CAM and arabinose. The plates were incubated o/n at 37 °C overnight after which colonies were counted and the fraction of CAM^R colony forming units that are AMP^R was determined.

5.4.15. Plasmid loss assay

Precultures of *E. coli* BW25113 containing a BAC encoding MapSPARTA (pBK116) or MapSPARTA^{TIR-E77A} (pBK124), or an empty BAC (pBeloBAC11), as well as pUC-empty (pBK117) in LB supplemented with CAM and AMP were

incubated o/n at 37 °C in a shaker incubator. The precultures were diluted 1:1000 in 1 mL fresh LB supplemented with CAM and arabinose, and further incubated at 37 °C in a shaker incubator for 24h. This procedure was repeated for nine days in a row. On days 0, 3, 6 and 9 cultures were diluted 1:10⁶ and 30 µL of the dilution was plated in duplicate on LB agar plates supplemented with either only CAM or both CAM and AMP. The plates were incubated o/n at 37 °C after which colonies were counted and the fraction of CAM^R cells that are AMP^R was determined. All experiments were performed as biological quadruplicates, and error bars indicate standard deviations.

5.4.16. Culture competition assay

Precultures of *E. coli* BW25113 containing a BAC encoding MapSPARTA (pBK116) or MapSPARTA^{TIR-E77A} (pBK124), or an empty BAC (pBeloBAC11), and either containing no additional plasmid or pUC-empty (pBK117) were inoculated in LB supplemented with CAM and AMP and were incubated in a shaker incubator at 37 °C o/n. The precultures were diluted 1:100 in 5 mL fresh LB supplemented with CAM and AMP and were incubated for another 3h at 37 °C in a shaker incubator. Subsequently, the amount of cells corresponding to 300 µL of OD_{600 nm} = 1 was pelleted by centrifugation for 10 min at 3,500 x g. The pellet was resuspended in 300 µL LB supplemented with CAM and arabinose. Cultures containing the same BAC but with and without pBK117 plasmids were mixed in a 24:1 ratio (with plasmid:without plasmid). The cell mixture was used to inoculate 1 mL fresh LB supplemented with CAM and arabinose. These cultures were incubated for 24 h at 37 °C in a shaker incubator. Daily for three days, each culture was diluted 1:1000 into a LB supplemented with CAM and arabinose. Each day starting from the precultures on day 0, cultures were diluted 1:10⁶ (1:4*10⁵ for the preculture mix) and 50 µL was plated on LB agar plates supplemented with CAM as well as CAM and AMP. The plates were incubated o/n at 37 °C overnight after which colonies were counted and the fraction of CAM^R cells that are AMP^R was determined. All experiments were performed as biological sextuplicates, and error bars indicate standard deviations.

5.4.17. Phage assays

- **Phage strains**

E. coli phages T1, T4, T7, Lambda-vir, M13 and Nami were used in this study. For their production, phages T1, T4, T7, and Lambda-vir were propagated in *E. coli* BL21-AI, as described previously ⁷³, and bacteriophage Nami in its host *E. coli* isolate R10256 following same procedure. Briefly, bacterial cultures at exponential growth phase (approximately 0.4 OD_{600 nm}) were infected with a phage lysate and incubated overnight. Then, cells were collected by centrifugation and the supernatant was filtered through 0.2 µm PES membranes. When required, phages were concentrated by addition of PEG-8000 and NaCl at final concentrations of 100 mg/mL and 1 M, respectively. Subsequently, cultures were incubated overnight at 4 °C, centrifuged at 11,000 x g at 4 °C for 60 min, and the phage-containing pellet was resuspended in the desired final volume of LB media. Bacteriophage M13 was produced in *E. coli* JM109(DE3) as described by manufacturer (NEB #N4040S). All phage stocks were stored at 4°C before their use and titer determined as indicated below.

- **Phage titer determination**

E. coli cultures in exponential growth phase (approximately 0.4 OD_{600 nm}) were used to titer phages stocks. For this, 100 µL of culture were mixed with 5 mL of 0.6 % LB agar at 45 °C and poured on a LB agar plate to form a bacterial lawn. Ten-fold dilutions of phages in LB were plated on the top of the bacterial lawn in 5-10 µL drops and dried for 20 min. Plates were incubated overnight upside down at 37 °C. To determine the phage titer, the number of centers of infection (plaques) were counted. *E. coli* strain BW25113 was used to titer phages T1, T4, T7, Lambda-vir, and Nami. To titer phage M13, the 0.6 % LBA was supplemented with 1 mM IPTG and 200 µg/mL of X-gal, and *E. coli* ER2738 was used.

- **Bacterial survival growth curves**

Overnight bacterial cultures of *E. coli* BW25113 containing a BAC encoding MapSPARTA (pBK116) or MapSPARTA^{TIR-E77A} (pBK124), or an empty

BAC (pBeloBAC11) were used to inoculate LB at an initial OD_{600 nm} of 0.05. At exponential growth phase (approximately 0.4 OD_{600 nm}), expression of SPARTA systems was induced with 0.2 % L-arabinose and then further incubated at 37 °C for 90 min. To obtain a standard initial cell density, cultures were normalized to an OD_{600 nm} of 0.5 (corresponding to approximately 1x10⁸ colonies forming units per mL (CFU/mL) following McFarland scale indications ⁷⁴). Then, these cultures were challenged with bacteriophages as follows. For each replicate, 190 µL of cells was dispensed into wells of flat-bottom 96-well plates and 10 µL of bacteriophages (or SM buffer for non-infection controls) was added to obtain a specific initial multiplicity of infection (MOI), as indicated in the figures. Well-plates were incubated in plate readers at 37 °C with constant shaking at 280 rpm, and automatic OD_{600 nm} reading (Epoch2 microplate reader; BioTek, Bad Friedrichshall) every 10 min during the course of the experiment. Please note that OD_{600 nm} determined in the pre-cultures differs from that in the plate reader (**Sup. Fig. S7**) due to different pathway lengths. All experiments were performed as biological triplicates and error bars indicate standard deviations.

• Phage M13 titering over time

Bacterial cultures of *E. coli* JM109(DE3) containing a plasmid encoding MapSPARTA (pBK126), or MapSPARTA^{TIR-E77A} (pBK128), or an empty plasmid (pBK133), were cultivated as described above. At exponential growth phase (approximately 0.4 OD_{600 nm}), expression of SPARTA systems was induced with 0.2 % L-arabinose and the OD_{600 nm} was normalized to 0.05. Then, M13 phage was added at an initial MOI of 10⁻⁵ and phage titer was determined at various time points as described above. All experiments were performed as biological triplicates and error bars indicate standard deviations.

• Efficiency of plating determination

To determine the efficiency of plaquing (EOP), phages were plated directly on induced cultures of *E. coli* BW25113 containing a BAC encoding MapSPARTA (pBK116) or MapSPARTA^{TIR-E77A} (pBK124), or an empty BAC (pBeloBAC11). For this, 200 µL of the bacterial cultures were mixed a specific, countable number of infectious particles (200-300 PFU/plate) and 4.5 mL of 0.6

% LB agar prewarmed at 45°C. The mixture was poured on top of 1 % LB agar plates to form a bacterial layer containing the infectious particles. After overnight incubation at 37 °C, the relative EOP was calculated dividing the number of plaques counted on each plate by the number of plaques formed in the strain containing the empty BAC. All experiments were performed as biological triplicates.

5.4.18. Phylogenetic analysis

- **Homologous sequences**

To perform the phylogenetic analysis of Ago and APAZ-domain containing proteins, homologous sequences were retrieved from various bacterial taxa. A total of 7249 annotated genomes with chromosome, scaffold, representative or reference levels were used from the RefSeq database ⁷⁵. HMMER (v3.3.1; hmmer.org) was used for the homology searches using an HMM model based on PIWI domain as query for Ago proteins and APAZ domain as query for the APAZ-domain containing proteins. All the resulting sequence hits i.e., 500 Ago homologs and 256 APAZ homologs (with an E-value < 0.001) were considered for further phylogenetic analysis. For all these sequences, protein domain architectures were identified using InterProScan (v5.51-85.0 ⁷⁶).

- **Phylogeny construction**

Homologous sequences were aligned using MAFFT E-INS-I algorithm with a maximum of 1000 iterations (v7.475 ⁷⁷). Alignment positions with 30% gap threshold (-gt 0.3) were removed prior to the phylogeny construction using trimAl (v1.4 ⁷⁸) for both Ago and APAZ families. Phylogenetic trees were built using maximum-likelihood method implemented in IQtree (v2.0.4 ⁷⁹) with automatic model selection from in-built ModelFinder ⁸⁰ and a maximum of 1000 rapid bootstraps (-bb 1000 -m MFP+MERGE). LG+F+R10 and LG+F+R6 were selected as the models of evolution based on the Bayesian Information Criterion (BIC), for Ago and APAZ trees, respectively. Obtained phylogenetic trees were visualized in iTOL (v6; itol.embl.de ⁸¹).

- **Gene neighbourhood (operons) analysis**

To identify APAZ-domain containing genes that are associated with short-Ago proteins in same operon, or any other gene family members even if in a different operon, five upstream as well as downstream genes of each short-Ago protein were analyzed. All-by-all similarity search was performed with all genes in the neighborhood using Diamond blastp (v2.0.7.145⁸²) with default settings (except --query-cover 80) to identify the homologous proteins. Obtained results were then visualized in Cytoscape (v3.7.1⁸³) to identify the clusters of the same gene family associated with each clade of short-Ago proteins.

- **Tanglegram**

Phylogenetic trees of SPARTA and TIR-APAZ families (saved in NEWICK format from iTOL) were loaded into R environment (v4.0.3) using *read.dendrogram* function of 'phylogram' R package (v2.1.0⁸⁴). Later a list of dendrograms was made using *dendlist* and a comparison was made using *tanglegram* function implemented in 'dendextend' R package (v1.15.1⁸⁵).

5.4.19. Quantification and Statistical Analysis

All statistical analyses were performed in R version 4.1.0, with exception of RNA sequencing correlation analysis, which was performed in Excel.

- ***In vitro* NAD/NADH cycling assays**

In vitro NAD/NADH cycling assays were performed in technical triplicates. The percentage of NAD⁺ consumed in each reaction was calculated as (absorption of sample / absorption of a sample containing only buffer and NAD⁺) * 100. The averages of technical triplicates are shown in **Fig. 2**, error bars indicate standard deviations.

- **ϵ -NAD assays**

All assays with ssDNA targets were performed in technical triplicates, detection of dsDNA targets with different T7 exonuclease concentrations. For determining the effect of mismatches (**Fig. 6A-B**), guide length preferences (**Sup. Fig. S10B**), and buffer pH preferences (**Sup. Fig. S10C**) the fluorescence was first corrected for background fluorescence determined using a no-target control, and

change in fluorescence min^{-1} was determined by calculating the slopes of linear components of the kinetic curves. For the detection of mismatches (**Fig. 6A-B**), the relative change in fluorescence compared to no mismatches was calculated as (change in fluorescence min^{-1} in sample / change in fluorescence min^{-1} in no mismatch control). For determining the effect of temperature (**Sup. Fig. S10A**) fluorescence was measured after 60 min incubation. For determining the limit of detection (**Fig. 6C-D**) the fluorescence was measured in kinetic mode and corrected with fluorescence of a no target control sample. For the detection of dsDNA the fluorescence was also measured in kinetic mode, but without subtracting the control (the control graph is added to the figure instead). For triplicate experiments in all panels the average are used to plot the graphs and error bars indicate standard deviations.

- **RNA and DNA sequences correlation analysis**

Correlation between total long RNA and SPARTA-associated small RNA sequence abundance and MapSPARTA-associated small RNA and DNA sequence abundance was tested for each comparison by calculating the Pearson correlation coefficient and corresponding p-value.

- **In vivo assays to assess effect of pUC-empty on cells expressing MapSPARTA**

All assays were performed in biological triplicates of which the averages are shown in **Fig. 5A** and **Sup. Fig. S4A**, error bars indicate standard deviations. Per timepoint (0 h, 4 h and 20 h) and per dependent variable (CFU, $\text{OD}_{600 \text{ nm}}$, total NAD level and total NADP level) a two-way ANOVA was performed to test for significant independent and interaction effects of SPARTA system (MapSPARTA or MapSPARTA^{TIR-E77A} or empty BAC) and presence of pUC-empty. The data was subjected to a Shapiro-Wilk test on linear model residuals and Levene's test to confirm normal distribution of the data and homogeneity of variances, respectively. Tukey's test was used for post-hoc pairwise comparisons.

- ***In vivo* assays to assess effect of pUC-mRFP, pET-mRFP and pC101-mRFP on cells expressing MapSPARTA**

All assays were performed in biological triplicates of which the averages are shown in **Fig. 5A** and **Sup. Fig. S4A**, error bars indicate standard deviations. For timepoint 4 h and as the dependent variable the total NAD level, and for timepoints 4h and 20 and as the dependent variable the CFU, two-way ANOVAs were performed to test for significant independent and interaction effects of presence of pUC-empty and T7-based transcription of mRFP. The data was subjected to a Shapiro-Wilk test on linear model residuals and Levene's test to confirm normal distribution of the data and homogeneity of variances, respectively. Tukey's test was used for post-hoc pairwise comparisons.

- ***In vivo* assays to assess effect of multicopy plasmids containing a non-transcribed target on cells expressing MapSPARTA and mRFP from a BAC**

All assays were performed in biological duplicates of which the average and individual datapoints are shown in **Fig. 5A** and **Sup. Fig. S4A**. For timepoint 4 h and as the dependent variable the total NAD level, and for timepoints 4h and 20 and as the dependent variable the CFU, two-way ANOVAs were performed to test for significant independent and interaction effects of presence of a multicopy plasmid and presence of a non-transcribed target on said plasmid. Normal distribution of the data and homogeneity of variances were assumed. Tukey's test was used for post-hoc pairwise comparisons.

- ***In vivo* assay to assess effect of pUC-mRFP and derivatives on cells expressing MapSPARTA**

Cell viability assays (measured as CFU/mL) were performed in biological triplicates of which the averages are shown in **Fig. 5B**, error bars indicate standard deviations. Per timepoint (4 h and 20 h) a two-way ANOVA was performed to test for significant independent and interaction effects of SPARTA system (MapSPARTA, MapSPARTA^{TIR-E77A} or empty BAC) and pUC-variant (no pUC, pUC-mRFP, pUC-mRFP- Δ T7- Δ RBBS or pUC-mRFP- Δ RBBS). The data was subjected to a Shapiro-Wilk test on linear model residuals and Levene's test to

confirm normal distribution of the data and homogeneity of variances, respectively. Tukey's test was used for post-hoc pairwise comparisons.

- **RFP fluorescence of cells expressing MapSPARTA**

RFP fluorescence and OD_{700 nm} in bacterial cultures was measured in biological triplicates of which the RFP fluorescence/ OD_{700 nm} averages are shown in **Fig. 5C**, error bars indicate standard deviations. A t-test was used to test for significant difference between cells expressing MapSPARTA and MapSPARTA^{TIR-E77A}. The data was subjected to a Shapiro-Wilk test on linear model residuals and Levene's test to confirm normal distribution of the data and homogeneity of variances, respectively.

- **Transformation assay**

Transformation assays were performed in biological triplicates that consisted of technical replicates that were averaged. The transformation efficiency was calculated as AMP^R+CAM^R CFU/CAM^R CFU. The average transformation efficiencies of the biological triplicates are shown in **Sup. Fig. S9C**, error bars indicate standard deviations. One-way ANOVA was used to test for significant effect of SPARTA system on transformation efficiency (MapSPARTA, MapSPARTA^{TIR-E77A} or empty BAC). The data was subjected to a Shapiro-Wilk test on linear model residuals and Levene's test to confirm normal distribution of the data and homogeneity of variances, respectively. As the one-way ANOVA was not significant, no post-hoc tests were performed.

- **Bacterial growth curves and phage M13 titering over time**

Both experiments were performed in biological triplicates. The averages of the curves are shown in **Sup. Fig. S6**, shadings (growth curves) or error bars (M13 titering over time) indicate standard deviations. For statistical analyses, the areas under the curves were calculated using the Trapezoidal rule. For every phage that was used a one-way ANOVA was performed to test for significant effect of SPARTA system on the area under the growth curve. The data was subjected to a Shapiro-Wilk test on linear model residuals and Levene's test to confirm normal distribution of the data and homogeneity of variances, respectively. Tukey's test was used for post-hoc pairwise comparisons.

- **Plasmid loss assay**

Plasmid loss assays were performed in biological quadruplicates and plated in duplicates that were averaged. The fraction of AMP^R cells was calculated as CFU on AMP+CAM/CFU on CAM. The average fractions of AMP^R cells of the biological triplicates are shown in **Sup. Fig. S9B**, error bars indicate standard deviations. A two-way ANOVA was performed to test for significant independent and interaction effects of SPARTA system (MapSPARTA or MapSPARTA^{TIR-E77A}) and timepoint (day 1 or day 9). The data was subjected to a Shapiro-Wilk test on linear model residuals and Levene's test to confirm normal distribution of the data and homogeneity of variances, respectively. As the two-way ANOVA was not significant, no post-hoc tests were performed.

- **Culture competition assay**

Culture competition assays were performed in biological sextuplets. The fraction of AMP^R cells was calculated as CFU on AMP+CAM/CFU on CAM. The average fractions of AMP^R cells are shown in **Fig. 6F**, error bars indicate standard deviations. A mixed two-way ANOVA was performed to test for significant independent and interaction effects of SPARTA system (MapSPARTA or MapSPARTA^{TIR-E77A}) and timepoint (day 0, 1, 2 or 3) as within-subjects variable. The data was subjected to a Shapiro-Wilk test on linear model residuals and Levene's test to confirm normal distribution of the data and homogeneity of variances, respectively. One-way ANOVA's and pairwise t-tests with Bonferroni corrections were used to analyse simple main effects and for pairwise comparisons, respectively.

5.4.20. Resource Availability

- **Materials availability**

Protein expression vectors encoding MapSPARTA (pBK086) and CrtSPARTA (pBK092) have been deposited on Addgene (plasmid numbers 183145 and 183146). Other newly generated materials used in this paper will be shared upon request.

- **Data and code availability**

The unprocessed image files, raw microplate reader data, and RNA/DNA-seq analyses used to prepare the figures in this manuscript have been deposited in Mendeley Data (<https://dx.doi.org/10.17632/7f4gsbhrpp.1>). The raw RNA-seq data have been deposited in the NCBI's Gene Expression Omnibus (GEO; accession number GSE191271). This paper does not report original code.

5.5. References

1. Peters,L. and Meister,G. (2007) Argonaute Proteins: Mediators of RNA Silencing. *Mol. Cell*, 26.
2. Vaucheret,H. (2008) Plant ARGONAUTES. *Trends Plant Sci.*, 13.
3. Shabalina,S.A. and Koonin,E. V. (2008) Origins and evolution of eukaryotic RNA interference. *Trends Ecol. Evol.*, 23.
4. Olovnikov,I., Chan,K., Sachidanandam,R., Newman,D.K. and Aravin,A.A. (2013) Bacterial Argonaute samples the transcriptome to identify foreign DNA. *Mol. Cell*, 51.
5. Swarts,D.C., Jore,M.M., Westra,E.R., Zhu,Y., Janssen,J.H., Snijders,A.P., Wang,Y., Patel,D.J., Berenguer,J., Brouns,S.J.J., et al. (2014) DNA-guided DNA interference by a prokaryotic Argonaute. *Nature*, 507.
6. Kuzmenko,A., Oguienko,A., Esyunina,D., Yudin,D., Petrova,M., Kudinova,A., Maslova,O., Ninova,M., Ryazansky,S., Leach,D., et al. (2020) DNA targeting and interference by a bacterial Argonaute nuclease. *Nature*.
7. Jolly,S.M., Gainetdinov,I., Jouravleva,K., Zhang,H., Strittmatter,L., Bailey,S.M., Hendricks,G.M., Dhabaria,A., Ueberheide,B. and Zamore,P.D. (2020) *Thermus thermophilus* Argonaute Functions in the Completion of DNA Replication. *Cell*, 182.
8. Jin,Z., Tu,Z., Zhang,A., Fu,L., Han,L., Xiang,Y., Xie,C. and Jin,M. (2019) The prokaryotic Argonaute proteins enhance homology sequence-directed recombination in bacteria. *Nucleic Acids Res.*, 47.
9. Lee,K.Z., Mechikoff,M.A., Kikla,A., Liu,A., Fitzgerald,K., Gimble,F.S., Solomon,K. V and Pandolfi,P. (2021) NgAgo possesses guided DNA nicking activity. 49.
10. Swarts,D.C., Hegge,J.W., Hinojo,I., Shiimori,M., Ellis,M.A., Dumrongkulraksa,J., Terns,R.M., Terns,M.P. and Van Der Oost,J. (2015) Argonaute of the archaeon *Pyrococcus furiosus* is a DNA-guided nuclease that targets cognate DNA. *Nucleic Acids Res.*, 43.
11. Zander,A., Willkomm,S., Ofer,S., Van Wolferen,M., Egert,L., Buchmeier,S., Stöckl,S., Tinnefeld,P., Schneider,S., Klingl,A., et al. (2017) Guide-independent DNA cleavage by archaeal Argonaute from *Methanocaldococcus jannaschii*. *Nat. Microbiol.*, 2.
12. Doxzen,K.W. and Doudna,J.A. (2017) DNA recognition by an RNA-guided bacterial Argonaute. *PLoS One*, 12.
13. Kaya,E., Kranzusch,P.J., Wilson,R.C., Strutt,S.C., Knoll,K.R., Doxzen,K.W. and Doudna,J.A. (2016) A bacterial Argonaute with noncanonical guide RNA specificity. *Proc. Natl. Acad. Sci.*, 113.
14. Lapinaite,A., Doudna,J.A. and Cate,J.H.D. (2018) Programmable RNA recognition using a CRISPR-associated Argonaute. *Proc. Natl. Acad. Sci.*, 115.
15. Kropocheva,E., Kuzmenko,A., Aravin,A.A., Esyunina,D. and Kulbachinskiy,A. (2021) A programmable pAgo nuclease with universal guide and target specificity from the mesophilic bacterium *Kurthia massiliensis*. *Nucleic Acids Res.*, 49.
16. Li,W., Liu,Y., Wang,F. and Ma,L. (2021) A programmable pAgo nuclease with RNA target preference from the psychrotolerant bacteria *Mucilaginibacter paludis*. *bioRxiv*.

17. Ryazansky,S., Kulbachinskiy,A. and Aravin,A.A. (2018) The Expanded Universe of Prokaryotic Argonaute Proteins. *MBio*, 9.
18. Makarova,K.S., Wolf,Y.I., van der Oost,J. and Koonin,E. V. (2009) Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biol. Direct*, 4.
19. Swarts,D.C., Makarova,K., Wang,Y., Nakanishi,K., Ketting,R.F., Koonin,E. V., Patel,D.J. and Van Der Oost,J. (2014) The evolutionary journey of Argonaute proteins. *Nat. Struct. Mol. Biol.*, 21.
20. Dayeh,D.M., Cantara,W.A., Kitzrow,J.P., Musier-Forsyth,K. and Nakanishi,K. (2018) Argonaute-based programmable RNase as a tool for cleavage of highly-structured RNA. *Nucleic Acids Res.*, 46.
21. Ma,J.-B., Yuan,Y.-R., Meister,G., Pei,Y., Tuschl,T. and Patel,D.J. (2005) Structural basis for 5'-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature*, 434.
22. Miyoshi,T., Ito,K., Murakami,R. and Uchiumi,T. (2016) Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. *Nat. Commun.*, 7.
23. Parker,J.S., Parizotto,E.A., Wang,M., Roe,S.M. and Barford,D. (2009) Enhancement of the Seed-Target Recognition Step in RNA Silencing by a PIWI/MID Domain Protein. *Mol. Cell*, 33.
24. Burroughs,A.M., Iyer,L.M. and Aravind,L. (2013) Two novel PIWI families: Roles in inter-genomic conflicts in bacteria and Mediator-dependent modulation of transcription in eukaryotes. *Biol. Direct*, 8.
25. Willkomm,S., Oellig,C.A., Zander,A., Restle,T., Keegan,R., Grohmann,D. and Schneider,S. (2017) Structural and mechanistic insights into an archaeal DNA-guided Argonaute protein. *Nat. Microbiol.*, 2.
26. Hegge,J.W., Swarts,D.C. and Van Der Oost,J. (2018) Prokaryotic argonaute proteins: Novel genome-editing tools? *Nat. Rev. Microbiol.*, 16.
27. Kim,S., Jung,Y. and Lim,D. (2020) Argonaute system of *Kordia jejudonensis* is a heterodimeric nucleic acid-guided nuclease. *Biochem. Biophys. Res. Commun.*, 525.
28. Horsefield,S., Burdett,H., Zhang,X., Manik,M.K., Shi,Y., Chen,J., Qi,T., Gilley,J., Lai,J.-S., Rank,M.X., et al. (2019) NAD⁺ cleavage activity by animal and plant TIR domains in cell death pathways. *Science*, 365.
29. Morehouse,B.R., Govande,A.A., Millman,A., Keszei,A.F.A., Lowey,B., Ofir,G., Shao,S., Sorek,R. and Kranzusch,P.J. (2020) STING cyclic dinucleotide sensing originated in bacteria. *Nature*, 586.
30. Johnson,C.M., Harden,M.M. and Grossman,A.D. (2020) An integrative and conjugative element encodes an abortive infection system to protect host cells from predation by a bacteriophage. *bioRxiv*.
31. Ofir,G., Herbst,E., Baroz,M., Cohen,D., Millman,A., Doron,S., Tal,N., Malheiro,D.B.A., Malitsky,S., Amitai,G., et al. (2021) Antiviral activity of bacterial TIR domains via immune signalling molecules. *Nature*, 600.

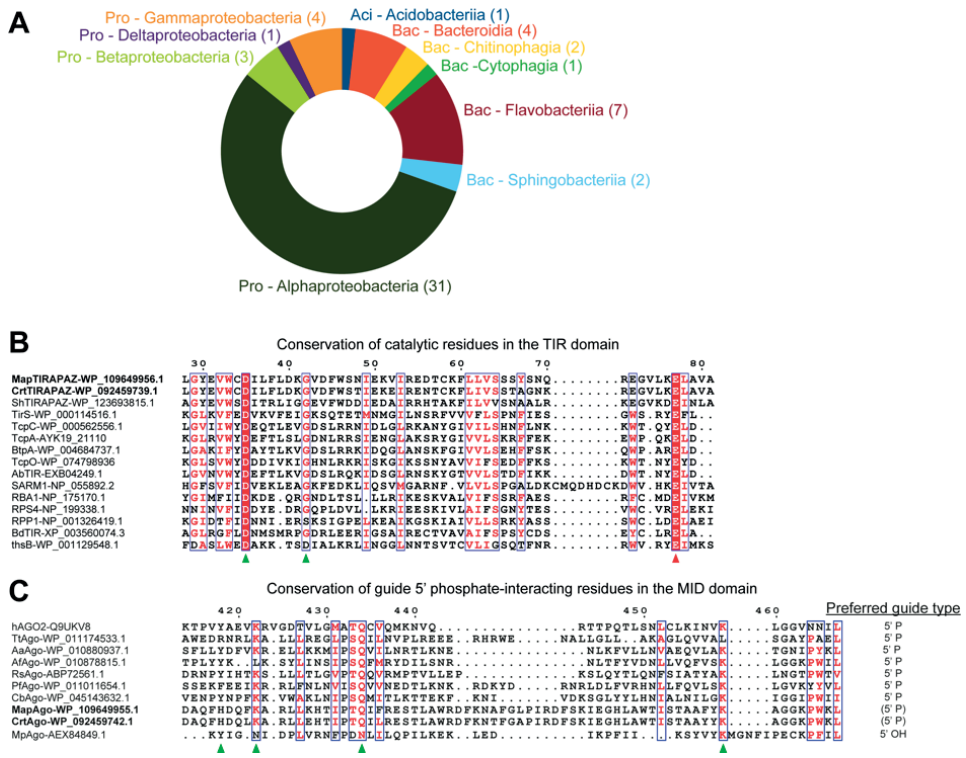
32. Wan,L., Essuman,K., Anderson,R.G., Sasaki,Y., Monteiro,F., Chung,E.H., Nishimura,E.O., DiAntonio,A., Milbrandt,J., Dangel,J.L., et al. (2019) TIR domains of plant immune receptors are NAD⁺-cleaving enzymes that promote cell death. *Science*, 365.
33. Essuman,K., Summers,D.W., Sasaki,Y., Mao,X., Yim,A.K.Y., DiAntonio,A. and Milbrandt,J. (2018) TIR Domain Proteins Are an Ancient Family of NAD⁺-Consuming Enzymes. *Curr. Biol.*, 28.
34. Frank,F., Sonenberg,N. and Nagar,B. (2010) Structural basis for 5' -nucleotide base-specific recognition of guide RNA by human AGO2. *Nature*, 465.
35. Maute,R.L., Schneider,C., Sumazin,P., Holmes,A., Califano,A., Basso,K. and Dalla-Favera,R. (2013) TRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma. *Proc. Natl. Acad. Sci. U. S. A.*, 110.
36. Nakanishi,K. (2021) Are Argonaute-Associated Tiny RNAs Junk, Inferior miRNAs, or a New Type of Functional RNAs? *Front. Mol. Biosci.*, 8.
37. Shizuya,H., Birren,B., Kim,U.J., Mancino,V., Slepak,T., Tachiiri,Y. and Simon,M. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl. Acad. Sci. U. S. A.*, 89.
38. Jeffrey,V. and Joachim,M. (1991) New pUC-derived cloning vectors with different selectable markers and DNA replication origins. *Gene*, 100.
39. Brautaset,T., Lale,R. and Valla,S. (2009) Positively regulated bacterial expression systems. *Microb. Biotechnol.*, 2.
40. Siguier,P., Filée,J. and Chandler,M. (2006) Insertion sequences in prokaryotic genomes. *Curr. Opin. Microbiol.*, 9.
41. Hasunuma,K. and Sekiguchi,M. (1977) Replication of plasmid pSC101 in *Escherichia coli* K12: Requirement for dnaA function. *MGG Mol. Gen. Genet.*, 154.
42. Wang,I.N., Dykhuizen,D.E. and Slobodkin,L.B. (1996) The evolution of phage lysis timing. *Evol. Ecol.*, 10.
43. Doron,S., Melamed,S., Ofir,G., Leavitt,A., Lopatina,A., Keren,M., Amitai,G. and Sorek,R. (2018) Systematic discovery of antiphage defense systems in the microbial pangenome. *Science*, 4120.
44. Gao,L., Altae-Tran,H., Böhning,F., Makarova,K.S., Segel,M., Schmid-Burgk,J.L., Koob,J., Wolf,Y.I., Koonin,E. V. and Zhang,F. (2020) Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science*, 1084.
45. Chen,J.S., Ma,E., Harrington,L.B., Da Costa,M., Tian,X., Palefsky,J.M. and Doudna,J.A. (2018) CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science*, 360.
46. Gootenberg,J.S., Abudayyeh,O.O., Lee,J.W., Essletzbichler,P., Dy,A.J., Joung,J., Verdine,V., Donghia,N., Daringer,N.M., Freije,C.A., et al. (2017) Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science*, 356.
47. Gootenberg,J.S., Abudayyeh,O.O., Kellner,M.J., Joung,J., Collins,J.J. and Zhang,F. (2018) Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a and Csm6. *Science*, 360.

48. Li,L., Li,S., Wu,N., Wu,J., Wang,G., Zhao,G. and Wang,J. (2019) HOLMESv2: A CRISPR-Cas12b-Assisted Platform for Nucleic Acid Detection and DNA Methylation Quantitation. *ACS Synth. Biol.*, 8.
49. Steens,J.A., Zhu,Y., Taylor,D.W., Bravo,J.P.K., Prinsen,S.H.P., Schoen,C.D., Keijser,B.J.F., Ossendrijver,M., Hofstra,L.M., Brouns,S.J.J., et al. (2021) SCOPE enables type III CRISPR-Cas diagnostics using flexible targeting and stringent CARF ribonuclease activation. *Nat. Commun.*, 12.
50. To,K.K.W., Chan,K.-H., Li,I.W.S., Tsang,T.-Y., Tse,H., Chan,J.F.W., Hung,I.F.N., Lai,S.-T., Leung,C.-W., Kwan,Y.-W., et al. (2009) Viral Load in Patients Infected With Pandemic H1N1 2009 Influenza A Virus. *J. Med. Virol.*, 82.
51. Pan,Y., Zhang,D., Yang,P., Poon,L.L.M. and Wang,Q. (2020) Viral load of SARS-CoV-2 in clinical samples. *Lancet Infect. Dis.*, 20.
52. Jet,T., Gines,G., Rondelez,Y. and Taly,V. (2021) Advances in multiplexed techniques for the detection and quantification of microRNAs. *Chem. Soc. Rev.*, 50.
53. Oh,M., Park,W.B., Choe,P.G., Choi,S.-J., Kim,J.-I., Chae,J., Park,S.S., Kim,E.-C., Oh,H.S., Kim,E.J., et al. (2016) Viral Load Kinetics of MERS Coronavirus Infection. *N. Engl. J. Med.*, 375.
54. Harrington,L.B., Burstein,D., Chen,J.S., Paez-Espino,D., Ma,E., Witte,I.P., Cofsky,J.C., Kyrpides,N.C., Banfield,J.F. and Doudna,J.A. (2018) Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. *Science*, 362.
55. Tal,N., Morehouse,B.B., Millman,A., Stokar-avihail,A., Avraham,C., Fedorenko,T., Yirmiya,E., Herbst,E., Brandis,A., Mehlman,T., et al. (2021) Cyclic CMP and cyclic UMP mediate bacterial immunity against phages. *Cell*, 184.
56. Ma,S., Lapin,D., Liu,L., Sun,Y., Song,W., Zhang,X. and Logemann,E. (2020) Direct pathogen-induced assembly of an NLR immune receptor complex to form a holoenzyme. *Science*, 370.
57. Martin,R., Qi,T., Zhang,H., Liu,F., King,M., Toth,C., Nogales,E. and Staskawicz,B.J. (2020) Structure of the activated ROQ1 resistosome directly recognizing the pathogen effector XopQ. *Science*, 370.
58. Ka,D., Oh,H., Park,E., Kim,J. and Bae,E. (2020) Structural and functional evidence of bacterial antiphage protection by Thoeris defense system via NAD⁺ degradation. *Nat. Commun.*
59. Essuman,K., Summers,D.W., Sasaki,Y., Mao,X., DiAntonio,A. and Milbrandt,J. (2017) The SARM1 Toll/Interleukin-1 Receptor Domain Possesses Intrinsic NAD⁺ Cleavage Activity that Promotes Pathological Axonal Degeneration. *Neuron*, 93.
60. Killackey,S.A., Rahman,M.A., Soares,F., Zhang,A.B., Abdel-Nour,M., Philpott,D.J. and Girardin,S.E. (2019) The mitochondrial Nod-like receptor NLRX1 modifies apoptosis through SARM1. *Mol. Cell. Biochem.*, 453.
61. Hopkins,E.L., Gu,W., Kobe,B. and Coleman,M.P. (2021) A Novel NAD Signaling Mechanism in Axon Degeneration and its Relationship to Innate Immunity. *Front. Mol. Biosci.*, 8.
62. Lopatina,A., Tal,N. and Sorek,R. (2020) Abortive Infection: Bacterial Suicide as an Antiviral Immune Strategy. *Annu. Rev. Virol.*, 7.

63. Page,R. and Peti,W. (2016) Toxin-antitoxin systems in bacterial growth arrest and persistence. *Nat. Chem. Biol.*, 12.
64. Meeske,A.J., Nakandakari-Higa,S. and Marraffini,L.A. (2019) Cas13-induced cellular dormancy prevents the rise of CRISPR-resistant bacteriophage. *Nature*, 570.
65. Dimitriu,T., Kurilovich,E., Łapińska,U., Severinov,K., Pagliara,S., Szczelkun,M.D. and Westra,E.R. (2021) Bacteriostatic antibiotics promote CRISPR-Cas adaptive immunity by enabling increased spacer acquisition. *Cell Host Microbe*, 30.
66. Correa,A.M.S., Howard-Varona,C., Coy,S.R., Buchan,A., Sullivan,M.B. and Weitz,J.S. (2021) Revisiting the rules of life for viruses of microorganisms. *Nat. Rev. Microbiol.*
67. Dym,O. and Eisenberg,D. (2001) Sequence-structure analysis of FAD-containing proteins. 10.
68. Bushnell,B., Rood,J. and Singer,E. (2017) BBMerge – Accurate paired shotgun read merging via overlap. *PLoS One*, 12.
69. Kim,D., Langmead,B. and Salzberg1,S.L. (2015) HISAT: a fast spliced aligner with low memory requirements Daehwan HHS Public Access. *Nat. Methods*, 12.
70. Liao,Y., Smyth,G.K. and Shi,W. (2014) FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30.
71. Williamson,M.P. (2005) Nuclear magnetic resonance spectroscopy techniques | Nuclear Overhauser Effect. *Encycl. Anal. Sci.*
72. Hecht,A., Endy,D., Salit,M. and Munson,M.S. (2016) When Wavelengths Collide: Bias in Cell Abundance Measurements Due to Expressed Fluorescent Proteins. *ACS Synth. Biol.*, 5.
73. Bonilla,N., Rojas,M.I., Netto Flores Cruz,G., Hung,S.-H., Rohwer,F. and Barr,J.J. (2016) Phage on tap-a quick and efficient protocol for the preparation of bacteriophage laboratory stocks. *PeerJ*, 4.
74. Mcfarland,J. (1907) The Nephelometer: An instrument for estimating the number of bacteria in suspensions used for calculating the opsonic index and for vaccines.
75. O'Leary,N.A., Wright,M.W., Brister,J.R., Ciufu,S., Haddad,D., McVeigh,R., Rajput,B., Robbertse,B., Smith-White,B., Ako-Adjei,D., et al. (2016) Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.*, 44.
76. Jones,P., Binns,D., Chang,H.Y., Fraser,M., Li,W., McAnulla,C., McWilliam,H., Maslen,J., Mitchell,A., Nuka,G., et al. (2014) InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, 30.
77. Katoh,K. and Standley,D.M. (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.*, 30, 772–780.
78. Capella-Gutiérrez,S., Silla-Martínez,J.M. and Gabaldón,T. (2009) trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25.
79. Minh,B.Q., Schmidt,H.A., Chernomor,O., Schrempf,D., Woodhams,M.D., Von Haeseler,A., Lanfear,R. and Teeling,E. (2020) IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.*, 37.

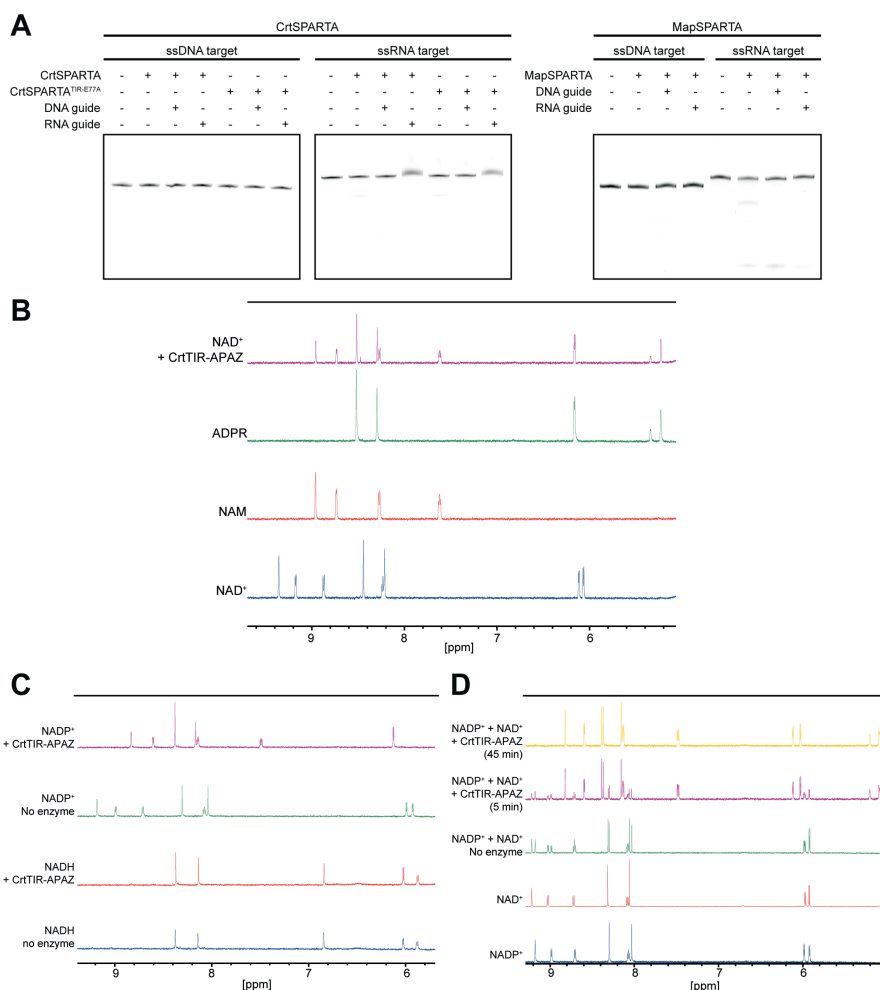
80. Kalyaanamoorthy,S., Minh,B.Q., Wong,T.K.F., Von Haeseler,A. and Jermin,L.S. (2017) ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods*, 14.
81. Letunic,I. and Bork,P. (2021) Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.*, 49.
82. Buchfink,B., Reuter,K. and Drost,H.G. (2021) Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods*, 18.
83. Shannon,P., Markiel,A., Ozier,O., Baliga,N.S., Wang,J.T., Ramage,D., Amin,N., Schwikowski,B. and Ideker,T. (2003) Cytoscape: A Software Environment for Integrated Models. *Genome Res.*, 13.
84. P. Wilkinson,S. and K. Davy,S. (2018) phylogram: an R package for phylogenetic analysis with nested lists. *J. Open Source Softw.*, 3.
85. Galili,T. (2015) dendextend: An R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics*, 31.

5.6. Supplementary figures

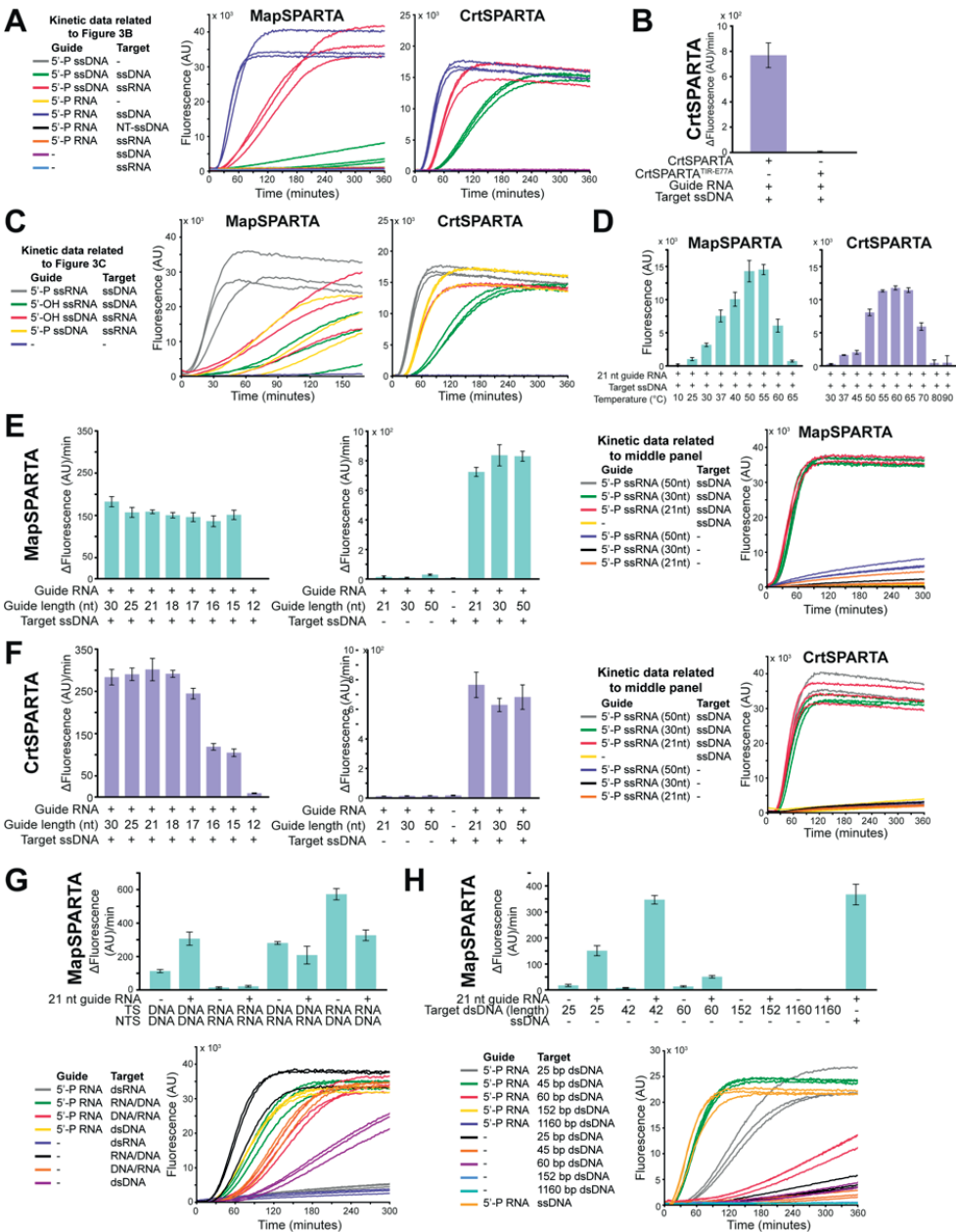


Supplementary Figure S1. Distribution of and conservation of catalytic residues in SPARTA systems, related to Fig. 1. **A)** Circle diagram indicating bacterial phyla (abbreviated) and classes in which SPARTA systems are found. Aci: Acidobacteria. Bac: Bacteroidetes. Pro: Proteobacteria. **B)** Multiple sequence alignment of TIR domains from SPARTA systems with various prokaryotic and eukaryotic TIR domains that demonstrate NADase activity. Green and red triangles indicate catalytic residues important for NADase activity, red triangle indicates residue substituted in CrtTIR-APAZ^{E77A} and MapTIR-APAZ^{E77A}. **C)** Multiple sequence alignment of MID domains from different prokaryotic Argonaute proteins and their preferred guide types. Green triangles indicate residues important for 5' phosphate binding.

Short pAgos trigger cell death upon detection of invading DNA

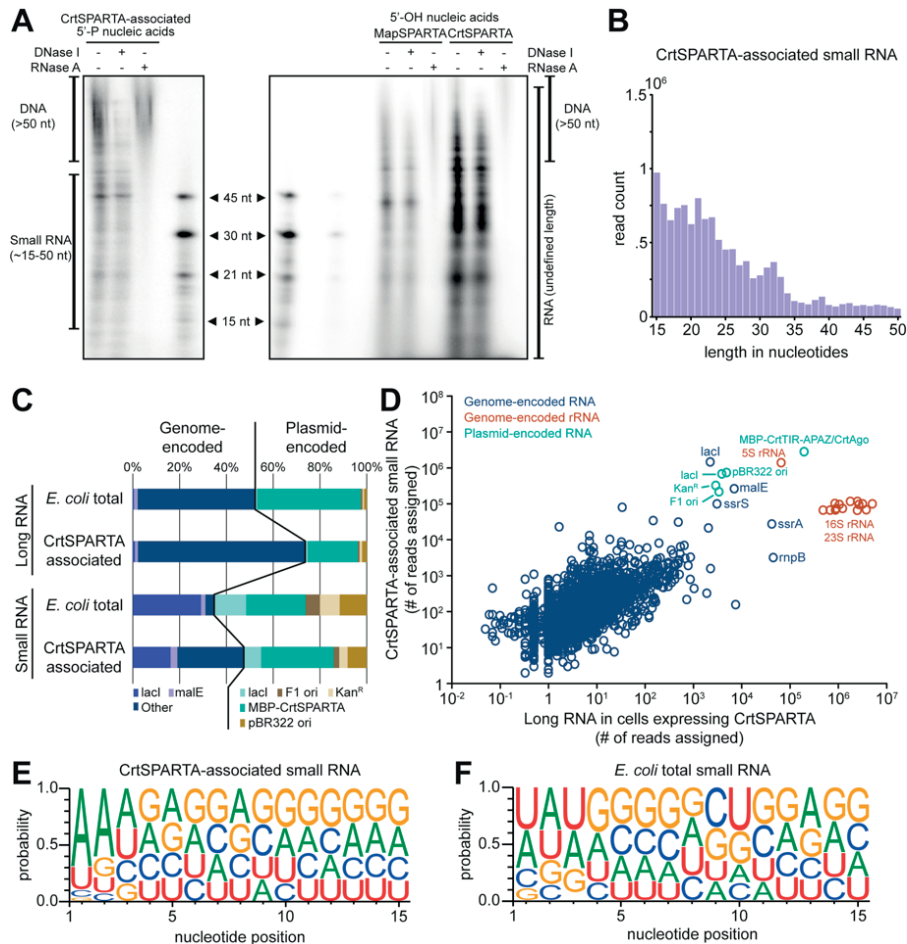


Supplementary Figure S2. CrtTIR-APAZ converts NAD(P)⁺ to nicotinamide and adenosine diphosphate ribose (phosphate), related to Fig. 2. A) SPARTA does not cleave nucleic acids in a guide-dependant manner. CrtSPARTA and MapSPARTA were incubated with different guides and Cy5-labeled DNA or RNA target in a 10:10:1 molar ratio (SPARTA:guide:target). Cleavage products were resolved on denaturing (7M) urea PAGE gels. B) TIR-APAZ converts NAD⁺ into ADPR and NAM. NMR spectra show the reaction products from NAD⁺ converted by CrtTIR-APAZ correspond to the spectra of nicotinamide (NAM) and adenosine diphosphate ribose (ADPR). C) TIR-APAZ degrades NADP⁺ but not NADH. NMR spectra of NADP⁺ (upper two panels) and NADH (lower two panels) incubated with and without CrtTIR-APAZ. D) TIR-APAZ shows no preference for NAD⁺ or NADP⁺. NAD⁺ and NADP⁺ were mixed in equimolar ratios and incubated with CrtTIR-APAZ. NMR spectra show that during incubation with CrtTIR-APAZ, NAD⁺ and NADP⁺ are converted with a similar reaction rate.



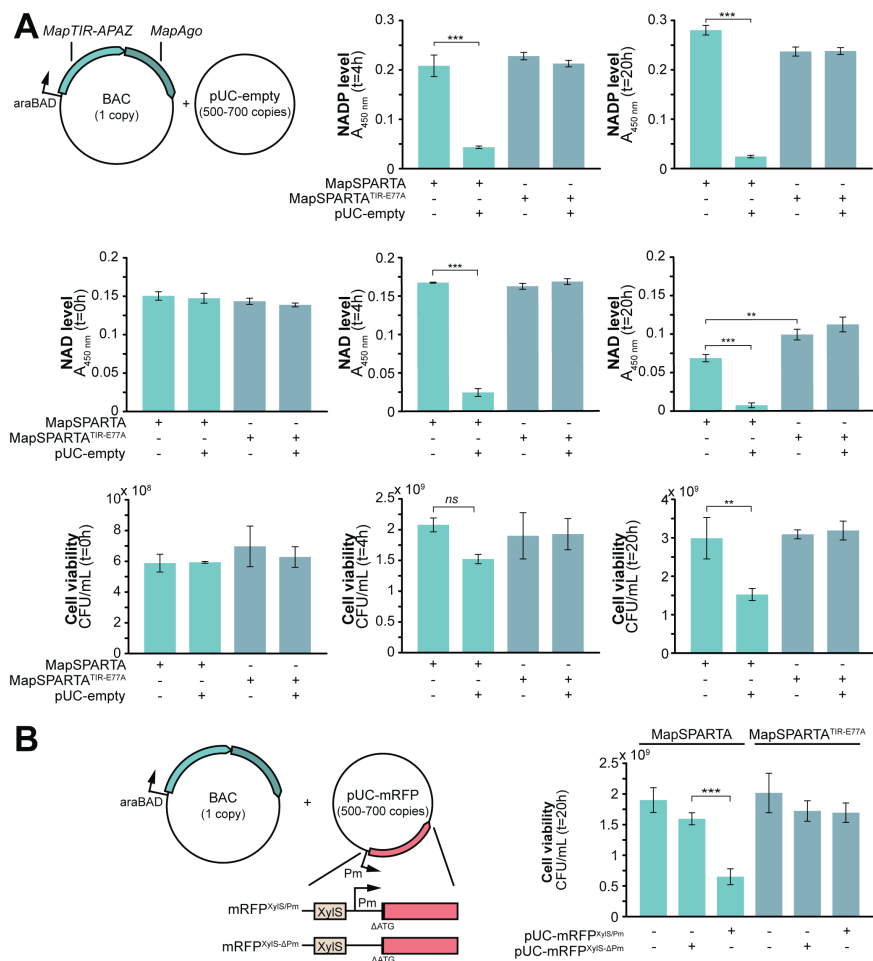
Supplementary Figure S3. SPARTA is preferentially activated upon guide RNA-mediated ssDNA target binding, related to Fig. 3. **A)** Kinetic data related to Fig. 3B. Graph indicates total fluorescence measured over time. **B)** The catalytic mutant CrtSPARTA^{TIR-E77A} does not show NADase activity even when incubated with guide RNA and target ssDNA. Graph indicates the maximum change in fluorescence over time. **C)** Kinetic data related to Fig. 3C. Graph indicates

[Sup. Fig. S3]... total fluorescence measured over time. **D)** Effect of incubation temperature on activity of MapSPARTA (left) or CrtSPARTA (right) incubated with guide RNA and target ssDNA. Graph indicates end-point fluorescence intensity after 60 min of incubation. **E-F)** Effect of guide RNA length on activity of MapSPARTA (**E**) or CrtSPARTA (**F**). Left and middle graphs indicate the maximum change in fluorescence over time. Right panels show kinetic data related to the middle panels and indicate total fluorescence measured over time. **G-H)** MapSPARTA is activated by RNA/DNA hybrids (**G**) and guide RNA-mediated detection of small dsDNA targets (**H**). Graphs indicate the maximum change in fluorescence over time. Lower panels show the kinetic data related to the upper panels and indicate total fluorescence measured over time. For all panels, CrtSPARTA, CrtSPARTA^{TIR-E77A}, or MapSPARTA was mixed with guide and target oligonucleotides in a 1:1:2 molar ratio and incubated with ϵ -NAD⁺. The average of three technical replicates is shown, error bars indicate standard deviations.

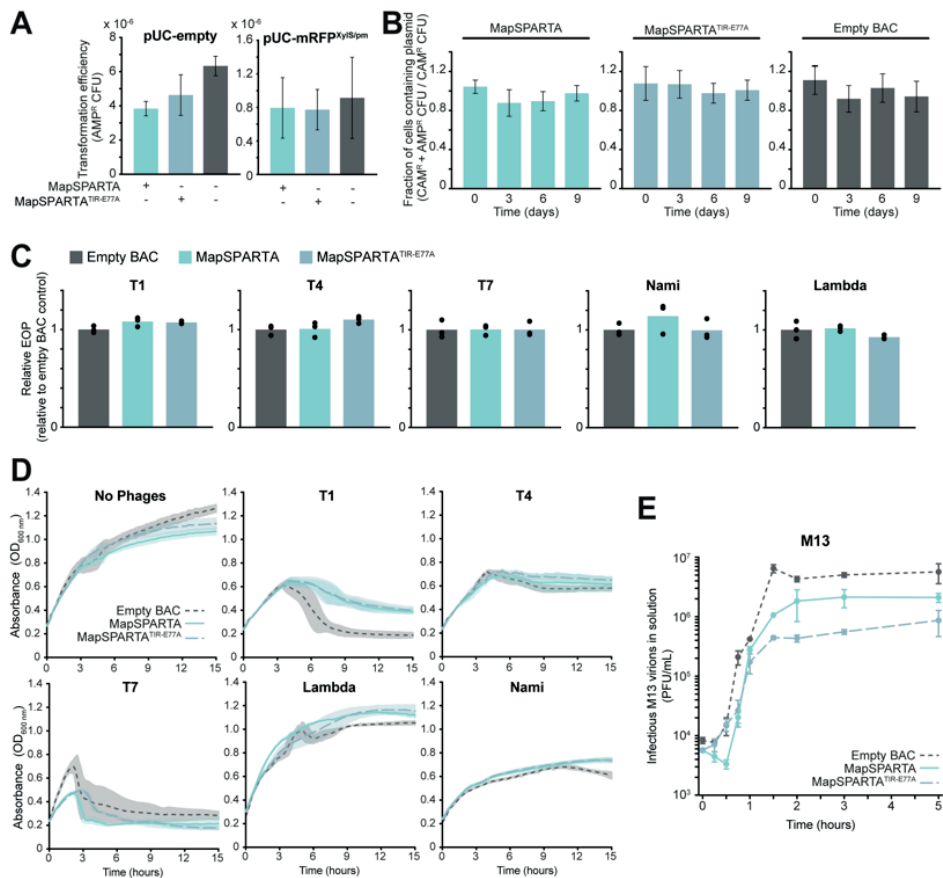


Supplementary Figure S4. CrtSPARTA associates with small RNAs targeting highly transcribed genes, related to Fig. 4. A) Left: CrtPARTA associates with 5' phosphorylated (5'-P) small RNAs. Right: MapSPARTA and CrtSPARTA associate with 5'-OH nucleic acids of undefined lengths. Nucleic acids that co-purified with CrtSPARTA or MapSPARTA were [γ - 32 P]-ATP labelled, treated with DNase I or RNase A, and resolved on a denaturing polyacrylamide gel. nt: nucleotides. **B)** Length distribution of small co-purified with CrtSPARTA as determined by sequencing. **C)** Percentages of long RNAseq or small RNAseq reads that align to specific genomic or plasmid sequences. RNA was extracted from *E. coli* expressing CrtSPARTA (*E. coli* total) or from purified CrtSPARTA (CrtSPARTA associated). rRNA-derived reads are excluded from this graph. **D)** CrtSPARTA-associated small RNA sequences and *E. coli*-extracted long RNA sequences correlate (Pearson Correlation coefficient $r > 0.99$, $p < 10^{-99}$). **E-F)** Small RNA associated with CrtSPARTA have a bias for 5' adenosine bases. Nucleotide bias of small RNAs co-purified with MapSPARTA (**E**) or extracted from *E. coli* (**F**). See also **Sup. Table 1**.

Short pAgos trigger cell death upon detection of invading DNA

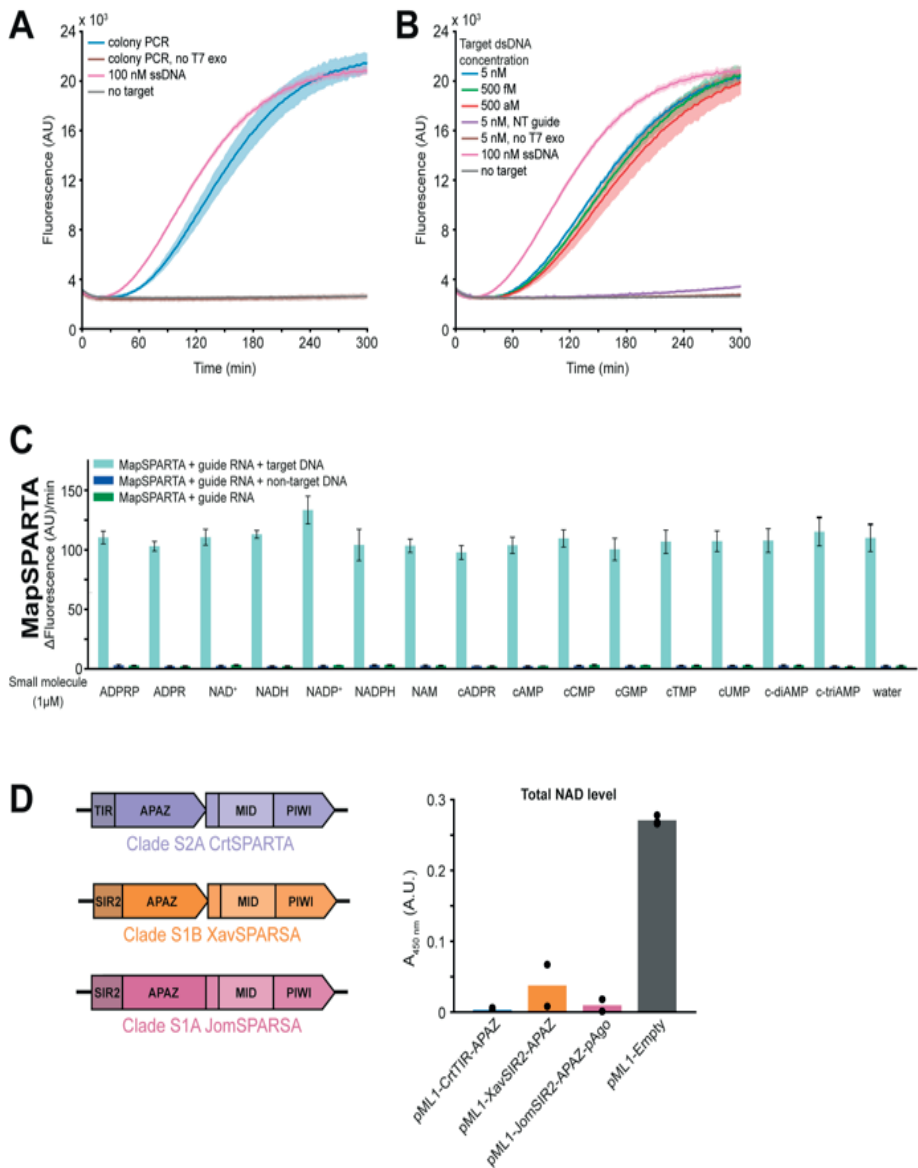


Supplementary Figure S5. Plasmid DNA triggers SPARTA-mediated NAD(P) depletion and cell death, related to Fig. 5. A) Plasmid DNA triggers SPARTA-mediated NAD(P) depletion and cell death. NAD and NADP levels and viability were determined in *E. coli* cultures expressing MapSPARTA or MapSPARTA^{TIR-E77A} in the presence or absence of the high copy pUC-empty plasmid. Panels showing NAD levels at t=4 and cell viability at t=20 are identical to those in Fig. 5A. **B)** Native RNA polymerase-mediated transcription of plasmid-encoded genes enhances SPARTA-mediated cell death. Cell viability was determined in *E. coli* cultures expressing MapSPARTA or MapSPARTA^{TIR-E77A}, either in absence of plasmid DNA or in presence of a pUC-derivate from which mRFP is transcribed but not translated (pUC-mRFP^{XylS-Pm}), or encoded but not transcribed (pUC-mRFP^{XylS-ΔPm}). The Pm promoter induces transcription by native RNA polymerases. Graphs show the average of three biological replicates. Error bars indicate standard deviations. N.s. = not significant; ** = p<.01; *** = p<.001.



Supplementary Figure S6. Effect of SPARTA on plasmid DNA and bacteriophages, related to Fig. 6. **A)** *E. coli* cells expressing MapSPARTA, MapSPARTA^{TIR-E77A} or harbouring an empty BAC were transformed with pUC-empty or pUC-mRFP^{XylS-Pm} and the transformation efficiency was calculated. The average of three biological replicates is shown, error bars indicate standard deviations. **B)** *E. coli* cultures harbouring pUC-empty and expressing MapSPARTA, MapSPARTA^{TIR-E77A} or harbouring an empty BAC were grown without selective pressure (Ampicillin) to maintain pUC-empty. The fraction of cells containing pUC-empty was determined on a daily basis for nine consecutive days. The average of four biological replicates is shown, error bars indicate standard deviations. **C)** *E. coli* cultures expressing MapSPARTA or MapSPARTA^{TIR-E77A}, or harbouring an empty BAC were grown in agar plates in the presence of various bacteriophages and the number of plaques relative to the strain harbouring an empty BAC was determined. The average of three biological replicates is shown, data points indicate results from individual replicates. **D)** *E. coli* cultures expressing MapSPARTA or MapSPARTA^{TIR-E77A}, or harbouring an empty BAC were grown in the presence of various bacteriophages (bacteriophage

[Sup. Fig. S6]...type indicated above each graph) and the OD_{600 nm} was monitored over time. Bacteriophages were added to *E. coli* cells with the following multiplicity of infection: T1 – 1.4×10^{-7} ; T4 – 2×10^{-7} ; T7 – 1×10^{-7} ; Lambda – 3×10^{-7} ; Nami – 1×10^{-7} . The average of three biological replicates is shown, shadings indicate standard deviations. **E)** *E. coli* JM109(DE3) cultures expressing MapSPARTA or MapSPARTA^{TIR-E77A}, or harbouring an empty vector were grown in the presence of chronic-infecting phage M13. The number of infectious M13 phage virions excreted into the bacterial culture was determined at multiple timepoints using *E. coli* ER2738, in which phage M13 is lytic. The average of three biological replicates is shown, error bars indicate standard deviations.



Supplementary Figure S7. SPARTA can be used to detect dsDNA and SPARSA systems degrade NAD⁺ *in vivo*, related to Fig. 7. A) SPARTA combined with colony PCR and subsequent T7 exonuclease treatment facilitates detection of dsDNA. MapSPARTA was mixed with guide RNA in a 1:1 molar ratio and colony PCR-amplified DNA treated with T7 exonuclease was added. After addition of ϵ -NAD⁺, total fluorescence was measured over time. No target negative control and 100 nM ssDNA positive control are identical to those in **Fig. 7F. B)** SPARTA mixed with target

[Sup. Fig. S7]... dsDNA and T7 exonuclease simultaneously facilitates detection of dsDNA at aM levels. MapSPARTA was mixed with guide RNA in a 1:1 molar ratio and target DNA and T7 exonuclease were added. After addition of ϵ -NAD⁺, total fluorescence was measured over time. No target negative control and 100 nM ssDNA positive control are identical to those in **Fig. 7F**. **C)** MapSPARTA is not activated by small molecules, and small molecules do not enhance guide RNA/target ssDNA-mediated activity of MapSPARTA. MapSPARTA was mixed with guide and target oligonucleotides in a 1:1:2 molar ratio and incubated with ϵ -NAD⁺. Graphs indicate the maximum change in fluorescence over time. **D)** Left: Schematic diagram of the operon structure and domain organization of SPARTA, SPARSA-1A, and SPARSA-1B systems. Right: SPARSA SIR2-APAZ proteins lower total NAD when heterologously expressed in *E. coli*. Total NAD levels were determined in *E. coli* cultures harbouring and empty expression vector pML1 or expressing CrtTIR-APAZ, clade S1A *Joostella marina* JomSIR2-APAZ-Ago or clade S1B *Xanthomonas vesicatoria* XavSIR2-APAZ. Graphs show the average of three (**A & C**) or two (**B**) technical replicates, or two biological replicates (**D**). Error bars and shadings indicate standard deviations.



6.

Phages

Nami and Shinka: isolation, characterization, and sequencing

In preparation: *Complete genome of two novel phages infecting UroPathogenic Escherichia coli (UPEC).* Cristian Aparicio-Maldonado, Franklin L. Nobrega, and Stan J.J. Brouns

6

Abstract

Myophages vB_EcoM_Nami and vB_EcoM_Shinka were isolated against a clinical uropathogenic *Escherichia coli* from sewage water. The complete double-stranded DNA genomes of 165,679 and 162,160 bp, respectively, were annotated and classified as *Tequatrovirus*. Both phages have a relatively broad host range in *E. coli* strains, and complete one life cycle in approximately 30 minutes. Phage Nami is of particular interest for phage therapy applications, as it seems to utilize a mechanism of infection that prevents the quick development of phage resistance in infected *E. coli*. Evolution experiments with phage Shinka further showcase the ability to evolve phages to overcome bacterial mutants insensitive to the wild-type phages. Overall, this work provides a morphological, phenotypic, and genomic characterization of two novel phages with features of interest for phage-based control of antibiotic-resistant *E. coli*.

6.1. Introduction

Escherichia coli is, predominantly, a commensal bacterial species of mammals and birds, but it is also found in other distinct ecological habitats. It includes some pathogenic subgroups such as enteropathogenic *E. coli* (EPEC), meningitis-associated *E. coli* (MNEC), and uropathogenic *E. coli* (UPEC), amongst others ¹⁻³, which makes it a clinically relevant species. In fact, *E. coli* was recently listed by the World Health Organization as one of the bacteria that pose the greatest threat to human health, and a priority pathogen for the development of new antimicrobial solutions ⁴.

In recent years, the use of bacteriophages – viruses of bacteria – to control bacterial pathogens has become a compelling alternative or complementary treatment to antibiotics ^{5,6}, and their use has proven successful in different occasions ^{7,8}. Here, we isolated two bacteriophages infecting the cefotaxime-resistant clinical *E. coli* strain R10256, isolated at Erasmus Medical Center (Rotterdam, Netherlands), and present their complete morphological, phenotypic and genotypic characterization.

6.2. Results

6.2.1. Morphological and phenotypic characterization of phages Nami and Shinka

- **Phage and plaque morphology**

Bacteriophages vB_EcoM_Nami (Nami) and vB_EcoM_Shinka (Shinka) were isolated from a wastewater treatment plant in Amsterdam (Netherlands). The purified phage samples were imaged using transmission electron microscopy (TEM) and both Nami and Shinka phages showed a clear head-tail structure (**Fig. 1a-b**) and therefore belong to the *Caudovirales* order of phages with double-stranded DNA. More specifically, they have the typical myovirus morphology with a long, contractile tail.

Phages Nami and Shinka have a small-medium plaque size of 0.3-1mm (**Fig. 1c-d**), similar to phage T4 ⁹.

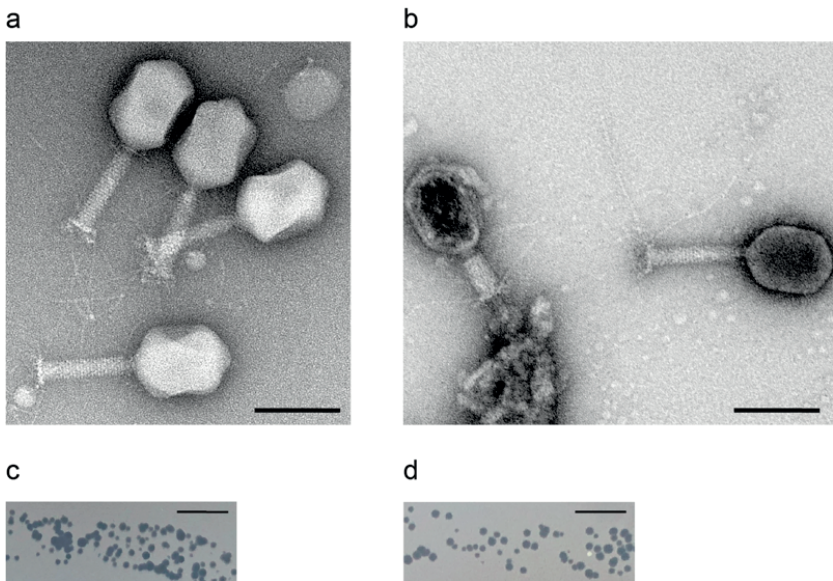


Figure 1. Morphological features of phages Nami and Shinka. a-b) Particle morphology of phages (a) Nami and (b) Shinka as assessed by transmission electron microscopy. Scale bar of 100 nm. c-d) Plaque morphology of phages (c) Nami and (d) Shinka. Scale bar of 5 mm.

• **One-step growth curve**

To characterize the replication cycle of the phages, we performed a one-step-growth curve experiment (Fig. 2). The one step growth curves show similar infection patterns, with latency periods of approximately 20 min, but burst sizes of 83 PFU/cell for Nami (Fig. 2a) and 53 PFU/cell for Shinka (Fig. 2b).

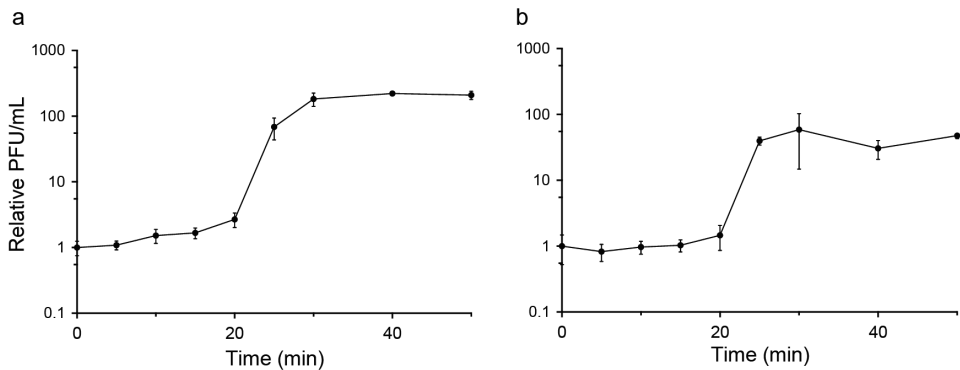


Figure 2. One-step growth curves of phages a) Nami and b) Shinka. The one-step growth curve was determined in *E. coli* strain R10256. Error bars indicate standard deviation of three individual replicates.

• **Host range**

To characterize the infection profile of phages Nami and Shinka, we tested their infectivity against the non-prophage-containing strains of the ECOR collection ^{10,11} and common lab strains. Both phages infected a range of strains (**Table 1**), displaying slightly different lytic spectra. The host range was broader for Shinka (17 of 56) than for Nami (14 of 56) on the ECOR collection. In addition, they produced lysis from without ¹² in 13 (Nami) and 10 (Shinka) of the ECOR strains.

Table 1. Host range of phages Nami and Shinka.

Nami-infected strains	ECOR: 56, 65
Shinka-infected strains	ECOR: 16, 38, 39, 40, 49,
Nami- and Shinka-infected strains	ECOR: 4, 11, 13, 19, 24, 33, 47, 51, 54, 55, 57, 63 Lab strains: BL21-AI, DH5 α , K12-JM109, K12-W1485
Non-infected strains	ECOR: 1, 2, 3, 5, 6, 7, 10, 12, 15, 17, 20, 21, 22, 25, 26, 27, 29, 30, 31, 32, 34, 35, 36, 37, 41, 42, 43, 50, 52, 59, 60, 62, 66, 68, 69, 71, 72

6.2.2. Genomic characterization of phages Shinka and Nami

Phages Nami (accession no. MZ502380) and Shinka (accession no. MZ502379) have double stranded DNA genomes of 165,679 bp and 162,160 bp, with 264 and 250 open reading frames (ORFs), respectively. Both phages have a tRNA cluster containing 10 tRNA coding sequences for amino acids Gln, Leu, Gly, Pro, Ser, Thr, Met, Tyr, Asn, and Arg. The GC content of the phages (35.5% for Nami and 35.4% for Shinka) is lower than the median GC content of *E. coli* (50.8%), a divergence that has been observed previously for phages infecting different species ^{13–15}.

Phage Nami is similar to phage EcNP1 (NC_054910.1), with a genome identity of 99.8 %, and a total of 247 genes with over 95 % protein identity. The 0.2% difference between the two phages corresponds to 17 genes with less of 95 % protein identify, including 10 of them with less than 70 % identity. These last ones include 4 genes coding for structural proteins (tail and capsid components), 4 metabolism-related genes (Mhr transcriptional regulator, polynucleotide kinase, ribonucleotide reductase, and alpha-glucosyl transferase) and 2

hypothetical proteins of unknown function. The closest homologue to phage Shinka is Shigella phage JK23 (MK962752.1), with a 98.1 % genome identity, and a total of 194 genes over 95 % protein identity. This difference is due to a total of 70 genes with less of 95 % protein identify, including 20 with less than 70 % identity. These last ones include 7 genes coding for structural proteins (baseplate, tail and capsid components), 5 metabolism-related genes (NDP reductase subunits, alpha-glucosyl transferase, inhibitor of host Lon protease) and 8 hypothetical proteins of unknown function. Whole genome alignment analysis demonstrated a sequence identity of 96.1 % between phages Nami and Shinka.

A phylogenetic analysis of the genomes of phages Nami and Shinka with their 10 closest relatives in the refseq database (**Fig. 3**) reveals their clustering in pairs with the closest relatives EcNP1 and CM8, respectively. The phylogenetic tree further demonstrated that Nami and Shinka cluster in a larger group with phages vB_EcoM_OR5505, fPS-65, HY01 and UFV-AREG1. The proximity of this group to *Escherichia* phage T4, indicates that phages Nami and Shinka can be classified as *Tequatrovirus*.

6 Closer inspection of the genomes of phage Nami and Shinka reveal features common to T4-like phages, such as short and long tail fiber genes that are involved in host recognition, and the presence of a set of genes that code for factors involved in the subversion of the host RNA polymerase for phage transcription. These include homologues of T4's Alc transcription terminator, Alt, AsiA and MotA proteins, and the gp33 and gp55 sigma factor subunits (reviewed in ¹⁶). Also similarly to phage T4, Nami and Shinka encode genes responsible for the glucosyl hydroxymethylation of cytosines of the phage genome (dCMP hydroxymethylase and glucosyltransferase), an epigenetic modification that is known to protect the phage DNA from degradation by bacterial defenses such as restriction-modification systems (R-M) ¹⁷ and some CRISPR-Cas subtypes (type I and type II) ¹⁸.

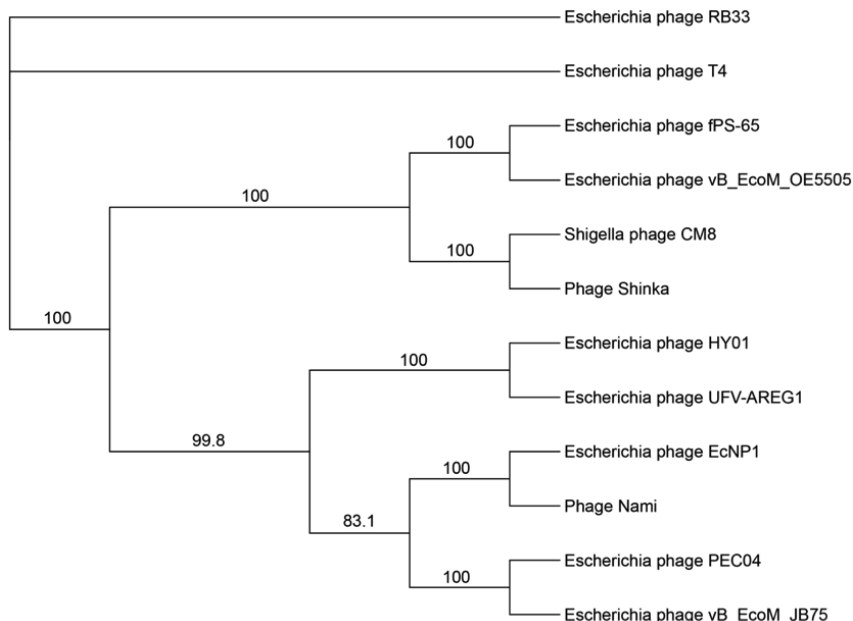


Figure 3. Phylogenetic tree of phages Shinka and Nami. Consensus tree of 1000 replicates generated with the Geneious Tree Builder, using the genetic distance model Tamura-Nei by the method Neighbor-Joining. Numbers indicate bootstrapping values.

6.2.3. Evolution of resistance to phages Nami and Shinka and their counter-acting responses

- **Phage Nami avoids development of resistance in lab strains of *E. coli***

It has been shown that exposure of bacteria to phages can lead to the appearance of phage resistant mutants in both the laboratory and clinical settings, which can complicate phage-based treatments¹⁹. To gather insights on the development of resistance to phages Nami and Shinka and the capacity of these phages to overcome resistance, we analyzed the development of bacteriophage insensitive mutants (BIMs) by *E. coli* strains infected with these phages. We evaluated bacterial growth in the absence and presence of phages Nami and Shinka over a course of 46 hours (**Fig. 4**), and observed a typical decrease upon phage infection, and a recovery of bacterial growth after approximately 16 hours for both phages in two *E. coli* strains (K12 and BL21-AI).

As this suggests the development of phage resistance, we diluted and spotted overnight grown cultures of infected *E. coli* and tested individual colonies for phage sensitivity using liquid cultures. For Shinka, 5.3% of the colonies (15 of 281) tested resistant to phage infection, but no resistance was observed for phage Nami, as all colonies still tested sensitive to phage infection. This surprising absence of resistance to phage Nami may be related to the uncommon fluctuation observed in the bacterial density in *E. coli* cultures infected with this phage. This fluctuation seems to suggest repetitive cycles of culture lysis and recovery, supporting the absence of phage resistance observed for the tested colonies.

This unique adaptive behavior might represent a scenario in which the phage has evolved to sense the host cell density or the phage-to-cell ratio and adjust its rate of infection to prevent extinction of the cell population (and therefore of a host in which to replicate), and/or to avoid evolution of phage resistance. Lysogenic phages have been shown to use different strategies to sense phage and host populations to make a lysis or lysogeny decision^{20,21}, so it is possible that virulent phages have adapted such strategies to regulate their rate of infection.

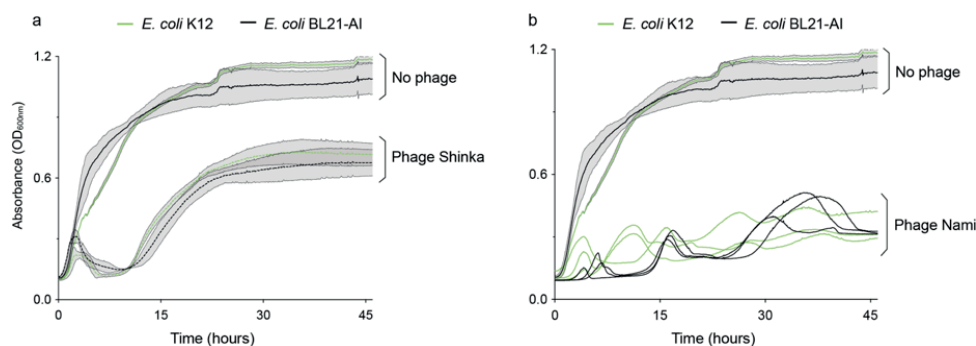


Figure 4. Bacterial growth profiles in absence or presence of phages Shinka and Nami. a) Growth of *E. coli* strains K12 and BL21-AI in the absence or in presence of phage Shinka. Initial multiplicity of infection (MOI): 10^{-5} . Filled areas inside dotted lines indicate standard deviation of three independent replicates. **b)** Growth of *E. coli* strains K12 and BL21-AI in absence and presence of phage Nami. Filled areas inside dotted lines indicate standard deviation of three independent replicates. Each line for infected cultures represents one individual replicate.

The ability of phage Nami to infect *E. coli* strains without the development of resistance is a desirable feature in phage therapy applications. However, the fluctuation behavior of infected cultures was observed in lab strains, but not in the clinical host strain (**Sup. Fig. 1**). It is possible that this strain codes for a counter-mechanism to phage Nami's strategy to prevent phage resistance, and it remains to be seen if such mechanism is shared among more pathogenic clinical strains.

• Phage Shinka can evolve to infect BIMs

Contrary to phage Nami, phage Shinka induced the development of phage resistance in *E. coli*. To determine if the phage is able to evolve to overcome phage resistance, we performed multiple sequential rounds of infection of mixed (1:1) cultures of the wildtype (WT) host strain and BIMs. After ten passages, we observed the appearance of a mutant Shinka phage (MutShinka) able to infect one BIM out of seven tested. Phage MutShinka has a 10-fold increased infectivity in the BIM compared to the original host strain, indicating its adaptation to the BIM host.

We then sequenced MutShinka to identify mutations responsible for the ability to infect the BIM. A total of eight mutations (**Table 2**) were found: One silent mutation in gene *gp58*, a recombination-related endonuclease; four mutations in genes coding for tail fiber proteins Gp146 (Y465H and S466F) and Gp233 (Q252H and D444N); and three mutations in gene *gp211*, a hypothetical protein, which change the C-terminal domain of the protein (PAEMLYHR → PTVNAISSLIL).

The mutations on tail fibers are likely related to changes in the phage-bacteria surface interactions, as previously described for other phages^{22,23}. Because Shinka is related to T4, it likely interacts with lipopolysaccharides (LPS) and/or the OmpC protein²⁴. The mutant bacteria might have changed these receptors, driving phage MutShinka to mutate to adapt to these changes. The mutations occurring in Gp211 might be related to transcription, as the homologue of this protein is phage T4 Alc, which is responsible for the host

transcription shutoff^{25,26}. While the WT version of Alc only allows transcription of DNA containing hydroxymethyl cytosine (hmCyt, phage DNA), mutant versions were found to allow transcription of non-modified cytosine DNA²⁵, but it is unclear if this is the case for the mutant version of Alc in MutShinka.

In summary, phage MutShinka acquired both surface-related and transcription-related mutations to adapt to infection of a bacterial mutant resistant to the original phage.

Table 2. Summary of mutations acquired by phage MutShinka compared to WT Shinka. Mutations indicated in bold for the nucleotide sequence. nt, nucleotide; pt, protein; SM, silent mutation; FS, frameshift; RS, residue substitution.

Gene	nt position displayed	Original nt seq.	Mutated nt sequence	Original pt seq.	Mutated pt seq.	Effect
gp58	169-171	ACG	ACA	T	T	SM
gp146	1393-1397	TATTCC	CATTTC	YS	HF	RS
gp211	427-end	CCGGCTGAAATGCTA TATCATCGTTAA	CCAAGTGTAAATGCTATA TCATCGTTAATCCTTTAA	PAEMLYHR	PTVNAISSLIL	FS
gp233	754-756	CAA	CAC	Q	H	RS
gp233	1330-1332	GAC	AAC	D	N	RS

6.3. Conclusions

Herein, we have isolated and characterized two new coliphages Nami and Shinka, which belong to the *Tequatrovirus* genus. They were both isolated for the clinical uropathogenic *E. coli* isolate R10256, showing ability to also infect a variety of non-pathogenic *E. coli* strains.

Phage Nami revealed important features for phage therapy applications, as bacteria were unable to evolve and avoid infection by this phage. For phage Shinka, we observed the appearance of BIMs, but also that phage variants could be quickly obtained that are able to overcome and infect some of the BIMs. Mutations acquired by the mutant phage MutShinka for infection of BIMs were present mainly in the tail fibers, but also in a transcription-related gene, suggesting that the BIM's resistance to the phage involved surface adaptations and also internal mechanisms related to host transcription takeover.

Importantly, phage Shinka could not overcome resistance of other BIM mutants, indicating a diversity of mechanisms used by *E. coli* to prevent infection by this phage.

In conclusion, it is remarkable how the adaptation profiles of both phages were different, when their genome identity is above 96%. This highlights the importance of studying phage-bacteria interactions in detail, as small differences can lead to very distinct outcomes when exposing bacteria to phages. In this case, one of the phages would lead to a clear resistance phenotype, while the other would not lead to this phenotype but it would also not lyse the bacterial population.

6.4. Material and methods

6.4.1. Bacterial strains and culturing

E. coli strains of the ECOR collection, clinical strain R10256 (Erasmus Medical Center, Rotterdam, Netherlands), and laboratory strains BL21-AI, DH5 α , K12-JM109, and K12-W1485, were grown in Lysogeny broth (LB) at 37 °C, with shaking at 180 rpm. For solid media assays, LB was supplemented with either 1.5% (w/v) agar (LBA, bottom layer) or 0.7 % (w/v) agar (top layer).

6.4.2. Phage isolation, culturing and titering

Phages were isolated from a wastewater treatment plant in Amsterdam (Netherlands). A ten-liter sample was filtered (0.22 μ m pore-size) using a Nalgene® vacuum system and concentrated to 500 mL using a VivaFlow 50® system. An *E. coli* R10256 starting culture was infected with 1 mL of concentrated sample and incubated overnight at 37 °C with shaking. Next day, the culture was centrifuged (4000 $\times g$, 10 min, 4 °C) and filtered, and diluted aliquots were plated on an *E. coli* R10256 lawn using the double-layer agar method, as previously described ²⁷. Plaques were individually picked and passaged multiple times in lawns of *E. coli* R10256 to obtain purified phages.

The purified phages were produced in liquid cultures of *E. coli* R10256, and tittered by spotting 10-fold serial dilutions in double-layer agar plates. After overnight incubation at 37 °C, plaque forming units (PFU) were counted and the titer estimated.

6.4.3. Transmission electron microscopy

For sample preparation, Formvar/Carbon grids with a mesh size of 400 (TED PELLA) were used. The grids were glow discharged using the easiGlow™ discharge system (PELCO) at 15 mA for 60 s in a 0.39 mbar atmosphere. An aliquot of 5 μ L of the sample was incubated for 1 minute, followed by 3 times washing with 10 μ L ultrapure water, and blotting with filter paper. For staining, 5 μ L of 1% NanoVan® (Nanoprobes) was applied and blotted away after 15

seconds. Image acquisition was performed using a Jeol JEM-1400 TEM with an acceleration voltage of 120 kV.

6.4.4. Host range analysis

Ten-fold serial dilutions of the phages were spotted onto double layer agar plates of non-prophage containing strains of the ECOR collection, as well as *E. coli* strains BL21-AI, DH5 α , K12-JM109, and K12-W1485. The plates were incubated overnight at 37 °C, and the phage plaques were observed to distinguish productive infection (lysis with individual phage plaques formed) and lysis from without (lysis without individual phage plaques).

6.4.5. One-step growth curve

Phages were added to 50 mL exponentially growing bacteria (OD_{600 nm} of 0.5) at an MOI of 0.01. Cultures were incubated at 37 °C with shaking at 180 rpm, and 200 μ L samples were taken periodically for 50 min and tittered as detailed above.

6.4.6. DNA extraction, sequencing and genome analysis

Phage DNA was extracted using a SDS-proteinase K protocol followed by organic-phase phenol-chloroform extraction as previously described ²⁸. Library preparation and sequencing were performed by Novogene (Hong Kong, China) using the Illumina HiSeq PE150 platform. Phages Nami and Shinka generated 8,958,504 and 7,845,994 raw reads, respectively. For quality control processing, these reads were trimmed, paired, and error-corrected, generating 264,800 and 39,976 reads respectively, with a corresponding coverage of 240x for Nami and 37x for Shinka. A random read sampling (for a 15x theory coverage) was used for *de novo* assembly using SPAdes ²⁹, resulting on a single contig for each phage. Annotation and open reading frame (ORF) prediction were performed using the RAST server ^{30–32} followed by manual curation using BLASTp ³³, HHPrep ³⁴ and Phyre2 ³⁵. Prediction of tRNAs was performed with tRNAscan-SE version 2.0 ³⁶. All bioinformatics tools were used with default parameters, and HHpred

alignment performed for databases PDB_mmCIF70, Pfam-A and NCBI_Conserved_Domains. For phylogenetic tree construction, selected genomes were downloaded from refseq and aligned using ClustalW 2.1. A Neighbour-Joining tree using the Tamura-Nei model was built using Geneious Tree Builder (Geneious® 11.1.5), with a bootstrap of 1000 replicates.

6.4.7. Phage infection growth curves

Bacterial cultures were grown until late exponential phase ($OD_{600\text{ nm}}$ of 1.0-1.2), normalized to $OD_{600\text{ nm}}$ of 0.5 (approximately 1×10^8 CFU/ml) and aliquoted in 96-well microtiter plates. Next, 10 μ l of phage suspension at an MOI of 10^{-5} were added to the wells, and bacterial growth was followed in an EPOCH2 microplate reader with $OD_{600\text{ nm}}$ measurements at 37 °C, every 5 min, with constant shaking at 270 rpm.

6.4.8. Development of phage resistance

Bacterial cultures of clinical strain R10256 grown until late exponential phase were used to inoculate 50 mL of LB to an initial $OD_{600\text{ nm}}$ of 0.05, then incubated at 37 °C with constant shaking at 180 rpm. At exponential growth phase ($OD_{600\text{ nm}}$ of 0.3 to 0.5), phage was added at an initial MOI of 1, incubated in same conditions. Following O/N incubation, 100 μ L of 10-fold serial dilutions of the culture were plated on LB plates to determine the BIM frequency. To assess the phage resistance phenotype, each of the selected colonies were grown until exponential phase and used to test phage infection by double-layer agar method.

6.4.9. Phage evolution

To study if phage Shinka could evolve to infect a BIM, a total of 10 co-culture passages were performed as described next. Cultures in steady-state growth phase, of the clinical strain R10256 (WT) and each of the obtained BIMs, were normalized in LB medium to an $OD_{600\text{ nm}}$ of 0.05 and combined in a 1:1 ratio in a total volume of 50 mL. After incubation at 37 °C with constant shaking at 180 rpm, when the $OD_{600\text{ nm}}$ was aprox. 0.3, phage Shinka was added at an initial

MOI of 10^{-5} , followed by 20h of incubation in the same conditions. Then, 1 mL of culture was centrifuged at $10.000 \times g$ for 1 min at 4°C and the supernatant tested for plaque formation on the WT and BIM strains using the double-layer agar method, as indicated previously. In the case of finding an evolved phage infecting the BIM strain, the supernatant was used to produce phage in the BIM strain and isolate individual plaques for further testing and characterization.

6.5. Data availability

Genome sequences have been deposited in Genbank under the accession numbers MZ502380 (vB_EcoM_Nami) and MZ502379 (vB_EcoM_Shinka). Raw sequencing data have been deposited in the European Nucleotide Archive (ENA) under accession number PRJEB46561.

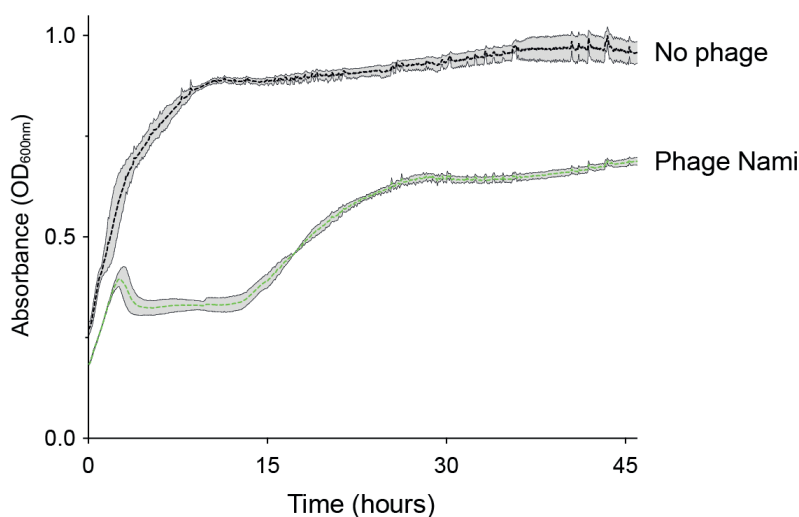
6.6. References

1. Fratamico,P.M., DebRoy,C., Liu,Y., Needleman,D.S., Baranzoni,G.M. and Feng,P. (2016) Advances in molecular serotyping and subtyping of *Escherichia coli*. *Front. Microbiol.*, 7.
2. Touchon,M., Perrin,A., De Sousa,J.A.M., Vangchhia,B., Burn,S., O'Brien,C.L., Denamur,E., Gordon,D. and Rocha,E.P.C. (2020) Phylogenetic background and habitat drive the genetic diversification of *Escherichia coli*.
3. Kaper,J.B., Nataro,J.P. and Mobley,H.L.T. (2004) Pathogenic *Escherichia coli*. *Nat. Rev. Microbiol.*, 2.
4. Tacconelli,E., Magrini,N. and World Health organization (2018) Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics.
5. Wittebole,X., De Roock,S. and Opal,S.M. (2013) A historical overview of bacteriophage therapy as an alternative to antibiotics for the treatment of bacterial pathogens. *Virulence*, 5.
6. Moelling,K., Broecker,F. and Willy,C. (2018) A Wake-Up Call: We Need Phage Therapy Now. *Viruses*, 10.
7. Schooley,R.T., Biswas,B., Gill,J.J., Hernandez-Morales,A., Lancaster,J., Lessor,L., Barr,J.J., Reed,S.L., Rohwer,F., Benler,S., et al. (2017) Development and use of personalized bacteriophage-based therapeutic cocktails to treat a patient with a disseminated resistant *Acinetobacter baumannii* infection. *Antimicrob. Agents Chemother.*, 61.
8. Dedrick,R.M., Guerrero-Bustamante,C.A., Garlena,R.A., Russell,D.A., Ford,K., Harris,K., Gilmour,K.C., Soothill,J., Jacobs-Sera,D., Schooley,R.T., et al. (2019) Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant *Mycobacterium abscessus*. *Nat. Med.*, 25.
9. Chen,Y. and Young,R. (2016) The last *r* locus unveiled: T4 RIII is a cytoplasmic antiholin. *J. Bacteriol.*, 198.
10. Ochman,H. and Selander,R.K. (1984) Standard reference strains of *Escherichia coli* from natural populations. *J. Bacteriol.*, 157.
11. Shibata,Y., Ugumori,C., Takahashi,A., Sekoguchi,A. and Maeda,S. (2014) Survey of lysogenic phages in the 72 strains of *Escherichia coli* collection of reference (ECOR) and identification of a phage derived from the ECOR52 strain. *Am. J. Biosci.*, 2.
12. Abedon,S.T. (2011) Lysis from without. *Bacteriophage*, 1.
13. Marinelli,L.J., Fitz-Gibbon,S., Hayes,C., Bowman,C., Inkeles,M., Loncaric,A., Russell,D.A., Jacobs-Sera,D., Cokus,S., Pellegrini,M., et al. (2012) *Propionibacterium acnes* bacteriophages display limited genetic diversity and broad killing activity against bacterial skin isolates. *MBio*, 3.
14. Pan,Y., Lin,T., Chen,C. and Tsai,Y. (2017) *Klebsiella* Phage Φ K64-1 Encodes Multiple Depolymerases for Multiple Host Capsular Types. *J. Virol.*, 91.

15. Dupuis,M.È. and Moineau,S. (2010) Genome organization and characterization of the virulent lactococcal phage 1358 and its similarities to *Listeria* phages. *Appl. Environ. Microbiol.*, 76.
16. Hinton,D.M. (2010) Transcriptional control in the prereplicative phase of T4 development. *Viol. J.* 71.
17. Lehman,I.R. and Pratt,E.A. (1960) On the structure of the glucosylated hydroxymethylcytosine nucleotides of coliphages T2, T4, and T6. *J. Biol. Chem.*, 235.
18. Vlot,M., Houkes,J., Lochs,S.J.A., Swarts,D.C., Zheng,P., Kunne,T., Mohanraju,P., Anders,C., Jinek,M., Van Der Oost,J., et al. (2018) Bacteriophage DNA glucosylation impairs target DNA binding by type I and II but not by type V CRISPR–Cas effector complexes. *Nucleic Acids Res.*, 46.
19. Egido,J.E., Costa,A.R., Aparicio-Maldonado,C., Haas,P.-J. and Brouns,S.J.J. (2021) Mechanisms and clinical importance of bacteriophage resistance. *FEMS Microbiol. Rev.*, fuab048.
20. Erez,Z., Steinberger-Levy,I., Shamir,M., Doron,S., Stokar-Avihail,A., Peleg,Y., Melamed,S., Leavitt,A., Savidor,A., Albeck,S., et al. (2017) Communication between viruses guides lysis-lysogeny decisions. *Nature*, 541.
21. Oppenheim,A.B., Kobiler,O., Stavans,J., Court,D.L. and Adhya,S. (2005) Switches in Bacteriophage Lambda Development. *Annu. Rev. Genet.*, 39.
22. Chapman-McQuiston,E. and Wu,X.L. (2008) Stochastic receptor expression allows sensitive bacteria to evade phage attack. Part II: Theoretical analyses. *Biophys. J.*, 94.
23. Chapman-McQuiston,E. and Wu,X.L. (2008) Stochastic receptor expression allows sensitive bacteria to evade phage attack. Part I: Experiments. *Biophys. J.*, 94.
24. Washizaki,A., Yonesaki,T. and Otsuka,Y. (2016) Characterization of the interactions between *Escherichia coli* receptors, LPS and OmpC, and bacteriophage T4 long tail fibers. *Microbiologyopen*, 5.
25. Kashlev,M., Nudler,E., Goldfarb,A., White,T. and Kutter,E. (1993) Bacteriophage T4 Alc protein: A transcription termination factor sensing local modification of DNA. *Cell*, 75.
26. Kutter,E.M., Bradley,D., Schenck,R., Guttman,B.S. and Laiken,R. (1981) Bacteriophage T4 alc gene product: general inhibitor of transcription from cytosine-containing DNA. *J. Virol.*, 40.
27. Adams,M.H. (1959) *Bacteriophages* New York, Interscience Publishers.
28. Clokie,M.R.J. and Kropinski,A.M. (2009) *Bacteriophages. Methods and Protocols Volume 1: Isolation, Characterization, and Interactions.*
29. Prjibelski,A., Antipov,D., Meleshko,D., Lapidus,A. and Korobeynikov,A. (2020) Using SPAdes De Novo Assembler. *Curr. Protoc. Bioinforma.*, 70.
30. Aziz,R.K., Bartels,D., Best,A., DeJongh,M., Disz,T., Edwards,R.A., Formsma,K., Gerdes,S., Glass,E.M., Kubal,M., et al. (2008) The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics*, 9.
31. Overbeek,R., Olson,R., Pusch,G.D., Olsen,G.J., Davis,J.J., Disz,T., Edwards,R.A., Gerdes,S., Parrello,B., Shukla,M., et al. (2014) The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.*, 42.

32. Brettin,T., Davis,J.J., Disz,T., Edwards,R.A., Gerdes,S., Olsen,G.J., Olson,R., Overbeek,R., Parrello,B., Pusch,G.D., et al. (2015) RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.*, 5.
33. Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, 215.
34. Gabler,F., Nam,S.Z., Till,S., Mirdita,M., Steinegger,M., Söding,J., Lupas,A.N. and Alva,V. (2020) Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. *Curr. Protoc. Bioinforma.*, 72.
35. Kelley,L.A., Mezulis,S., Yates,C.M., Wass,M.N. and Sternberg,M.J. (2016) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.*, 10.
36. Chan,P.P. and Lowe,T.M. (2019) tRNAscan-SE: Searching for tRNA genes in genomic sequences. *Methods Mol. Biol.*, 1962.

6.7. Supplementary figure



Supplementary Figure 1. Bacterial growth profiles of clinical *E. coli* strain R10256 in absence or presence of phage Nami. Initial multiplicity of infection (MOI): 10^{-5} . Filled areas inside dotted lines indicate standard deviation of three independent replicates.

Appendix

A stylized illustration of a person in a white suit climbing a white ladder. The ladder is positioned diagonally across the frame. The background is a solid light purple color. A film strip with various colored frames (pink, blue, yellow, green) is draped over the ladder and extends across the background. The person is at the bottom of the ladder, reaching up.



Acknowledgements (Agradecimientos)

Every journey must have a beginning and an end; if not it would not be a journey, but just moving. The PhD journey seems to have lasted for almost 6 years, however it feels to have been going on since long before that. Some people consider it as the culmination of the continuous studies since young age. Others consider it as the beginning of your scientific or professional career. For me, it means closing a book that contains not just the chapters inside it, but also previous others that are not written about.

Therefore, I must acknowledge not just those who contributed to reaching the goal during this period, but all those who made me arrive to the starting line of the journey.

First of all, I would like to thank both of my (co-)promoters: Stan and Franklin. We met almost 6 years ago now, it was time of huge uncertainty for me, and solidifying this PhD position set a huge base for next years to come. **Stan**, in all this long time working together, I have to thank for many things I have learnt, for all the feedback that felt hard but at the end taught a lot, and for all the good times you provided me with, as a member of your group. I hope that you also got some good things from having me there. Bedankt! **Franklin**, what could I say? I feel you are the person which has helped me the most these years, having you next to me daily to provide guidance, not just professionally, but personally too. I have always tried to absorb all the

Todo viaje tiene un principio y un fin; de lo contrario, no sería un viaje sino movimiento. El viaje del Doctorado parece haber durado los último 6 años, pero se siente como que empezó mucho antes de eso. Algunos consideran que es la culminación de los estudios que has hecho desde niño. Otros como el punto inicial de tu carrera científica y profesional. Para mí, significa el cierre de un libro que contiene no solo los capítulos que incluye, sino muchos otros anteriores de los cuales no se ha escrito.

Es por esto que siento la necesidad de agradecer no solo aquellos que contribuyeron a alcanzar este objetivo durante este proceso, sino también aquellos que me hicieron llegar a la línea de salida de este viaje.

A

knowledge I saw on you, as well as disagree with and discuss it in other moments. But overall, what I have enjoyed the most are those hours-long conversation in which we could get lost in words about anything that would cross our mind. Also special thanks for helping me take my first steps after Delft, as well as I am sure you will help in further steps that are ahead to come. As I always say: you have been like my big brother in science, being harsh when needed and supportive when required. Obrigadíssimo!

Next, I must fully thank two friends that I have made during these years in Delft, that have supported me in many aspects during those hard and good times of the last years. **Rita** and **Fabio**, choosing you both as **Paranymphs** was a water-clear decision as soon as it crossed my mind. The irony of “discovering” your country similar to mine but so different, applies also to how you both surprised me with all things we have in common while being different at the same time. At the moment of writing this I am sure that, despite how much support you have already given, the biggest part is yet to come. Muito obrigado, amigos!

Aunque no lo ponga en primer lugar por el tipo de libro que es, no tengo que decir cuantísimo agradezco vuestro apoyo y presencia (a pesar de la distancia), no solo durante estos años, sino desde que me trajisteis a este mundo. **Mamá y Papá**, no sirve ningún extenso párrafo que os escriba aquí para todo lo que debo agradecerlos. Por eso, solo voy a deciros que a pesar de esos malos cambios de humor que he tenido (especialmente) en los últimos meses, jamás se me ha pasado por la cabeza que todo gesto y palabra no fuera con el mejor corazón. No quiero dedicaros muchas líneas aquí, sino el máximo tiempo posible en el futuro. **Carlos**, enanillo, llevamos unos años sin coincidir de forma constante, debido a que cada uno estamos tomando nuestra ruta en la vida. Me entristece pensar que estamos madurando por separado, pero a la vez me alegra que, a pesar de ello, siempre estamos y estaremos lo más cercano posible, para cualquier cosa que necesitemos. Para eso están los hermanos! Tengo la sensación de que es justo ahora cuando tú vas a empezar tu viaje vital más interesante, cuando has dado un salto a una experiencia única en la vida. Disfrútala!

Con la **gran familia** que tengo por suerte, se me hace difícil mencionar aquí a todos y cada uno de los 18 tíos/as, 27 primos/as y los 28 hijos/as de algunos. He de agradecer a todo/as por la gran cantidad de momentos y sentimientos que implican compartir una familia. **Tita Paquí**, tú tienes un lugar especial junto al **Tito Agustín**. Fuiste mi segunda madre y crecí contigo gran cantidad de tiempo desde niño. Gracias por todos esos momentos que me cuidaste cuando mi madre no podía, gracias por esos paseos en bici hasta el colegio, y gracias por todo lo que siempre me has querido. **Tita Marí** y **Tito Benardino**, aunque no estuvisteis en Almagro permanentemente, cada vez que veníais era como si siempre estuvierais, así como vuestra casa fue la mía cuando la necesité en Madrid, gracias. **Encarni** y **Agus**, también crecí con vosotros en casa, fue único el teneros como inspiración y ejemplo. Agus, siempre me vienen a la memoria esos paseos a la piscina, con ambos en la bici y cargados de mochilas. **Domingo**, gracias por todas esas travesuras que compartimos de niño, siempre me sacan una sonrisa. **Bernar** y **David**, aunque la diferencia de edad no lo hizo fácil, en los últimos años agradezco esas buenas charlas que hemos tenido con una birra en mano. **Isa** y **Tito Alberto**, mis padrinos, tengo mucho que agradecerlos como referente en la vida. **Tita Pepa**, gracias por esas meriendas en tu casa y el cariño que recibía. **Isa** y **Oswaldo**, nuestra conexión como primos es especial y agradezco muchísimo teneros cerca. Un gran honor ser el padrino de **Alvaro** y **Daniela**, gracias en especial por esa oportunidad de empezar a mostrarme la importancia de tener responsabilidades en la vida desde temprana edad.

Otra familia algo menos convencional, pero que siempre he disfrutado y aprendido mucho de ellos, es aquel grupo de amigos con los que he crecido desde que tengo memoria. **Consuelo**, **Manolo**, **Laura**, **Marí Carmen**, **Manolo**, **Manolito**, **Marcos**, **Irene**, **Javí**, **Virgínia**, **María**. ¡Que buenos ratos he pasado con vosotros! Cuantas risas, buenas comidas e increíbles momentos. Gracias por arroparme y hacerme sentir como en casa en cada visita a Almagro, y sentir como si nunca hubiera estado viviendo fuera.

Los grupos de amigos cambian constantemente durante la vida, es algo dinámico y no debe verse como algo negativo, pues te da opciones a cambiar y evolucionar como persona. Pero en Almagro, ha sido siempre especial.

A

Teresa, Lorena, Roberto, Miguel Ángel, Ismael, Carlos, y muchísimos otros. Teneos allí fue una gran oportunidad de volver a mis orígenes, recordar quien era y soy, devolverme a la mente de aquel niño con el sueño de ser científico algún día. Vuestras charlas cerveceras me ayudaron a recordarlo y seguir ahí en esos momentos que creía no tener fuerza para continuar.

*Llegar a Toledo fue un salto enorme, y donde debo considerar que mi camino en la ciencia comenzó, compartiendo experiencias con muchas personas, las cuales sería imposible nombrar en su totalidad. **Álvaro, Víctor, Pablo, Chema, Andrés**, gracias por todo. Por todas esas prácticas de laboratorio llenas de risas y pasión. Por esos trabajos (de Macro) en la Resí hasta las mil de la mañana. Por todas esas largas noches antes de los exámenes, encontrándonos en la cocina a comer algo y acabar sin estudiar por horas. Por todas esas fiestas en Dodici en el que a veces tocaba llevar a alguien en brazos de vuelta. Gracias por haber estado y estar aun aquí. **Lorena, Andrea, Patri, María, Mounia**, muchas gracias también por formar parte de esto; no hubiera sido lo mismo sin vosotras compartiendo cada momento. *Gracias Canteros!* And again, thanks **Álvaro** for such nice piece of biography reflecting quite who I am and have been, in such a small amount of lines.*

Querida Yoly, Toledo también me trajo la oportunidad de conocerte y compartir contigo una significativa parte de mi vida. Gracias por crecer conmigo en un importante proceso de madurez. Gracias por haberme abierto los ojos al mundo externo y la gran experiencia de viajar. Y gracias por haber influido en ser quien soy hoy. Nuestro viaje tuvo un fin, pero considero que gracias a ti tuve el coraje de empezar este doctorado lejos de mi familia y amigos de siempre; y así empezar una nueva etapa en la vida. Danke schön!

And reaching this point is when it all switches to English in my life. Equal to other places, Delft brought to me a whole change in life, full of experiences. It is been a time in which not only I did a PhD in Science, but also in life. I learnt new ways of thinking my life, opening my eyes, and living new aspects of it. Many people crossed with me during this period, contributing to all these rich moments. It will be hard to summarize in few lines, or even in many...

During my PhD time, I had an outstanding group of people around me in the Brouns' lab. Such really dynamic group, but with a really strong core. **Becca** and **Jochem**, you always gave me great support as fellows, but also many wonderful plans that helped me understand the country and the culture. Becca, thanks for those moments, the social committee and your guidance through the PhD. Jochem, I fully thank you for introducing me to the world of climbing, we need to catch up with it outdoors soon. I'm looking forward to visit you guys in NZ!! Patrick, we always shared the office and therefore many moments during those years. It was great to have those conversations about random topics that you would bring up. **Seb**, our interaction was somehow distant and close at same time. Thanks for always ended up having a great moment and talk when we enjoyed a nice Belgium beer together. **Tobal, Patri y Zeus**, *fuisteis un hallazgo especial en Holanda y un apoyo enorme para seguir recordando las raíces mientras estuve allí. Gracias por todos esos momentos.* **Rita**, I need to thank you once more at least, not just as paranymp but also as lab and office mate, you always made me remember the importance of seeing the bright side and brought a smile in the bad moments. **Teunke** and **Anna**, you kept the lab running as an essential pillar and allowed to keep doing experiments non-stop, thanks. **Boris** *muchas gracias por eso buenos y diversos momentos y por ese tan rico Ron Cubano que aún conservo y disfruto.* **Ana Gerós, Erik, Jasper** and many other transitory people, hanging out with you was really entertaining and refreshing, helping to keep up with the day to day. Also thanks to my students, **Jeroen, Ilma, Andrea**, you participated on making this thesis possible with your efforts. BN department, thanks to many of you for making the work place enjoyable with all the Borrels and events. Also I would like to thank those lab members of Nobrega's group that have supported me during these last months finishing the PhD from Southampton. **Thomas**, we have known each other for a while and sharing *This life* with you was always joyful, if we were in the gym, in the lab, or outdoors. **Daniela, Maggie**, thanks for those moments outside the lab that helped me disconnect from this stressful life. **Yi Wu**, you remind me to my-early-self, spending all those long days and weekends in the lab. Keep it up, it will pay off. **Jake, Moi, Sasha, Iolanta**, our

interactions was sort of brief, but I deeply appreciate those hours of climbing, those party moments and the nice chats over a beer.

And I arrive to the core of my PhD life, my family during the pandemic. Those that will never split from who I am, regardless of infinity in its meaning. The group The Game has been one of the best things that happened to me, ever. **Fabio, Ruben y Josema**, I really do not know what to say here that you still need to know. Like with my blood family, I would need thousands of lines. Our connection was really strong since we started to meet. But living through such hard times together just made our connection eternal. With you, I opened my mind to other worlds that I never imagined with all those trips, by bike, by train, by bus, or by plane. On them, we asked *why are there boundaries* while observing a *skyline*. Together, *vibin' out* was sometimes a *Risk*, with the impression that *we ain't feeling time*. But we got *blessed* to discover how to *go back home* at any time, from anywhere. Now, with no doubt, I acknowledge that I will *die with a smile*. **Vlad, Ion, Zamo, Manu, Nacho, Peibol, Francческа, Bojana, Illo, Javi, Ines, Lorenz, Nicolò, Luigi, Celía, Juls, Alev**, and many more passing through Delft, much thanks for all the things lived together, all those moments shared and all the beers at TPM and Bouwpub. Thanks for all of it!

Tantos los años de educación, tantos maestros y profesores que han contribuido a enseñarme todo lo que se hoy, y así poder llegar hasta aquí. Imposible incluir aquí a todos, pero agradezco toda aportación recibida. **Paquí, Jose, Choncha, Angel**, gracias por esos años en el **Calvín**. Con vosotros descubrí que la ciencia era mi pasión en la vida. **Aracelí, C. Arribas, Charo, Boiko, Susana**, gracias por enseñarme esos primeros pasos, en clase y en el lab, de lo que la ciencia realmente significa. **Isabel**, tengo la sensación de que sin ti nunca nada hubiera sido posible. Fuiste tú quien me catapultaste a dar un gran paso en la ciencia, con el que forjaría el comienzo de mi futuro. Fuiste tú la que me enseñaste con cuanta pasión es necesario vivir la ciencia. Fuiste tú quien despertó en mí la parte más fuerte de ese problem-solving para pensar y disfrutar sobre cuestiones científicas. Y lo que más agradezco es estar aún en contacto. Aquí, solo diré: ATG ATA CTC GGA CGT GCA TGT ATA GCT TCA TAG !

En mi opinión, algo que también debo agradecer es aquel deporte que hice principal y constantemente durante todos los años que crecí en Almagro; Taekwondo. Esto me dio una gran disciplina y percepción, lo cual considero me ha ayudado en el doctorado, así como en otros aspectos de la vida. Gracias **Rogelio** por contribuir en esta parte de la educación que recibí y me ha hecho en parte quien soy hoy día.

Equally, there have been many lab supervisors that have lead me into the good practice in a lab, as well as the good thinking about scientific problems. He de empezar por dos científicos que admiro; **Alma** y **Roberto**. Nunca recibí supervisión vuestra, pero compartimos el lugar de mi primera estancia de laboratorio; gracias por vuestro apoyo y ánimos a hacerlo. **Ágata** y **Eduardo**; gracias por hospedarme en estas primeras prácticas en el IQOG e introducirme en el mundo de la microbiología y bioquímica. **Rodrigo**; gracias por hospedarme en el HNP en Toledo y enseñarme aspecto de neurociencia y proteínas. **Carolina** e **Isabel**; gracias por hospedarme en Plantas y enseñarme aspectos de patología de plantas y RNA. **Quentin** and **Becky**, thank you very much for hosting me at Oxford University and teach me fundaments of immunology and virology. **Margarita** y **Modesto**; gracias por el tiempo que pase en el CBMO y enseñarme mis primeros pasos en bacteriófagos y bioquímica avanzada. Margarita, donde estés ahora, siempre fuiste una inspiración por tu pasión por la ciencia hasta el último momento. **E. Charpentier**, thank you very much for introducing me to the world of defense systems in the Max Plank Institute at Berlin.

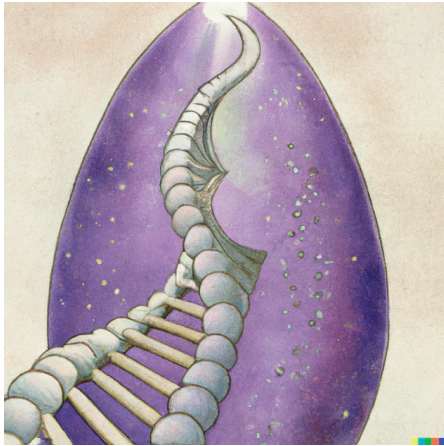
Finally, I would like to thank you, **Jeremy**, as well as **NBIC**, for supporting these last months in the UK, allowing me to take my first steps after the PhD but still before finishing it. Our interaction has been mostly brief, but still has allowed me to see a future and find the energy needed to finish this journey. And equally thanks to **DALLE-2** and their creators; it is been a great way to put the cherry on the top. Amazing experience (that not easy necessarily) having the opportunity to use an A.I. in the making of my PhD thesis book.

A

Original chapter-title images

List of the original A.I.-generated images used for the title page of each chapter

Chapter 1



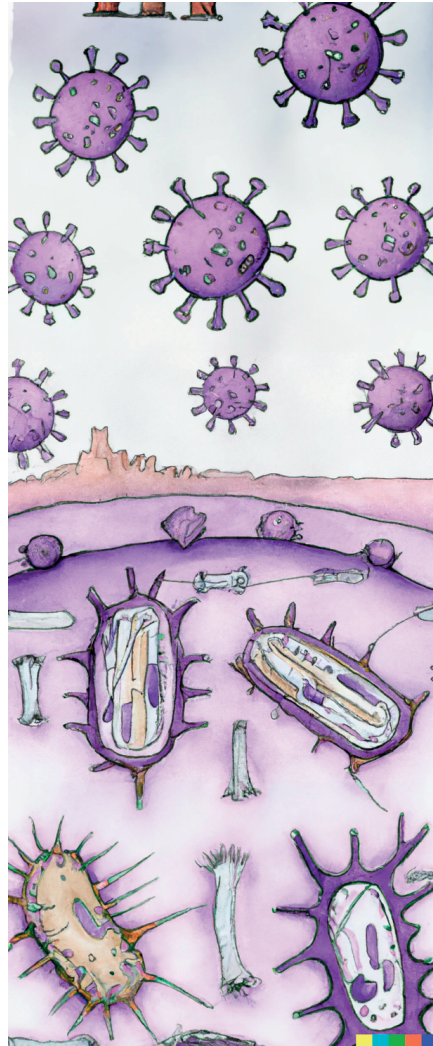
DALL-E interpretation of:
“Ukiyo-e drawing of DNA helix reaching
an end in the universe”

Chapter 3



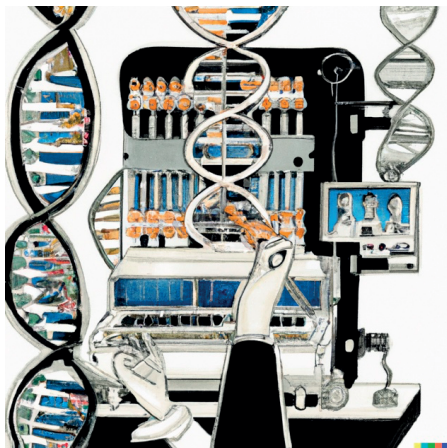
DALL-E interpretation of:
“DNA helix broken apart by hands,
Ukiyo-e”

Chapter 2



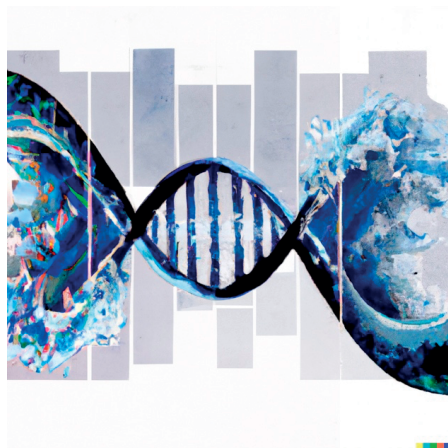
DALL-E interpretation of:
“A pencil and watercolour drawing of the
battlefield of bacteria and viruses”

Chapter 4



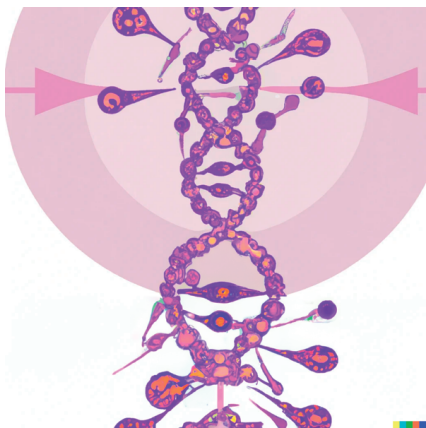
DALL-E interpretation of:
“Machines elaborating DNA
modifications, Luigi Serafini”

Chapter 6



DALL-E interpretation of:
“DNA painting in the style of The Great
Wave of Kanagawa”

Chapter 5



DALL-E interpretation of:
“Symmetric picture of DNA structure
with proteins around, Ukiyo-e”

Appendix



DALL-E interpretation of:
“Abstract scientist working on a huge
DNA helix”

A

About the author

Cristian Aparicio Maldonado was born on a Sunday 25th April 1993, exactly 40 years after the article "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid" was published in the journal *Nature*. In Almagro, Ciudad Real, he lived and grew up sharing his life with his parents, brother and an extended family. Outstanding in whatever he set his mind to, studies or Taekwondo, he finally chose that pure and unadulterated science would be his future.



In 2011 he began his studies in Biochemistry at the Faculty of Environmental Sciences and Biochemistry in Toledo where he continued with his childhood friends and met new ones. He received an undergraduate research scholarship at the Dep. of Plant Genetics after having previously spent time at the Institute of Organic Chemistry (CSIC) and at the National Paraplegic Hospital. He completed his Bachelor's thesis, "Mechanism of HIV vaccine adjuvant activity in NET formation", at the Sir William Dunn School of Pathology (Oxford Univ.) with Prof. Q. Sattentau. Loyal, sardonic and with an intimidating determination, Cristian came through his time at university at the top of his class. Unfortunately for those around him, he never learned to properly sing.

In 2015, he began his studies in the Master of Virology, at the University Complutense of Madrid. Being this topic previously subject of multiple discussions as his main interests within life sciences, he did his Master's Thesis, "Characterization of Translesion Synthesis capacity of DNA polymerases of B family" in the group of Prof. M. Salas (CBMSO). In his eagerness to learn how science is done in cutting-edge laboratories, he did a research internship in the group of Prof. E. Charpentier (Berlin).

With a courageous character and a clear idea of how a person should be trained and valued, he decided to start pursuing a PhD degree at the TU Delft, in the group of Dr. S.J.J. Brouns, at The Netherlands in 2017. Far from home, he knew how to make an adventure of his time there. Passionate and demanding scientist, with a clear philosophy of how scientific goals should be, he embarked on the challenge of understanding bacteriophages, understanding them as a way to combat the antibiotic resistance crisis. He has explored realities that were not his own, surrounding himself with friends who helped him to expand his horizons even further. He has travelled all over Europe, carrying a backpack as suitcase, enjoying life as adventure. This stage has culminated in this thesis, accrediting him with the doctoral degree.

Kindly by: **Álvaro Serrano Navarro**

Sobre el autor

Cristian Aparicio Maldonado nació un domingo 25 de abril de 1993, justo 40 años más tarde de la publicación en la revista Nature del artículo “Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid”. En Almagro, Ciudad Real, vivió y creció compartiendo su vida con sus padres, su hermano y una gran familia. Sobresaliente en lo que se propusiera, en los estudios y en el Taekwondo, eligió que la ciencia pura y sin adulterar sería su futuro.

En 2011 comenzó sus estudios de Bioquímica en la Facultad de Ciencias Ambientales y Bioquímica en Toledo donde continuó con sus amigos de la infancia y conoció nuevos compañeros. Recibió la beca para la investigación como *undergraduate* en el departamento de Genética de plantas, tras previamente haber realizado estancias en el Instituto de Química Orgánica del CSIC y el Hospital Nacional de Paraplégicos. Realizó su TFG, “Mechanism of HIV vaccine adjuvant activity in NET formation”, en la Sir William Dunn School of Pathology (Oxford Univ.), con el Prof. Q. Sattentau. Leal, socarrón y con una determinación que podía resultar intimidante, Cristian superó su etapa universitaria siendo de los mejores de su promoción. Desgraciadamente para los que le rodeaban, nunca aprendió a cantar.

En 2015 comenzó sus estudios en el Máster de Virología, en la Universidad Complutense de Madrid. Siendo este tema previamente objeto de múltiples discusiones y lecturas al tratarse de uno de sus máximos intereses dentro de las ciencias de la vida. realizó su TFM, “Characterization of Translesion Synthesis capacity of DNA polymerases of B family” en el laboratorio de la Prof. M. Salas, en Centro de Biología Molecular Severo Ochoa. En su afán por conocer cómo se hace ciencia en laboratorios de vanguardia, realizó una estancia de verano en el laboratorio de la Prof. E. Charpentier.

De carácter valiente y con una idea clara de cómo se debía formar y valorar una persona como él, decidió comenzar a perseguir el grado de doctor en la TU Delft, en laboratorio de Dr. S.J.J. Brouns, en Países Bajos en 2017. Lejos de su hogar natal, supo hacer de su tiempo ahí una aventura. Científicamente apasionado y exigente, con una filosofía clara de cuáles deben ser los objetivos de la ciencia, se embarcó en el desafío de entender los bacteriófagos, considerándolos como vía para combatir la crisis de la resistencia a antibióticos. Ha sabido explorar realidades que no eran la suya y rodearse de amigos que le ayudaron a expandir aún más sus horizontes. Ha viajado por toda Europa, cargando una mochila como maleta, disfrutando de lo que la vida le ha ofrecido. Esta etapa tiene como culminación la presente tesis, que lo acredita con el grado de doctor.

Amablemente por: **Álvaro Serrano Navarro**

A

Curriculum vitae

Dec. 2021 Present	Research fellow NBIC and University of Southampton Southampton, United Kingdom
April 2017 Jan. 2023	PhD in Molecular Microbiology and Biochemistry Technological University of Delft Delft, The Netherlands
July 2016 Dec. 2016	Research Intern Max Planck Institute for Infection Biology Berlin, Germany
Oct. 2015 July 2016	Master of Science in Virology Universidad Complutense de Madrid Madrid, Spain
May 2015 Aug. 2015	Undergraduate Research Intern Sir William Dunn School of Pathology. Oxford Univ. Oxford, United Kingdom
Sep. 2011 June 2015	Bachelor of Science in Biochemistry Universidad de Castilla-La Mancha Toledo, Spain
Sep. 2009 June 2011	High Education Diploma IES Antonio Calvin Almagro, Spain



List of publications

- 2022 **Cristian Aparicio-Maldonado***, Jack P.K. Bravo*, Franklin L. Nobrega, Stan J.J. Brouns, David W. Taylor. *Structural basis for broad anti-phage immunity by DISARM*. Nature Communications. Doi: 10.1038/s41467-022-30673-1
- 2022 **Cristian Aparicio-Maldonado**, Gal Ofir, Andrea Salini, Rotem Sorek, Franklin L. Nobrega, Stan J.J. Brouns. *Class 1 DISARM provides anti-phage and anti-conjugation activity by unmethylated DNA recognition*. bioRxiv. Manuscript submitted to mBio. Doi: 10.1101/2021.12.28.474362
- 2022 Balwina Koopal, Ana Potocnik, Sumanth K. Mutte, **Cristian Aparicio-Maldonado**, Simon Lindhoud, Jacques J. M. Vervoort, Stan J. J. Brouns, Daan C. Swarts. *Short prokaryotic Argonaute systems trigger cell death upon detection of invading DNA*. Cell. Doi: 10.1016/j.cell.2022.03.012.
- 2022 Yi Wu, Anne van den Hurk, **Cristian Aparicio-Maldonado**, Simran Krishnakant Kushwaha, Claire M. King, Yaqing Ou, Thomas C. Todeschini, Martha R.J. Clokie, Andrew D. Millard, Yilmaz Emre Gençay, Franklin L. Nobrega. *Defence systems provide synergistic anti-phage activity in E. coli*. bioRxiv. Doi: 10.1101/2022.08.21.504612
- 2021 Julia E. Egido, Ana Rita Costa, **Cristian Aparicio-Maldonado**, Pieter-Jan Haas, Stan J. J. Brouns. *Mechanisms and clinical importance of bacteriophage resistance*. FEMS Microbiology Reviews. Doi: 10.1093/femsre/fuab048.
- 2020 Alex N. Salazar, Franklin L. Nobrega, Christine Anyansi, **Cristian Aparicio-Maldonado**, Ana Rita Costa, Anna C Haagsma, Anwar Hiralal, Ahmed Mahfouz, Rebecca E McKenzie, Teunke van Rossum, Stan J. J. Brouns, Thomas Abeel. *An educational guide for nanopore sequencing in the classroom*. PLOS Computational Biology. Doi: 10.1371/journal.pcbi.1007314.



- 2017 Modesto Redrejo-Rodríguez, Carlos D. Ordóñez, Mónica Berjón-Otero, Juan Moreno-González, **Cristian Aparicio-Maldonado**, Patrick Forterre, Margarita Salas, Mart Krupovic. *Primer-Independent DNA Synthesis by a Family B DNA Polymerase from Self-Replicating Mobile Genetic Elements*. Cell Reports. Doi: 10.1016/j.celrep.2017.10.039.
- (2023) **Cristian Aparicio-Maldonado**, Franklin L. Nobrega, Stan J.J. Brouns. *Isolation and characterization of two new Tequatrovirus phages infecting an uropathogenic drug-resistant E. coli strain*. Manuscript in preparation.
- (2023) **Cristian Aparicio-Maldonado**, Amanda van der Sijs, Ana Rita Costa, Stan J. J. Brouns, Franklin L. Nobrega. *Semi-automated electron microscopy to assess quality on clinical bacteriophage preparations*. Manuscript in preparation.



*"You cannot suffer the **past** or the **future**...cause they don't exist.*

*What you are suffering is your **memory** and your **imagination**."*

- Sadhguru -

Therefore, enjoy the **present**, in which coincidentally,

the meaning of life
is
the engine of a film.

