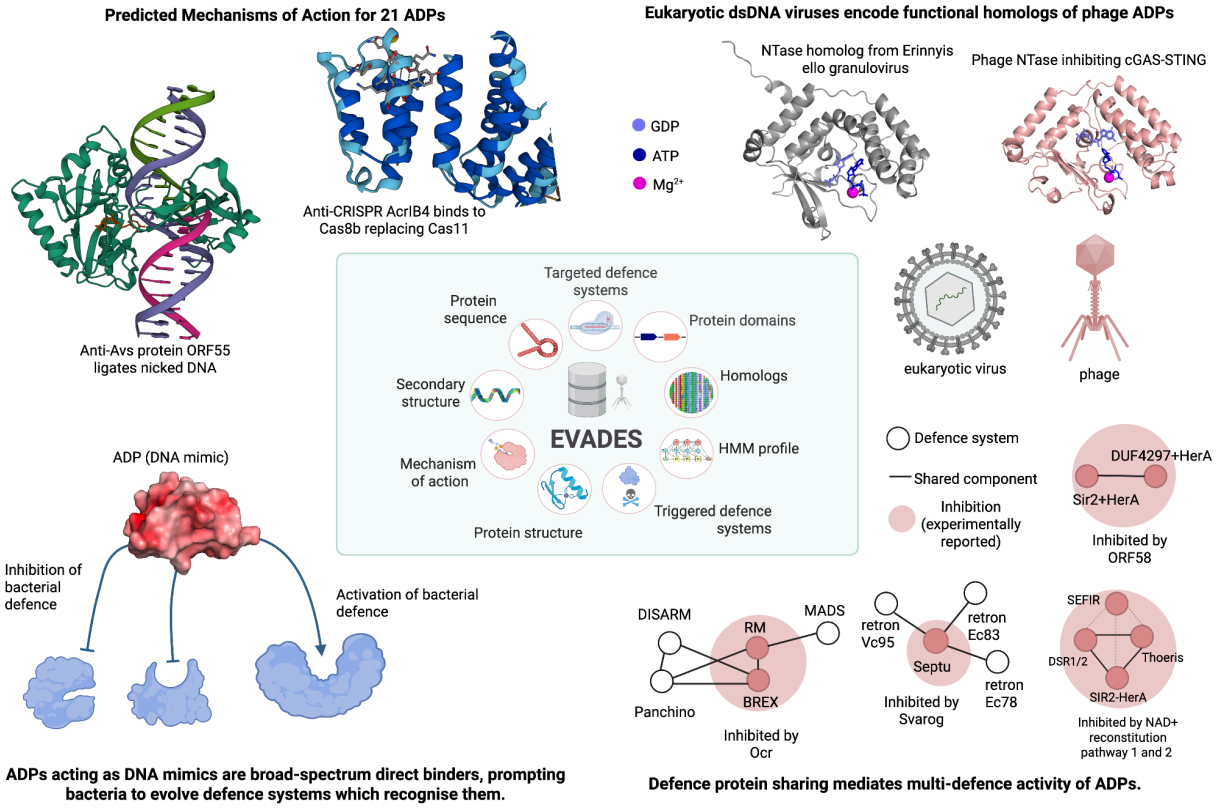
Elucidating the mechanisms of action and evolutionary relationships of phage anti-defence proteins

Khalimat Murtazalieva[1,2*], Evangelos Karatzas[1], Alessio Yang[3#], Federico Corona[3#], Jiawei Wang[1,4], Athanasios Typas[3] and Robert D. Finn[1*]

[1] European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Cambridge CB10 1SD, United Kingdom
[2] University of Cambridge, Clare College, Cambridge CB2 1TN, UK
[3] European Molecular Biology Laboratory, Molecular Systems Biology Unit, Meyerhofstraße 1, 69117 Heidelberg, Germany
[4] Department of Life Sciences, University of Bath, Bath BA2 7AY, UK
# These authors contributed equally to this work.
* Correspondence: kam@ebi.ac.uk (K.M.), rdf@ebi.ac.uk (R.D.F.)



- The EVADES database aggregates knowledge on phage anti-defence proteins (ADPs), facilitating hypothesis generation.
- Mechanisms of action are proposed for 21 uncharacterised ADPs.
- ADPs that function as DNA mimics can block diverse bacterial defence systems.
- Homology of targets of ADPs underpin their broad counter-defence capacities and enable prediction of further targets.
- Eukaryotic double-stranded DNA viruses employ functional homologs of ADPs to evade host immune responses.

SUMMARY

Phages and bacteria are locked in a molecular arms race, with phage anti-defence proteins (ADPs) enabling them to evade bacterial immune systems. To streamline access to information on ADPs, we developed the Encyclopaedia of Viral Anti-DefencE Systems (EVADES), an online resource containing sequences, structures, protein family annotations, and mechanisms of action (MoA) for 268 ADPs. Through computational structural analysis we predicted MoAs for 21 uncharacterised ADPs. We demonstrate the utility of EVADES by exploring different characteristics of ADPs: (i) DNA mimic ADPs exhibit broad-spectrum activity against bacterial defence systems; (ii) protein sharing across defence systems enables multi-defence inhibitory activity of ADPs; and (iii) eukaryotic dsDNA viruses encode ADP homologs, suggesting conserved immune evasion strategies across domains of life. These results highlight the broader relevance of phage ADPs in understanding interactions between prokaryotic or eukaryotic viruses and their hosts. EVADES is freely accessible at https://www.ebi.ac.uk/finn-srv/evades.

KEYWORDS

phage-host interplay, phage anti-defence proteins, protein evolution, eukaryotic viruses

INTRODUCTION

Biological arms races—the adaptation between competing sets of co-evolving genes or organisms—are ubiquitous in nature. One such example is the perpetual arms race between bacteria and their viruses, bacteriophages (hereafter termed phages), which has driven the evolution of myriads of defence systems in bacteria that help them protect against phage infections[1]. Currently, known bacterial defence systems employ a wide range of molecular mechanisms to prevent phage replication. These include degradation of phage nucleic acids, interference with or blocking essential molecular processes for the phage, or mechanisms that involve targeting the host cell before the phage manages to replicate (collectively called abortive infection)[1,2]. In response, phages have developed an array of anti-defence systems aimed at circumventing or neutralising these bacterial defences, thereby ensuring successful infection and replication[3].

Since the discovery of the first anti-defence proteins (ADPs) against restriction-modification (RM) enzymes in the 1970s[4,5], and the explosion of anti-CRISPR mechanism discovery in the 2010's[6], the number of characterised systems has grown significantly[7], particularly over the past decade. This underscores the need for a resource that systematically classifies and describes ADPs to enhance our understanding of their mechanisms and evolutionary diversity. In this study, we compiled a comprehensive collection of ADPs, including their sequences, experimental or predicted secondary and tertiary structures, functional domains, and mechanisms of action (MoA). However, many entries have not been functionally characterised in terms of MoA, questioning whether they directly target defence systems or act via indirect mechanisms. Recognising this uncertainty, we incorporated Evidence Code Ontology (ECO) labels[8] to indicate the type of evidence supporting each entry, thereby improving interpretability and confidence in the data. Complementing this, we conducted computational analyses and proposed putative MoAs for 21 previously uncharacterised ADPs. Multitarget activity is an emerging theme among phage ADPs. Striking examples of broad acting ADPs are phage kinases that counteract defence systems via promiscuous phosphorylation. The T7 protein kinase exhibits anti-defense activity against many nucleic-acid binding defense systems (e.g. Retron-Eco9 and DarTG1) of *E. coli*[9], and so does its homologous *Salmonella* phage JSS

protein kinase (Dnd, CRISPR–Cas, QatABCD, SIR2+HerA and DUF4297+HerA systems)[10]. ADPs can also act on conserved components or functions shared by different bacterial defence systems. Well-characterised examples include: (i) Thoeris anti-defence proteins 1 and 2 (Tad1/2)[11–13], which inhibit both Thoeris and CBASS systems by sequestering cyclic nucleotides; (ii) Ocr, which binds to PglX[14] of the Bacteriophage Exclusion (BREX) system—and homologous HsdM methyltransferase in RM defence; (iii) S-adenosyl-methionine (SAM) lyase[15], which cleaves SAM, a molecule utilised by RM and BREX defences; and (iv) NAD+ reconstitution pathways 1 and 2 (NARP1/2), which evade many defence systems that deplete NAD+, including Thoeris, DSR1, DSR2, SIR2-HerA, and SEFIR[16]. Lastly, there are also ADPs that have poorly understood multitarget activity. ORF58[17], with an uncharacterised MoA, evades both Sir2+HerA and DUF4297+HerA defences, which share structural similarity[18]. We hypothesise that multipurpose ADPs may be widespread, as bacterial defence systems have evolved through protein domain shuffling across different defence systems[19].

Many ADPs are sensed by other defense systems as triggers[20–24]. These cascade relationships are likely a general wiring principle of bacterial defence systems, which have learned to detect anti-defence activities[25]. Activation of a bacterial defense system by a phage ADP may occur through direct interaction with the defence components. For instance, the PARIS defence senses the DNA mimic Ocr from *Escherichia* phage T7 through binding[26] and various retrons are activated by DNA methylation (Dam or Dcm), which is used by phages to circumvent RM systems[23,27]. Bacterial defences may also recognise phage ADPs indirectly. For example, it has been proposed that the retrons Se72 and Ec48 “guard” the RecBCD system in *E. coli*, detecting the inhibition of the RecBCD complex—for instance, by phage *λ* Gam protein or the *T7* Gp5.9 protein—and triggering abortive infection[28]. The Panoptes defence synthesises decoy signalling nucleotides, which are sensed by its effector, OptE, preventing membrane disruption. The phage sponge protein Acb2 activates the Panoptes defence by sequestering these signalling nucleotides and thereby triggering the OptE effector protein[29]. To better understand the properties of ADPs that trigger bacterial defence systems, we examined their features to identify commonalities.

Bacterial defense systems seem to also be ancestral to different components of eukaryotic immunity[30,31]. For instance, nucleotide-binding oligomerization domain–like receptors (NLRs) of the STAND superfamily (receptors in eukaryotes involved in innate immunity) are homologous to the bacterial Avs-defence system, which recognises conserved viral proteins[32]. Evolutionary links between phages and eukaryotic dsDNA viruses have been also extensively described[33–35]. For example, it has been hypothesised that polintoviruses were the first group of eukaryotic dsDNA viruses that evolved from bacteriophages, and that they subsequently gave rise to most large DNA viruses of eukaryotes[33]. This shared evolutionary history, together with the structural and functional homology between the phage anti-CBASS protein 1 (Acb1) and the poxvirus cGAMP phosphodiesterase (inhibits the animal cGLR response)[36], implies that ADPs are conserved between bacterial and eukaryotic viruses. Consistently, phage-encoded Thoeris anti-defence protein 4 (Tad4) and anti-CBASS protein 3 (Acb3) can inhibit related innate immune systems in plants, animals and humans[37], suggesting that retaining phage-evolved ADPs could be beneficial for eukaryotic viruses. Hence, we decided to investigate further the evolutionary and functional relationships of ADPs that counteract defence systems shared between bacteria and eukaryotes.

In this study, we systematically survey the literature for phage ADPs, investigate their functions, targets and potential to act as triggers for other defense systems using bioinformatics. We also extend the systematic comparisons to systems evading eukaryotic immunity. Since ADP

discovery is still at its infancy, we use phage genomic information (diversity islands containing anti-defence genes) and AlphaFold 3[38] to predict ultra-high confidence direct binders among uncharacterised phage proteins to effectors of bacterial defence systems. By uniting structural, evolutionary, and functional insights, our work provides a foundation for advancing the understanding of phage ADPs and their roles in viral–host interplay across domains of life.

RESULTS

Diversity of phage ADPs

We have established the Encyclopaedia of Viral Anti-DefencE Systems (EVADES), which catalogues 268 ADPs that form components of 247 anti-defence systems[9–13,15–18,21,23,24,32,37,39–146], some of which are multicomponent systems (see Methods – Data Collection). Collectively, they target 45 families of bacterial defences. The EVADES database encapsulates a significantly broader diversity of phage ADPs compared to other resources (Figure 1A, Supplementary Table 1). For example, AntiDefenseFinder catalogues 176 ADPs across 156 systems[147]. The dbAPIS database describes 80 experimentally validated ADPs, along with their remote homologs[148]. In contrast, resources such as Anti-CRISPRdb[149,150], and AcrHub[151] focus exclusively on anti-CRISPR (Acr) proteins, with AcrHub containing 76 such proteins.

Beyond its broad range of ADPs entries, EVADES provides detailed MoAs, ECO labels, protein sequences, experimental and predicted structures, profile-hidden Markov models (HMMs), information on protein family membership and homologs in eukaryotic dsDNA viruses, and additional data—including the potential to trigger other defence systems. Only dbAPIS[148] and AcrHub[151] provide comprehensive information, but are restricted in their scopes.

We classified known ADPs into categories based on their MoAs (Figure 1B and Supplementary Figure 1): (i) ADPs with an unknown MoA (n = 108); (ii) ADPs that bind to bacterial defence components (n = 102), including 20 new AlphaFold 3 predictions from this study; (iii) ADPs that degrade or sequester signalling molecules/cofactors utilised by bacterial defences (n =17); (iv) ADPs that modify phage components to evade recognition (n = 15); (v) ADPs that synthesise or restore essential molecules depleted by bacterial defences, including one predicted in this study (n=10); (vi) ADPs that add post-translational modifications to bacterial defence systems (n = 6); (vii) uncategorised ADPs (n = 6); and (viii) proteins that bind and protect components targeted by host defences (n = 4). While 70 out of 136 ADPs with experimentally characterised MoAs have known structures in the Protein Data Bank (PDB)[152], the vast majority of uncharacterised ADPs (119 out of 129) lack structural data. To address this and gain insights into potential MoAs, we predicted protein structures for all ADPs lacking a known structure and performed structural similarity searches (see methods).

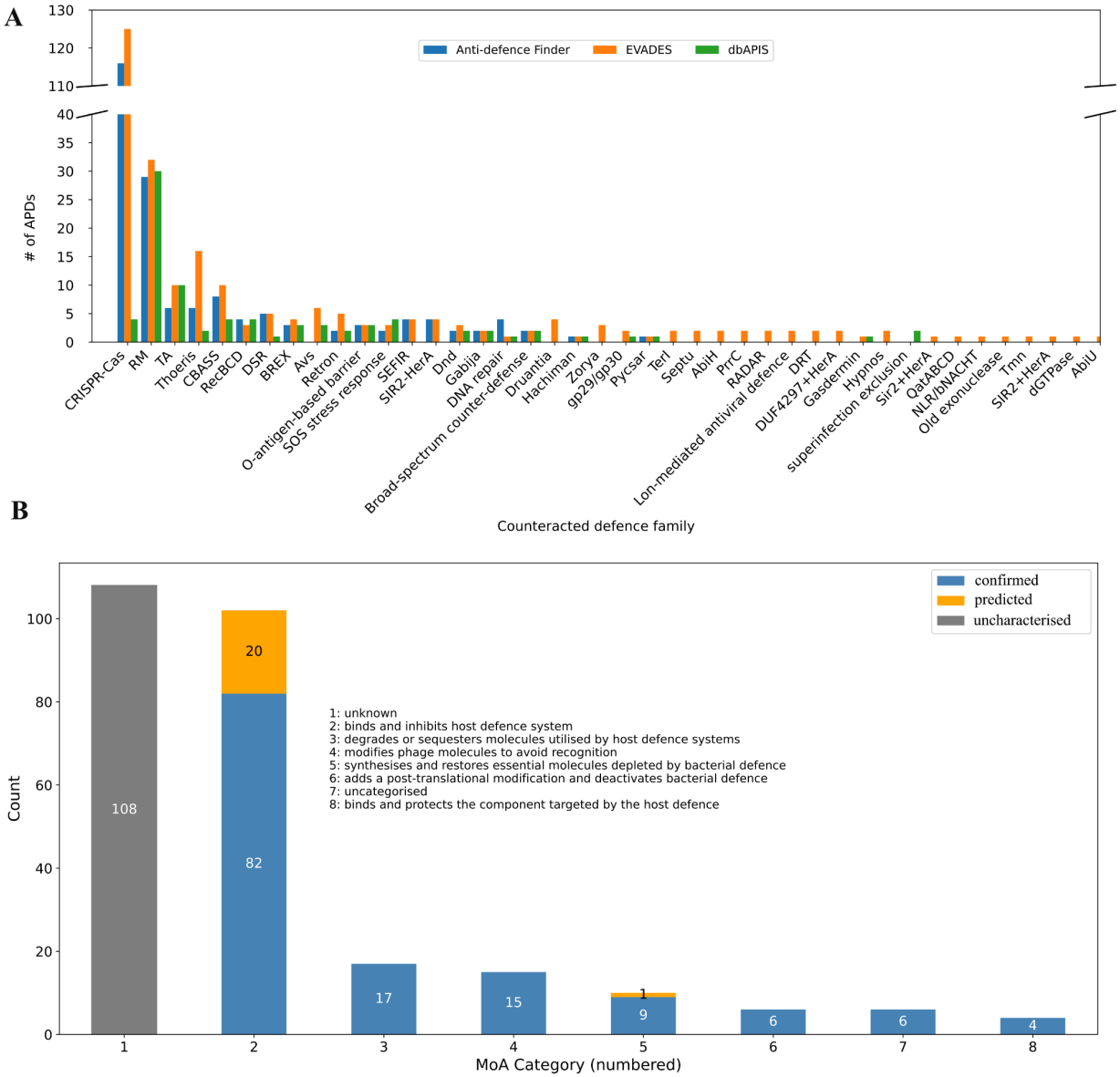


Figure 1. Overview of the database.

(A) Distribution of anti-defence proteins (ADPs) in AntiDefenseFinder (blue), EVADES (orange) and dbAPIS (green) by the family of targeted defence systems. The Y-axis was truncated due to disproportionately high counts of anti-CRISPR proteins compared to other defence systems. **(B)** Distribution of ADPs in EVADES. Bars represent the number of proteins assigned to each MoA category (1–8), split to experimentally confirmed MoAs (blue) and newly predicted ones from this study (orange). The total count for each category is displayed on the corresponding bar.

We also compared the sequences of ADPs to infer evolutionary relationships between them and found many pairs with high similarity at the sequence level (Supplementary Figure 2A). Interestingly, some of these pairs comprised proteins that counteract different defence families. For instance, Tad2[12] and AcrIIA7[54] shared high sequence and structure-based similarity

(Supplementary Figure 2B-C, p-value = 2.55×10^{-8}). Tad2 sequesters gcADPR[12], a signalling molecule utilised by the Thoeris defence system. AcrIIA7 has been reported as type II-A anti-CRISPR protein with unknown MoA[54]. As type II CRISPR–Cas systems are not known to utilise cyclic signalling molecules, this similarity was unexpected. Although AcrIIA7 contains a variable central fragment that aligns poorly with Tad2, the rest of the sequence, including conserved residues, align well (Supplementary Figure 2C). Moreover, AcrIIA7 encodes residues similar to those at Tad2's active site. Based on these results, we concluded that AcrIIA7 is highly likely to sequester gcADPR, similar to Tad2 (Supplementary Figure 2B-C). During the course of this work, this functional prediction was validated by an experimental study[11]; however, the authors were unable to confirm the anti-CRISPR activity previously reported in a functional metagenomic screen[54]. Nevertheless, this example illustrates how structural computational biology approaches can elucidate molecular function based on similarity with other proteins. It also highlights the importance of including information on the nature of evidence supporting the anti-defence activity of database entries.

Inferring multi-defence activity of ADPs from defence protein-sharing network

Multi-defence activity of ADPs can be mediated by exploiting similar protein components of different defence systems. To explore this phenomenon further, we constructed a network based on structural similarity among bacterial defence proteins. We focused on defence proteins that are either directly targeted by ADPs or are presumed to be targeted via uncharacterised mechanisms of action, and expanded it by including other defence proteins that shared significant similarity (Supplementary Table 2). We then added ADPs as triangular nodes to indicate inhibitory interactions. The presence of shared components among defence systems suggests that some ADPs may also inhibit other systems that exhibit similarity to their known targets (Supplementary Figure 3A). For instance, it has been shown that Ocr binds to PglX, a component of the BREX defence system. PglX shares high structural similarity with gp28 from the Panchino defence, drmMI from Defense Island System Associated with Restriction-Modification (DISARM), and HsdM from the type I RM system. While inhibition of the type I RM system by Ocr has been demonstrated—specifically via binding between the HsdM and HsdR subunits— we hypothesise that Ocr may also inhibit DISARM and Panchino defence systems by binding to drmMI and gp28, respectively. This hypothesis is supported by AlphaFold 3 predictions of complexes between Ocr and drmMI or gp28. While these predictions are of a moderate confidence (0.6 > ipTM score < 0.8), the binding interface of Ocr was similar in the two complexes (Supplementary Figure 3B).

20 ADPs with uncharacterised MoA are likely to act as direct binders

Since direct binders represent the most frequent characterised MoA category, we hypothesised that some of the ADPs with uncharacterised function may act via the same mechanism. To investigate this, we co-folded ADPs with uncharacterised MoAs alongside the proteins from the targeted defence systems (Supplementary Table 3). This analysis predicted 13 high-confidence interactions and 10 moderate-confidence interactions between ADPs and their corresponding defence systems (Supplementary Table 4).

While examining the results, we found that the complex between AcrIIA13 and *Staphylococcus aureus* Cas9 (*Sau*Cas9) had been confirmed by an almost identical structure (PDB accession: 7ENI). However, as this PDB entry lacked a corresponding publication, it had not been initially included in our database. Additionally, we identified a recent paper that experimentally

confirmed the interaction between Cas5e and AcrIE3—predicted by AlphaFold 3, with an interface predicted template modelling (ipTM) score of 0.84 and predicted template modelling (pTM) score of 0.84—which had also not been included in our database[63]. We subsequently updated the MoA annotations for AcrIIA13 and AcrIE3, and added the other predictions, indicating that they are based on structural models.

We analysed our predictions alongside the original publications reporting ADPs to gain mechanistic insights and assess whether our predicted results align with available experimental data. For example, we predicted a high-confidence complex (ipTM = 0.92, pTM = 0.85) between AcrIC5 and *Pseudomonas aeruginosa* Cas8c (*Pa*Cas8c, NCBI GenPept accession: UEM35121.1). AcrIC5 has been reported to block *P. aeruginosa* CRISPR interference (CRISPRi), suggesting that it prevents target DNA binding[40]. Consistent with this, our prediction places AcrIC5 near the protospacer adjacent motif (PAM) binding site on PaCas8[153]. Notably, AcrIC5 is a highly acidic molecule (Figure 2A). The interaction between the PAM motif in the target DNA and the PAM-binding site on Cas proteins is essential for DNA binding. AcrIC5 may therefore displace the DNA and sterically block its binding. This prediction is consistent with the observation that AcrIC5 inhibits CRISPRi by preventing target DNA binding[40]. In contrast to AcrIC5, AcrIC3 was shown to be unable to inhibit CRISPRi in the same study[40], suggesting that it acts downstream of DNA binding. Consistent with a downstream target (obstructing Cas3 recruitment or DNA cleavage), we predicted a high-confidence (ipTM = 0.91, pTM = 0.90) interaction between the Cas3 endonuclease (GenBank: UEM35119.1)—spanning all four of its domains—and AcrIC3 (Figure 2B).

AcrIB4 is a puzzling example of an anti-CRISPR protein, as it shares similarity to the C-terminal fragment of *Listeria seeligeri* Cas8b, raising the possibility that it may replace Cas8b during aberrant Cascade assembly. However, experiments have shown that AcrIB4 does not prevent the assembly of Cas8b or other core subunits (Cas7, Cas6, and Cas5) into the Cascade complex, and it has been reported to act upstream of target DNA binding[58]. The authors were unable to generate a functionally tagged version of AcrIB4 to directly assess its incorporation into the Cascade complex[58]. To understand the potential functional role of AcrIB4, we predicted a high-confidence interaction between AcrIB4 and the *Listeria seeligeri* Cas8b (ipTM = 0.91, pTM = 0.76), involving the C-terminal residues of Cas8b that mediate binding to Cas11 (Figure 2C). Cas11 is translated from an internal ribosome-binding site within the 3′ end of the *cas8b* gene and typically stabilises the non-target DNA strand during R-loop formation by positioning it above the crRNA–target DNA heteroduplex[154]. Based on this, we propose that AcrIB4 replaces Cas11 in the complex, thereby destabilising the R-loop formation.

AcrIIA26 has been characterised as a potent inhibitor of *Streptococcus pyogenes* Cas9 (*Spy*Cas9). It does not affect the formation of the Cas9–single guide RNA (sgRNA) complex but prevents this complex from binding to DNA, suggesting that it interferes at a downstream stage of the targeting process[41]. Consistent with this data, we predicted a high-confidence direct interaction (ipTM = 0.87, pTM = 0.79) between AcrIIA26 and *Spy*Cas9, involving the negatively charged surface of AcrIIA26 and residues in *Spy*Cas9, typically involved in target DNA binding (Figure 2D).

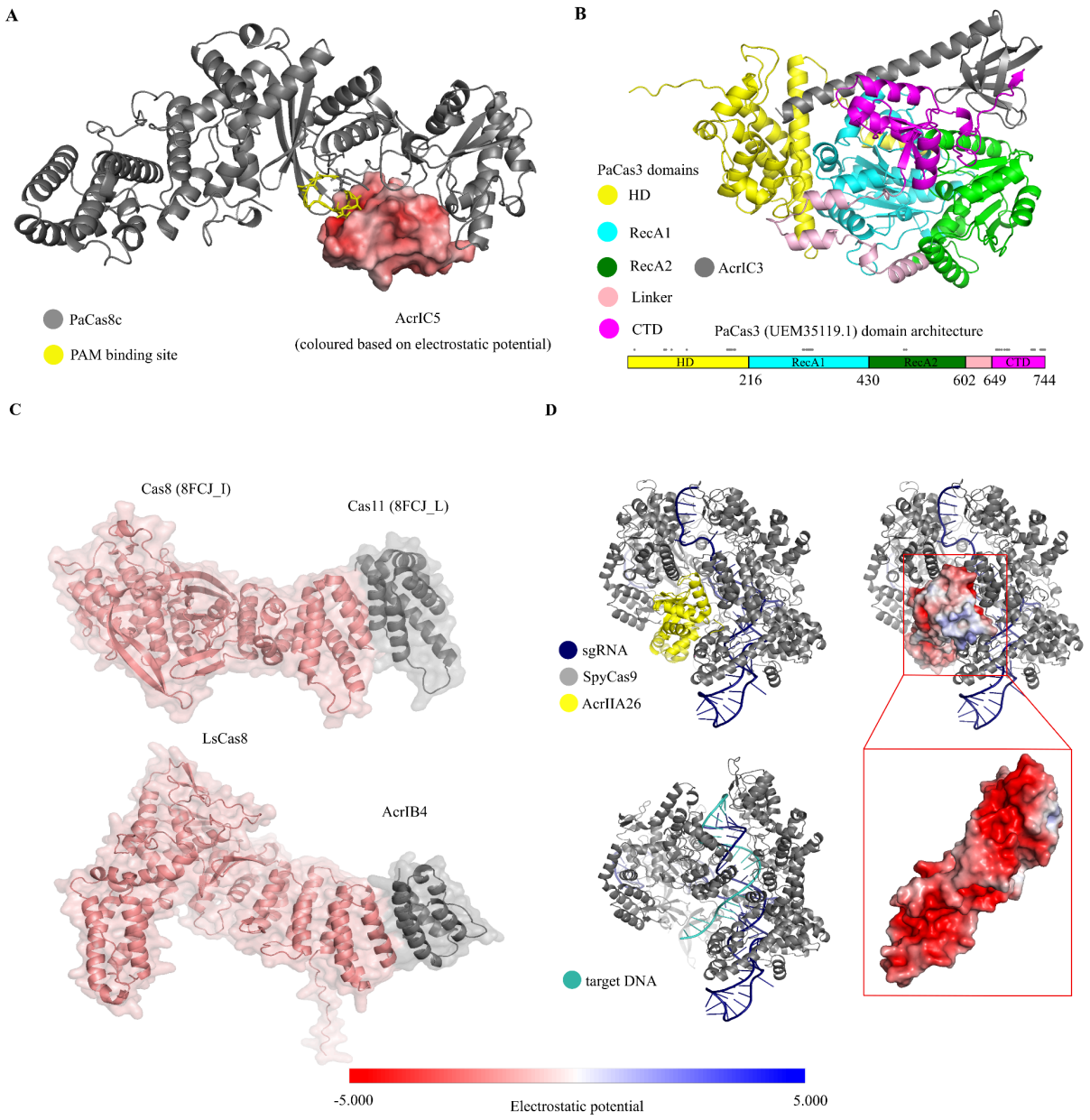


Figure 2. Putative mechanisms of action for ADPs based on structural model analysis.

(A) AcrIC5 (surface rendered and coloured based on electrostatic potential – colour code is shown below the entire Figure) is a negatively charged protein that binds to Cas8c (grey, ribbon representations), occluding its protospacer adjacent motif (PAM)-binding site (yellow, side chains shown a sticks) and preventing Cas8c from interacting with the PAM site on the target DNA. **(B)** AcrIC3 (grey) binds to Cas3, with interactions spanning all four domains of Cas3: the HD nuclease (yellow), RecA1 (cyan), RecA2 (green), and C-terminal domain (CTD, magenta). Underneath the structure is a linear representation of Cas3 domain architecture, which is coloured as in the structure. Grey dots indicate the regions involved in interacting with AcrIC3. **(C)** Typically, the C-terminal domain of *Ls*Cas8b interacts with Cas11 (top), which stabilises

Cas8 within the Cascade. AcrIB4 is likely to bind to Cas8b, replacing Cas11 in the Cascade. Cas8b from *Listeria seeligeri* (*Ls*Cas8b, coloured salmon pink) is predicted to bind AcrIB4 (grey) via its C-terminal domain (bottom). **(D)** AcrIIA26 (yellow, top-left) is predicted to form a high-confidence complex with *Streptococcus pyogenes* Cas9 (*Spy*Cas9, grey) and single guide RNA (cyan, sgRNA). The complex modelled between *Spy*Cas9, sgRNA, and target DNA (cyan, bottom-left) suggests that AcrIIA26 is likely to occlude the *Spy*Cas9's DNA-binding site. In the complex between *Spy*Cas9, sgRNA, and AcrIIA26, the latter is coloured based on electrostatic potential (top-right). We rotated AcrIIA26 (135 degrees around y axis and then 55 angles around y axis, both in counterclockwise directions) to reveal the *Spy*Cas9 interaction interface (bottom-right).

Phage ADPs trigger bacterial defence systems

The mechanisms by which bacterial defence systems sense phage infection remain poorly understood. However, several studies have identified the molecular triggers of these systems[28,155]. Notably, Stokar-Avihail *et al.* conducted a systematic study aimed at uncovering phage determinants sensed by bacterial defences, including putative triggers of defence systems[28]. Their approach involved isolating phage mutants capable of escaping bacterial defences, sequencing them, and identifying mutations in either the protein-coding genes or in the intergenic regions adjacent to such genes. Such mutations were found in genes encoding structural proteins, components of the replication machinery, ADPs and proteins of unknown function. To extend this work, we compared entries in EVADES with those reported by Stokar-Avihail *et al.* to identify potential functions or relationships to newly discovered ADPs. We found that the putative Borvo trigger from phage SECphi18 is, in fact, the entry mga47 in EVADES, which has been reported as a DNA polymerase that is likely capable of replicating DNA ADP-ribosylated by the host DarTG system[130]. The putative ShosTA trigger (mutations resulting in the absence of the protein enable escape from the defense system) is identical to the T7 protein kinase, which inhibits the Retron-Eco9 and DarTG1 defence systems[9], and 53.7% sequence identity over its full length with the *Salmonella* phage JSS1 protein kinase, which inhibits the Dnd, CRISPR–Cas, QatABCD, SIR2+HerA, and DUF4297+HerA systems[10]. Additionally, among the putative triggers, the Avs type I trigger from phage SBSphiC caught our attention, as it was annotated as deoxyuridylate (dUMP) hydroxymethylase. It shared 97% sequence identity over its full length with the deoxyuridylate (dUMP) hydroxymethylase of *Bacillus* phage SP-10, as well as structural similarity to deoxycytidylate hydroxymethylase from phage T4 included in the database (FATCAT *p*-value = 4.88^{e-12}). The dUMP hydroxymethylase of *Bacillus* phage SP-10 is involved in genome hypermodification and protection from restriction enzymes, similar to the deoxycytidylate hydroxymethylase from phage T4, suggesting that the Avs type I trigger may act as an ADP. We further searched for structural homologs of the type I Avs trigger in the PDB and found that it has extremely high similarity to cytidine monophosphate (CMP) hydroxymethylase MilA (Figure 3A, PDB: 5JNH; FATCAT 2.0 p-value =0.00^{e+00})[156]. However, there are indications that it may possess different substrate specificity from MilA. Zhao *et al.* reported that residues K133 and A176 determine MilA's specificity for CMP over dCMP, and mutations K133R and A176S shift its specificity towards dCMP[156]. Structural alignment of MilA with the Avs trigger revealed that R139 in the Avs trigger aligns with K133 in MilA and S183 aligns with A176 (Figure 3A). These observations suggest that the Avs trigger is more likely to function as a dCMP or dUMP hydroxymethylase[28]—known to be involved in DNA hypermodification—rather than as a CMP hydroxymethylase. AlphaFold 3 also predicts a

high-confidence complex between the type I Avs trigger and dUMP (ipTM = 0.97, pTM = 0.96), followed by a slightly lower confidence complexes with dCMP (ipTM = 0.94, pTM = 0.95) and CMP (ipTM = 0.92, pTM = 0.95); these scores are sufficiently similar that differences should be interpreted cautiously. While these scores are not substantially different, collectively these results suggest that the type I Avs trigger is likely to act as an ADP involved in phage DNA hyper modification.

An analysis in viral databases (see methods) revealed that phage ADPs reported to trigger bacterial defence systems (Supplementary Table 5) have, on average, more homologs than other ADPs (416 vs. 223, p-value 0.03, Figure 3B). These ADPs predominantly target bacterial systems that degrade phage nucleic acids and activate defences that induce abortive infection (Figure 3C). This pattern may be explained by evolutionary dynamics: bacteria are more likely to develop new defence strategies against widely prevalent phage ADPs. By coupling a system that degrades phage nucleic acids with an abortive-infection-inducing defence, bacteria can maximise their protection while minimising collateral damage to their own population due to autoimmunity, should the former fail to eliminate the phage. Further supporting this interpretation, the expression of abortive infection (Abi)-inducing defence systems, such as Gabija and Nezha, has been shown to be negatively regulated by active CRISPR systems. Inhibition of CRISPR systems leads to the overexpression of these CRISPR-supervised immune systems (CRISIS), reinforcing the model of layered immune regulation[25,157].

Interestingly, five out of ten ADPs triggering bacterial defence systems are also DNA mimics. Given that there are 11 DNA mimic ADPs in the database (Supplementary table 6), this suggests that DNA mimics are significantly more likely to trigger bacterial defence systems (Fisher's exact test p-value 9.53×10^{-6}, Odds ratio 42), but we acknowledge the total number is small. Nevertheless, we hypothesised that DNA mimics may bind a broad range of bacterial defence proteins that target nucleic acid[99], placing evolutionary pressure on bacteria to develop counter-counter adaptations—defence systems capable of recognising them. To investigate this, we co-folded DNA mimics with bacterial defence systems and predicted 20 high-confidence interactions (Supplementary Table 7), and 10 out of 20 of these interactions were with bacterial defence proteins binding DNA and 9 out of 20 were with the effector proteins. Among the 9 effectors, 4 have uncharacterised mechanisms of action, while 3 target DNA: BrxC, which binds host DNA as part of the BrxBCXZ complex in the BREX defence system[14]; HamA, the nuclease effector of the Hachiman defence that degrades both host and phage DNA[158]; and KwaB, which binds phage DNA to block replication in the Kiwa defence[159]. Unexpectedly, the remaining 2 effectors were cyclases from type I and type IV CBASS defence systems.

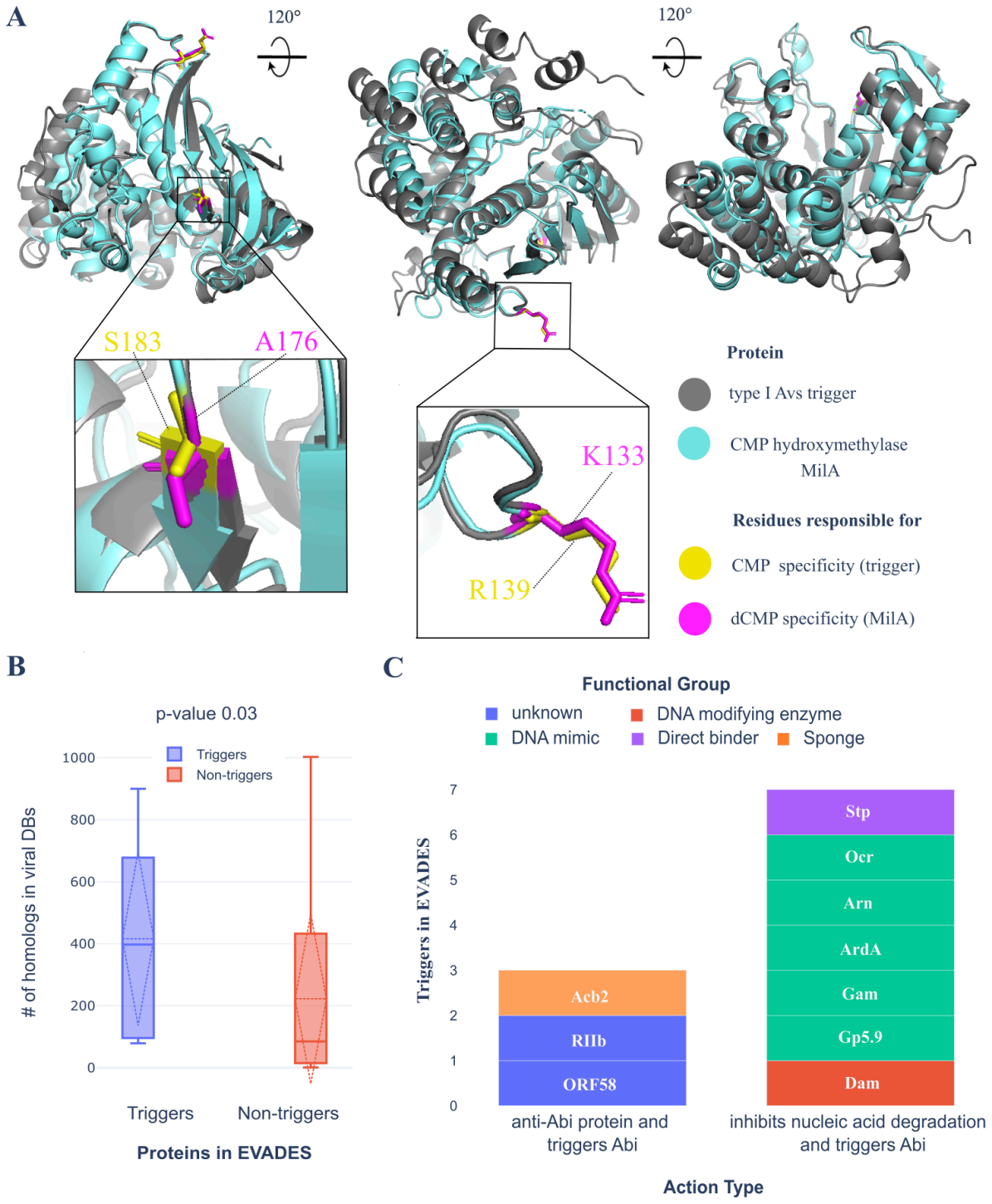


Figure 3. Phage ADPs trigger other defence systems.

(A) Structural superposition of the Avs type I trigger from *Bacillus* phage SBSphiC (grey) and cytidine monophosphate hydroxymethylase MilA (cyan). Alignment was obtained with FATCAT 2.0 (p-value = 0.00^{e+00}). Residues associated with CMP specificity are rendered as sticks and

coloured in yellow, while those associated with dCMP specificity are in magenta and also rendered as sticks. The Avs type I trigger has substitutions likely responsible for dCMP specificity. **(B)** Boxplots showing the number of homologs of ADPs (those known to trigger bacterial defence systems [blue] vs. other ADPs [red]) in phage genomes from GenBank. **(C)** Counts of ADPs found in EVADES that trigger other defence systems, organised based on the mechanism of action (Abortive Infection, Abi) and coloured by the functional group.

Eukaryotic viruses encode several homologs of phage ADPs

In order to assess conservation of defence evasion strategies between bacterial and eukaryotic viruses, we looked for structural homologs of phage ADPs in a database of ~68,000 proteins encoded by eukaryotic viruses, including ~43,000 proteins originating from dsDNA viruses[160]. We identified structural homologs for 27 ADPs, predominantly anti-RM proteins (n=10), followed by Acrs (n=5), anti-Avs (n=2) proteins, and anti-PrrC proteins (n=2). Interestingly, several of these ADPs evade defence systems that are evolutionarily conserved between bacteria and eukaryotes[32,161,162,163]. For instance, the phage encoded ORF55 and ORF83 proteins evade the bacterial type III Avs (antiviral STAND) defence system[17], which shares homology with the STAND superfamily of pattern-recognition receptors involved in eukaryotic innate immunity[32]. Similarly, phage encoded nucleotidyltransferases (NTases) inhibit the STING pathway, a well-known immune signalling mechanism found in both bacteria and eukaryotes[53]. We next conducted a detailed analysis of these structural similarities to evaluate whether the underlying mechanisms of action are also shared between phage and eukaryotic viruses.

Anti-Avs ORF55 is a DNA ligase shared by bacterial and eukaryotic viruses

ORF55 is an anti-Avs type III protein of phage *DruSM1*[17]. Bacterial Avs systems detect phage-encoded proteins such as terminases and portal proteins, triggering homo-oligomerisation and subsequent activation of the N-terminal effectors domains. Among these, DNA nucleases are the most prevalent effectors, inducing non-specific DNA cleavage upon activation. Interestingly, it has been demonstrated that bacterial Avs can also recognise portal proteins and terminases from eukaryotic viruses, including those from human herpesvirus 8[32].

ORF55 has been annotated as a DNA ligase[17]. Based on this annotation, we hypothesised that it may counteract the non-specific dsDNA endonuclease[32] activity of the type III Avs effector proteins by ligating DNA fragments cleaved during Avs activation. To investigate this possibility, we searched for structural homologs of ORF55 in PDB. The closest match was the *Chlorella* virus DNA ligase (PDB: 2Q2U, p-value = 0.00^{e+00})[164], a well-characterised enzyme involved in DNA damage repair[165] (Figure 4A). This DNA ligase represents the minimal functional unit of the ATP-dependent DNA ligases and exhibits intrinsic nick-sensing activity, binding with high affinity to duplex DNA containing a single 3′-OH–5′-PO4 nick[166]. Structural alignment revealed that ORF55 retains 11 out of 13 active site residues conserved in the *Chlorella* virus DNA ligase (Figure 4B)[164]. We used AlphaFold 3 to model the structure of ORF55 in the complex with the DNA fragment co-crystallised with the *Chlorella* virus DNA ligase, which produced a high-confidence complex (ipTM 0.88, pTM = 0.91) with a similar DNA binding interface (Figure 4A). Given that the *Chlorella* virus ligase efficiently ligates nicked DNA, performs poorly on one-nucleotide gaps, and fails entirely on two-nucleotide gaps[166], we modelled three complexes: ORF55 with ATP and (i) nicked DNA, (ii) one-nucleotide gapped DNA, and (iii) two-nucleotide

gapped DNA. AlphaFold 3 predicted with higher confidence the nicked (ipTM = 0.95, pTM = 0.95) and one-nucleotide gapped complexes (ipTM = 0.95, pTM = 0.95) than the two-nucleotide gapped complex (ipTM = 0.87, pTM = 0.91). These results suggest that ORF55 from the DruSM1 phage shares a similar substrate preferences to the *Chlorella* virus DNA ligase, favouring nicked and one-nucleotide gapped DNA.

We identified homologs of ORF55 in 66 different species of eukaryotic dsDNA viruses infecting a wide range of hosts, including amoebae, algae, insects, invertebrates, birds, and mammals, most notably humans. The ORF55 homologs that have been functionally characterised and are involved in DNA repair, for example, the eukaryotic African swine fever virus (ASFV) encodes a homologous DNA ligase. ASFV replicates and assembles virions in the cytoplasm of infected cells, primarily macrophages, which are highly enriched in reactive oxygen species (ROS) and subject to extensive DNA damage. The ASFV DNA ligase has been reported as a crucial component of its DNA repair system[167]. Although viral DNA ligases are generally non-essential for eukaryotic virus DNA replication, since these viruses can utilise cellular DNA ligases for replication, they are essential for conferring resistance to DNA-damaging agents[168]. Normally, the cytosol of eukaryotic cells does not contain dsDNA, and several nucleases act to degrade any cytosolic DNA that arise[169–171]. Therefore, maintaining genome integrity is extremely important for DNA viruses, as DNA damage can trigger antiviral responses such as interferon (IFN)[172] or inflammasome activation[173], ultimately limiting viral propagation.

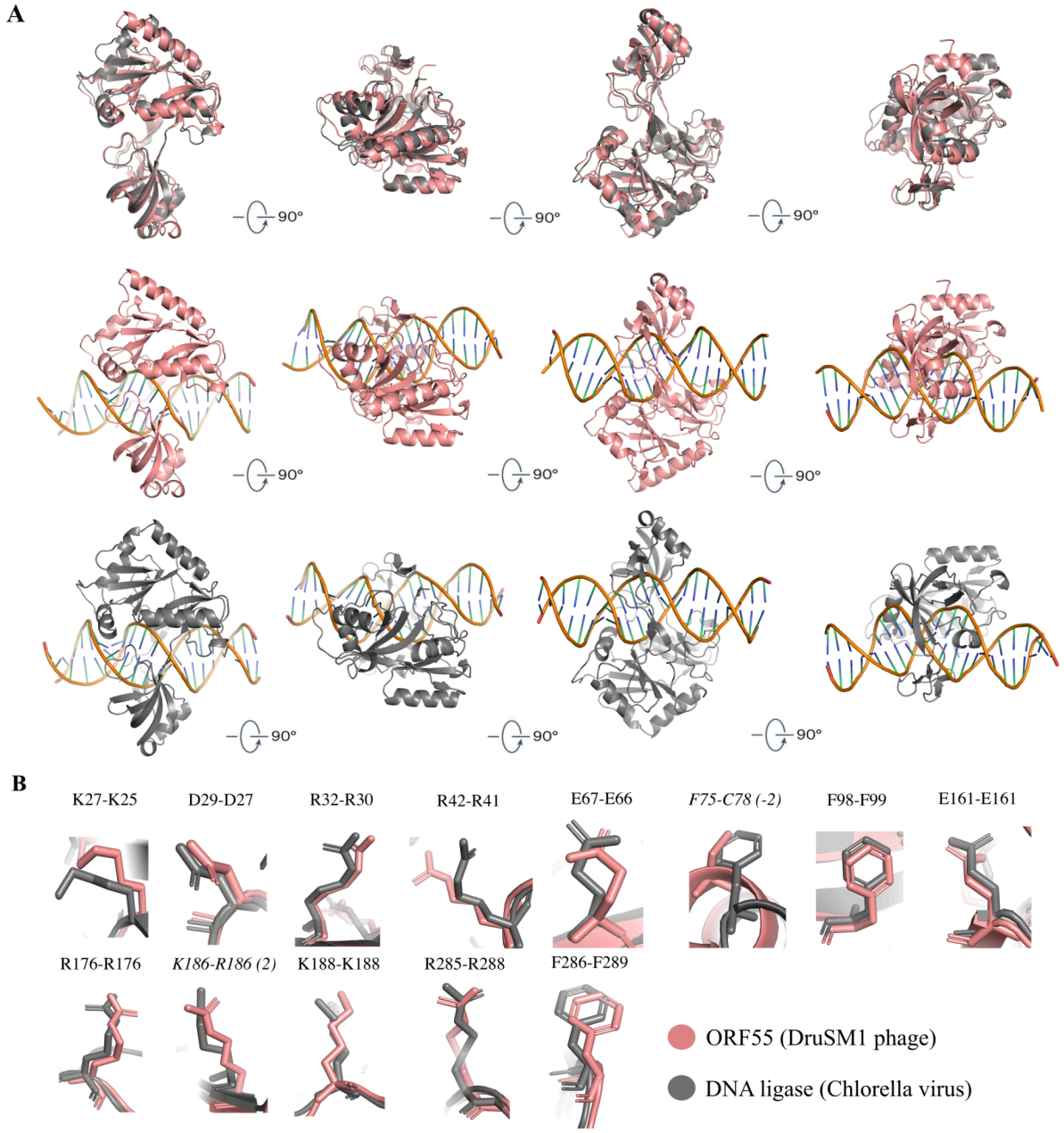


Figure 4. ORF55 shares high similarity with the DNA ligase from Chlorella virus.

(A) From left to right, each view represents a 90° rotation of the previous view around the x-axis. Top: superposition of the *Chlorella* virus DNA ligase (salmon, ribbon representation, PDB: 2Q2U) and ORF55 (grey, ribbon representation) generated using FATCAT 2.0 (p-value 0.00^{e+00}). Middle: the DNA ligase from *Chlorella* virus in complex with DNA (PDB: 2Q2U). Button: a high-confidence complex between ORF55 and the DNA fragment from the complex with the DNA ligase from *Chlorella* virus. **(B)** Thirteen residues contributing to the active site of DNA ligase (grey) are rendered as sticks aligned with the closest residues (rendered in the same way) from ORF55 (salmon).

Anti-Avs protein ORF83 is likely to bind AMP and is encoded by viruses infecting multiple amoeba species

ORF83 is an anti-type III Avs protein, originally annotated as “Smf; Predicted Rossmann fold nucleotide-binding protein DprA/Smf involved in DNA uptake”[17]. We predicted a high-quality structural model of ORF83 using AlphaFold 3 (pLDDT 95, pTM = 0.92) and searched for structural homologs in PDB[174]. The top-scoring matches were *Cg*1261 (PDB ID: 5WQ3, E-value = $4.39e^{-7}$), a type II cytokine-activating protein from the Lonely Guy (LOG) family, along with TT1465, a hypothetical lysine decarboxylases (PDB ID: 1WEK, E-value = $7.95e^{-7}$), and DprA (PDB ID: 3UQZ, E-value = $3.49e^{-5}$). Although TT1465 was initially predicted to be a lysine decarboxylase, experimental studies revealed that it lacks decarboxylase activity and instead exhibits phosphoribohydrolase activity[175]. Both *Cg*1261 and TT1465 were shown to hydrolyse AMP into adenine and ribose 5-phosphate[175]. However, among these, only *Cg*1261 has been experimentally confirmed to possess cytokine-activating activity and the ability the ability to convert N6-(isopentenyl)-adenosine 5′-monophosphate (iPRMP) into isopentenyladenine via dephosphoribosylation.

To investigate whether ORF83 might have similar phosphoribohydrolase activity, we modelled its interaction with AMP using AlphaFold 3, yielding a high-confidence complex (ipTM = 0.9, pTM = 0.88; Figure 5A). Notably, the AMP-binding region of ORF83 aligns with the corresponding AMP-interacting region of *Cg*1261 (Figure 5A). Structural alignment between the active site of *Cg*1261 and ORF83 revealed that they share two invariant residues (R115/R89 and T175/T112), along with a conservative (E178/D114) and a neutral substitution (K116/N90). Given the high structural similarity between ORF83 and *Cg*1261, the confidently predicted AMP interaction, and the partial conservation of active site residues, we propose that ORF83 likely possesses phosphoribohydrolase activity towards AMP. However, the residue composition of ORF83 diverges from that of *Cg*1261 at the prenyl-binding site, thus providing no structural evidence that ORF83 binds iPRMP or functions as a cytokine-activating protein.

We examined the weaker structural similarity between ORF83 and DprA, but ORF83 lacks the equivalent key residues involved in DprA self-dimerisation and RecA binding[176]. AlphaFold predicts a low-likelihood interaction with RecA (ipTM = 0.18, pTM = 0.61). Instead, ORF83 likely binds AMP and functions as a phosphoribohydrolase, though experimental validation is needed to confirm this and clarify its role in evading type III Avs defence.

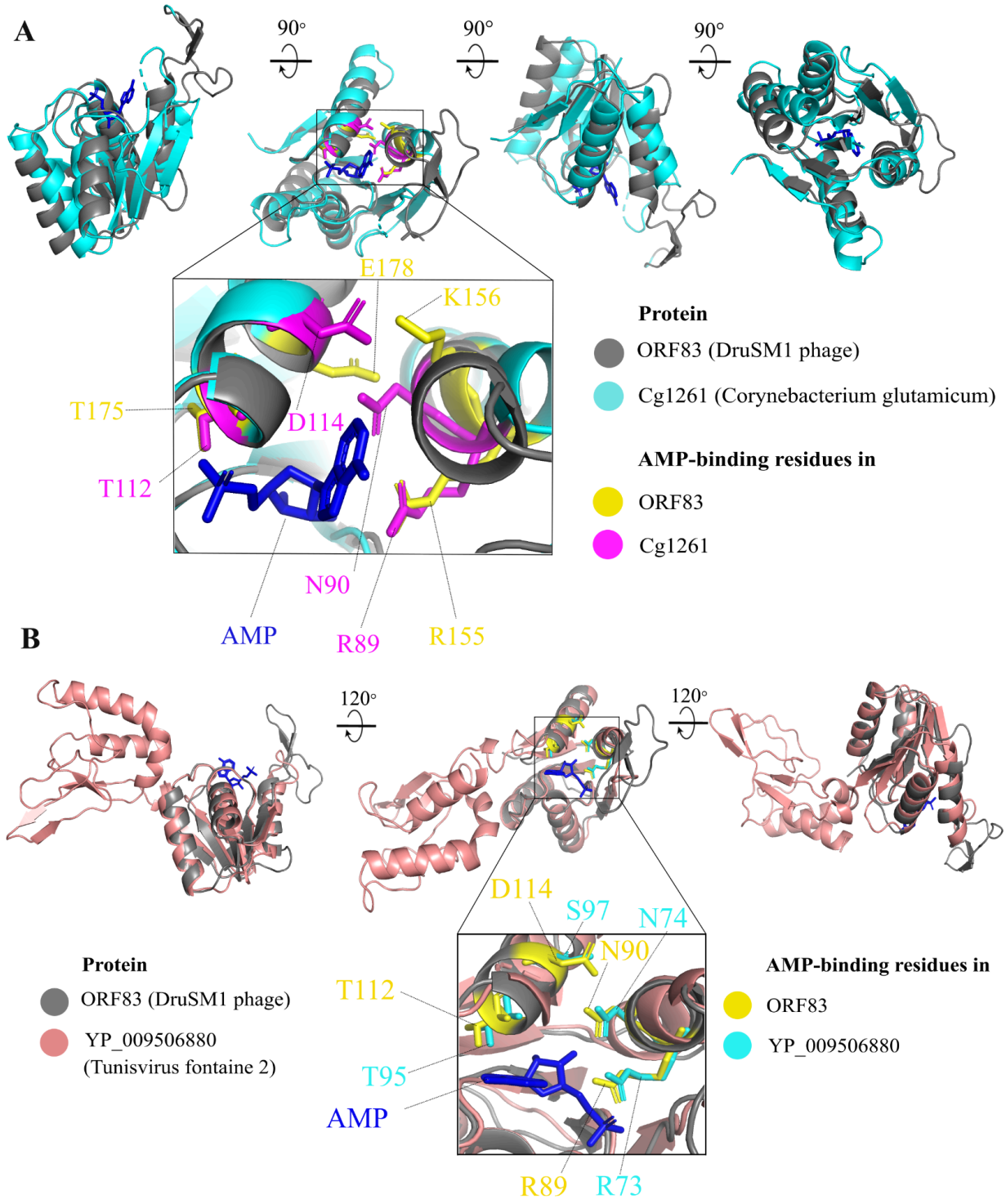


Figure 5. ORF83 is a homolog of Cytokine-Activating Protein Cg1291, binds AMP, and shares similarity with proteins in eukaryotic dsDNA viruses.

(A) Four different views of the superimposition of the AMP (blue, stick representation)-ORF83 (grey, ribbon representation) complex (AlphaFold 3 prediction) with the type II cytokine-activating protein *Cg*1261 (cyan, ribbon representations) of the LOG family, generated using FATCAT 2.0 (p-value = 4.39^{e-7}), each orientation is related by a 90° rotation around the x-axis. The second view provides the best orientation of the AMP-binding site, which is expanded in the bottom insert. Residues in *Cg*1261, characterised as part of the AMP-binding active site, are coloured magenta, and the closest residues in ORF83 are coloured yellow (corresponding to AMP-binding residues predicted by AlphaFold 3). **(B)** Three different views of the superimposition of the complex between AMP (blue, stick representation) and a protein from *Tunisvirus fontaine 2* (salmon, ribbon representation, NCBI GenPept accession: YP_009506880) with the phage protein ORF83 (grey, ribbon representation), generated using FATCAT 2.0 (p-value = 2.99^{e-13}) each related by a 120°.

We found 14 homologs of ORF83 in eukaryotic viruses that all infect various amoeba species. Interestingly, many homologs are longer than ORF83 and contain the domain DUF4326 at the C-terminus. For instance, YP_009506880 from *Tunisvirus fontaine 2* shares extremely high structural similarity to ORF83 (p-value = 2.99^{e-13}; Figure 5B). We predicted its complex with AMP using AlphaFold 3 (ipTM = 0.85, pTM = 0.77). The alignment between ORF83 and YP_009506880 showed that they share the same putative AMP-binding active site (R89-R73, N90-N74, and T112-T95 in ORF83 and YP_009506880, respectively), differing only by one neutral substitution (D114-S97). Overall, these results demonstrate functional homology between ORF83, the anti-Avs protein from phages, and YP_009506880, a protein from eukaryotic viruses.

Phage nucleotidyltransferase homologs in eukaryotic viruses

The cyclic GMP–AMP synthase (cGAS)-STING pathway is a conserved immune system in bacteria and eukaryotes. It is characterised by sensory cGAS proteins that synthesise cyclic nucleotides in response to viral infections. These nucleotides bind the STING domain-containing effector proteins, which initiate downstream immune responses. In bacteria, these effectors typically cleave NAD+, whereas in eukaryotes they activate the expression of antiviral genes.

An NTase enzyme from *Bacillus* phage Bcp1 (NCBI GenPept accession: YP_009031408) synthesises a competing cyclic nucleotide, 3′,3′-cyclic GMP-AMP (3′,3′-cGAMP), which inhibits the cGAS-STING response[53]. We identified 12 homologs of this phage NTase in a range of eukaryotic viruses that infect insects, suggesting a potential role in suppressing the eukaryotic cGAS-STING pathway. For instance, YP_009091875 from *Erinnyis ello* granulovirus shows high structural similarity to the phage NTase (Figure 6A; FATCAT 2.0 p-value 3.10^{e-10}). Importantly, Ho *et al.* identified two catalytic motifs for phage NTases: GS-x-AY[GAN]T-x4-SDxD and NP-x-h2[DE]. The second motif (NP-x-h2[DE]) is unique to phage NTases within the broader NTase superfamily. Notably, the *Erinnyis ello* granulovirus protein exhibits comparable similarity to the consensus active site motifs as the Bcp1 NTase (Figure 6B), matching 10 out of 14 conserved positions, whereas the phage NTase with experimentally verified inhibitory activity against cGAS-STING response matches 11 out of 14.

Nucleoside triphosphates (NTPs) serve as ligands in cyclic nucleotide synthesis, with metal ions such as Mg2+ typically acting as catalytic cofactors[177]. To investigate functional mechanisms of these NTases further, we modelled interactions between NTPs used in cyclic nucleotide synthesis (ATP, GTP and UTP) with both the phage NTase and its homolog in *Erinnyis ello* granulovirus using AlphaFold 3. All predicted complexes yielded ipTM scores above 0.8 (Figure 6C, Supplementary Table 8), indicating high-confidence predictions, and the predicted ipTM scores for complexes of phage and viral proteins show strong correlation (Pearson correlation = 0.865, p-value 0.0261), likely indicating similar substrate preferences. The highest-scoring complex produced for both proteins includes two molecules of UTP and three ions of Mg^{2+}. We show that the highest-confidence complex predicted for protein from *Erinnyis ello* granulovirus uses a similar configuration of the active site as *E. cloacae* CD-NTase CdnD (*Ec*CdnD), in which two Mg^{2+} ions coordinate the donor NTP and one Mg^{2+} ion coordinates the acceptor NTP (Figure 6D)[177]. Moreover, we visualised the residues forming bounds with NTPs and ions and showed that conserved motif SDxD is particularly important, similar to the E-clade of CD-NTases[177] (Figure 6D).

Given the high structural similarity between the phage NTase and putative NTases from the dsDNA eukaryotic viruses, their conserved active site residues, and high-confidence predictions of protein-NTPs-Mg2+ complex formation, we postulate that the viral proteins also synthesise cyclic nucleotides. This activity may enable eukaryotic dsDNA viruses to evade the cGAS-STING response.

Additionally we examined matches between proteins in EVADES and IMG/VR database and identified homologs of Tad1 in three high-confidence viral genome fragments assigned to the order *Durnavirales*. Members of this order are double-stranded RNA (dsRNA) viruses known to infect eukaryotic hosts and these fragments encode RNA-dependent RNA polymerase (RdRp) which serves as a hallmark of RNA viruses. Given that Tad1 can sequester universal cyclic nucleotides utilised by both bacterial CBASS and eukaryotic cGAS immune pathways, such as 2′,3′-cGAMP[178], we speculated that these viral homologues may contribute to evasion of eukaryotic cGAS-mediated immunity. To investigate this hypothesis, we modelled the complex between the *Durnavirales*Tad1 and 2′,3′-cGAMP and it yelled a high-confidence complex (ipTM = 0.90, pTM = 0.91). Structural comparison of the predicted DurnaviralesTad1–2′,3′-cGAMP complex with the experimentally determined Tad1–2′,3′-cGAMP structure (PDB ID: 8KBG) revealed an almost identical overall architecture (Supplementary Figure 4), indicating a conserved mechanism of cyclic dinucleotide recognition. However, because these homologs derive from metagenomic fragments rather than complete viral genomes, definitive evidence that *Durnavirales* species employ Tad1-like proteins to subvert host immunity remains to be established.

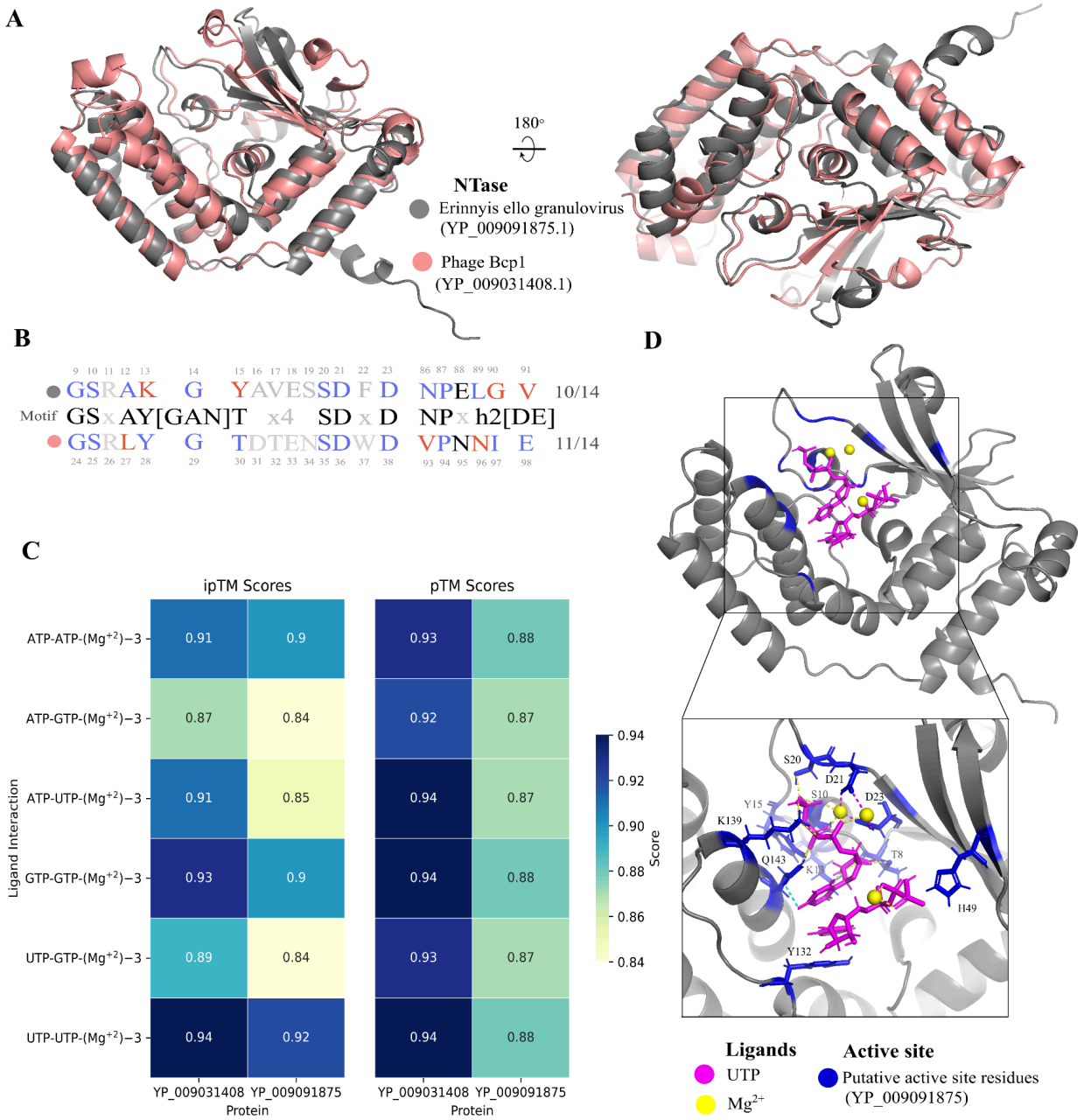


Figure 6. Structural and Functional Analysis of NTase from Bacillus Phage Bcp1 and Erinnyis ello Granulovirus.

(A) Left, structural alignment of the cyclic nucleotide-synthesising NTase from *Bacillus* phage Bcp1 (salmon, NCBI GenPept accession: YP_009031408) and its homolog from *Erinnyis ello* granulovirus (grey, NCBI GenPept accession: YP_009091875), performed using FATCAT 2.0 (p-value: 2.74^{e-10}). Right, the view of the structural alignment is rotation 180° around the X-axis. **(B)** Alignment of catalytic site motifs of phage NTases (middle row) to the experimentally

validated NTase from *Bacillus* phage Bcp1 (bottom row) and its homolog from the dsDNA eukaryotic virus (YP_009091875 from *Erinnyis ello* granulovirus, top row). Matches are in blue, mismatches in red, and residues matching variable motif positions in the sequence motif are in grey. **(C)** Heatmap of confidence scores (ipTM and pTM) for complexes between phage or viral NTases with pairs of NTPs and three ions of Mg^{2+} produced by AlphaFold 3. **(D)** Predicted complex of *Erinnyis ello* granulovirus NTase (grey) with two molecules of UTP (magenta) and three ions of Mg2+ (yellow), the active site residues in the NTase are in blue.

Putative ADPs acting via direct binding from phage diversity hotspots

Diversity hotspots are genomic regions marked by elevated sequence variability, typically flanked by conserved regions or genes, and commonly found in phage and plasmid genomes. These regions emerge through mechanisms such as mobile genetic element insertions, recombination, and hypermutation, and are often enriched in accessory genes related to defence, anti-defence, virulence, and antimicrobial resistance[179]. Notably, similar to bacterial defence systems, phage anti-defence genes frequently co-localise within these hotspots[7,46,53]. We hypothesised that many uncharacterised proteins encoded near anti-defence genes in phage diversity hotspots may contribute to immune evasion. It has been demonstrated[37] that AlphaFold 2 Multimer[180] can predict phage proteins that directly bind and inhibit bacterial defence systems, validating anti-defence activity in 6 out of 21 predicted interactors. Building on this approach, we developed a similar pipeline based on AlphaFold 3 (Supplementary Figure 5), which outperforms AlphaFold 2 in predicting protein-protein interactions[38], to identify phage proteins that interact with bacterial defence systems—focusing specifically on uncharacterised proteins prevalent in diversity hotspots. Using this pipeline, we conducted a large-scale screening of proteins with unknown function from phage diversity islands and identified 46 predicted to bind bacterial defence proteins with high confidence. These predicted ADPs target proteins from 27 distinct families of bacterial defence systems, with ISG15-like systems being the most frequently targeted. As interactions with effector proteins are more likely to result in defence inhibition, we prioritised candidates predicted to bind such effectors, yielding 29 proteins targeting 15 different defence systems (Supplementary Table 9). Yirmiya *et al.* reported that AlphaFold predictions are more reliable when remote homologs are also predicted to bind the same target protein. Accordingly, we co-folded remote homologs of phage proteins with their corresponding defense proteins; two of these pairs passed the threshold confidence score of 0.8. We then incorporated these results to refine the prioritisation of candidate ADPs for future experimental validation.

DISCUSSION

We have curated a comprehensive database of phage ADPs, called EVADES. This database encompasses a larger diversity of ADPs than other resources, such as Anti-DefenseFinder[147] or dbAPIS[148]. More importantly, it provides insights into the MoA of the APDs, as well as catalogues both experimental and predicted structures, information about protein family membership, and the potential to trigger bacterial defence systems. We also categorised phage ADPs based on their MoA. While most ADPs remain uncharacterised, the most prevalent functional characterised category is direct binders–phage proteins that directly interact with bacterial defence proteins to inhibit their activity. We believe this functional information can facilitate biological insights. For instance, knowing that the majority of known ADPs are direct

binders, we hypothesised that some uncharacterised ADPs may inhibit bacterial defences through binding. To test this hypothesis, we modelled the interactions between uncharacterised ADPs and the defence systems they inhibit using AlphaFold 3, and predicted interactions for 20 ADPs. Where possible, we cross-referenced the predicted complexes and experimental data and found that our predictions align with the experimental results. One exciting example is AcrIB4, a protein that was reported as a Cas8b homolog (sharing 38% identity with the last 90 residues), but was not replacing Cas8b in the Cascade complex[58]. We predicted that AcrIB4 binds to Cas8b, specifically interacting with the residues in Cas8b that bind to the Cas11 protein, which stabilises the Cascade complex. Cas11 is translated from an internal ribosome-binding site within the 3′ end of the *cas8b* gene, suggesting that AcrIB4 is homologous to Cas11. We propose that AcrIB4 is likely to replace Cas11, destabilising the Cascade complex. To the best of our knowledge, this is the first example of an Acr acting through complex subunit poisoning by replacing Cas11.

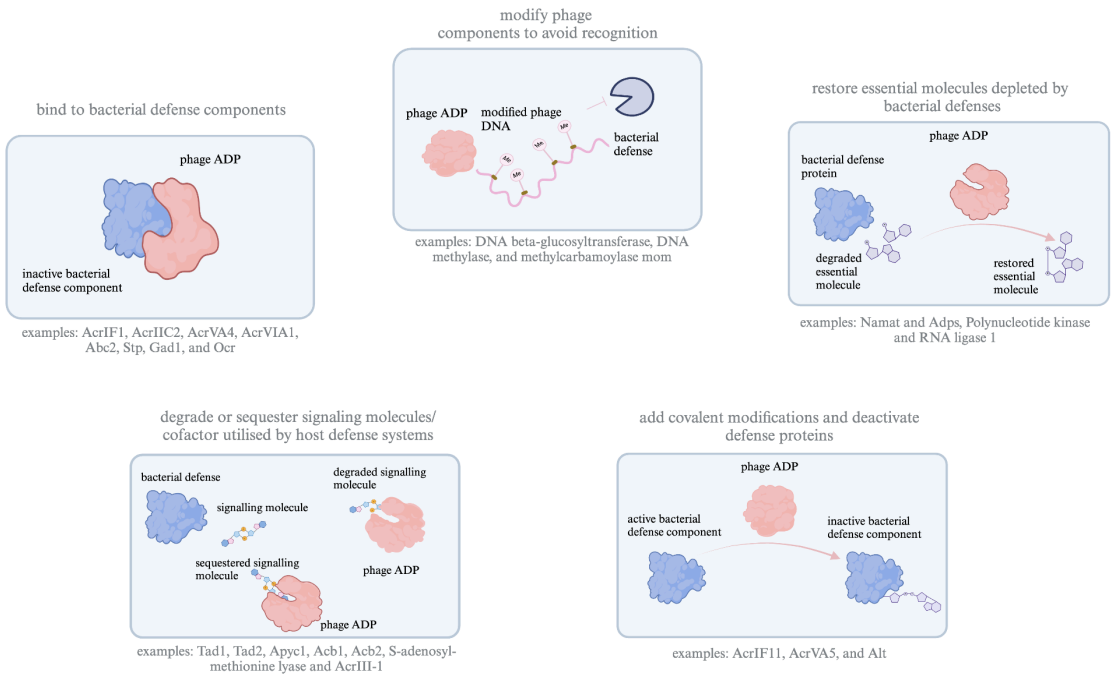
Many ADPs, while inhibiting one defence system, may actually trigger another. These ADPs are widespread and inhibit ubiquitous defence systems including those that cleave phage genetic material. At the same time, those may trigger a secondary line of defences that induce abortive infection. We speculate that this could be a common phenomenon in phage–host interactions, as bacteria evolve counter-counter-defence systems in response to phage counter-defence strategies. Furthermore, we hypothesise that bacterial defence systems operate in layers, where systems that cleave foreign nucleic acids and cause minimal collateral damage serve as the first line of defence. In contrast, systems that trigger abortive infection may sense the inhibition of these first-line defences, providing maximum protection while minimising self-damage. Such interactions between defence systems may also explain the distribution of these systems in bacterial populations, where certain pairs are frequently encoded together due to their evolutionary advantages[181,182]. We showed that ADPs that trigger bacterial defence systems are enriched in those ADPs that act as DNA mimics. We predicted broad-spectrum direct interactions between DNA mimics and bacterial defence proteins, unsurprisingly, particularly those that bind DNA. This layering of defence systems could be explained by the evolutionary pressure exerted by DNA mimics on bacteria, as they can inhibit numerous DNA-binding defence systems, thereby driving the evolution of bacterial counter-counter adaptations. Finally, we found that two proteins reported previously as triggers of bacterial defence systems are identical to ADPs in our database.

We identified numerous homologues of phage ADPs in eukaryotic viruses, suggesting that immune evasion strategies can be conserved across bacterial and eukaryotic viruses. For example, eukaryotic dsDNA viruses encode structural and functional homologues of phage nucleotidyltransferases, which in phages inhibit the bacterial STING effector by producing competing cyclic nucleotides. Because the cGAS–STING pathway is also conserved in eukaryotes, these competing oligonucleotides could similarly suppress the eukaryotic cGAS–STING response—a key antiviral defense—thereby facilitating viral infection. This conservation also provided mechanistic insights into the activity of the phage anti-Avs protein ORF55, informed by its similarity to the minimal nick-sensing DNA ligase found in *Chlorella* virus. Together, these findings highlight how studies of phages can illuminate the immune evasion strategies of eukaryotic viruses, underscoring the potential for reciprocal discovery between bacterial and eukaryotic viral systems.

Finally, we conducted a large-scale computational screen to identify new ADPs among uncharacterised phage proteins, using a previously established approach and focusing on protein encoded in phage diversity islands with anti-defence genes. Thereby, we predicted 29 new candidate ADPs targeting effector proteins of bacterial defence systems. Altogether, EVADES not only consolidates known ADPs into a single, richly annotated resource, but also facilitates elucidation of their mechanisms of action, exploration of their evolutionary dynamics, and discovery of new ADPs.

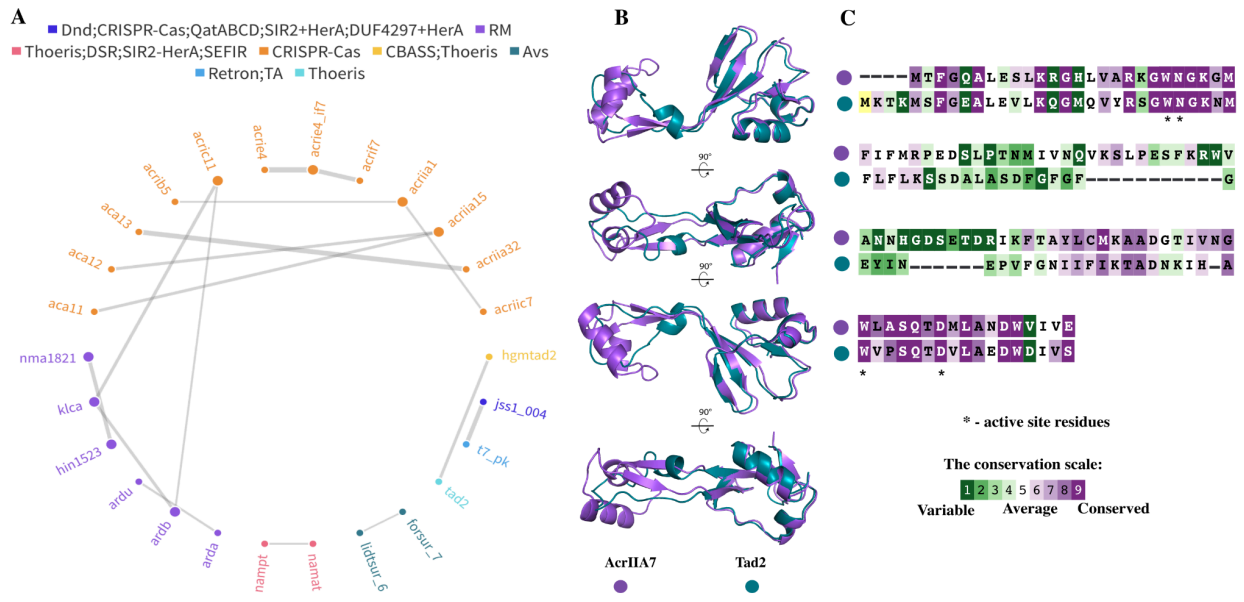
Limitations of the study

Our resource paper provides a comprehensive collection of phage ADPs and proposes several hypotheses which are based on in-depth computational analysis, but are yet to be supported by experimental validation. To enable the users of EVADES to understand those functions that are predicted versus experimentally validated, MoA are labelled with evidence code ontologies. In this paper and the resource, we have focused on high-confidence predictions, reducing the chances of false positives. The analysis in the section devoted to ADPs acting as triggers is based on a small number of entries from the literature, and while the results we present are statistically significant, they may change as more data become available. Although we demonstrated that ORF83 likely binds AMP and has a conserved structure between bacterial and eukaryotic viruses, we have yet to establish the detailed MoA to substantiate a claim that the homolog in eukaryotic viruses also plays a role in immune evasion. Only two out of twenty-nine predicted direct binders maintained high confidence scores (>0.8) in the analysis conducted for remote homologs, suggesting that some of the predictions could be false positives.



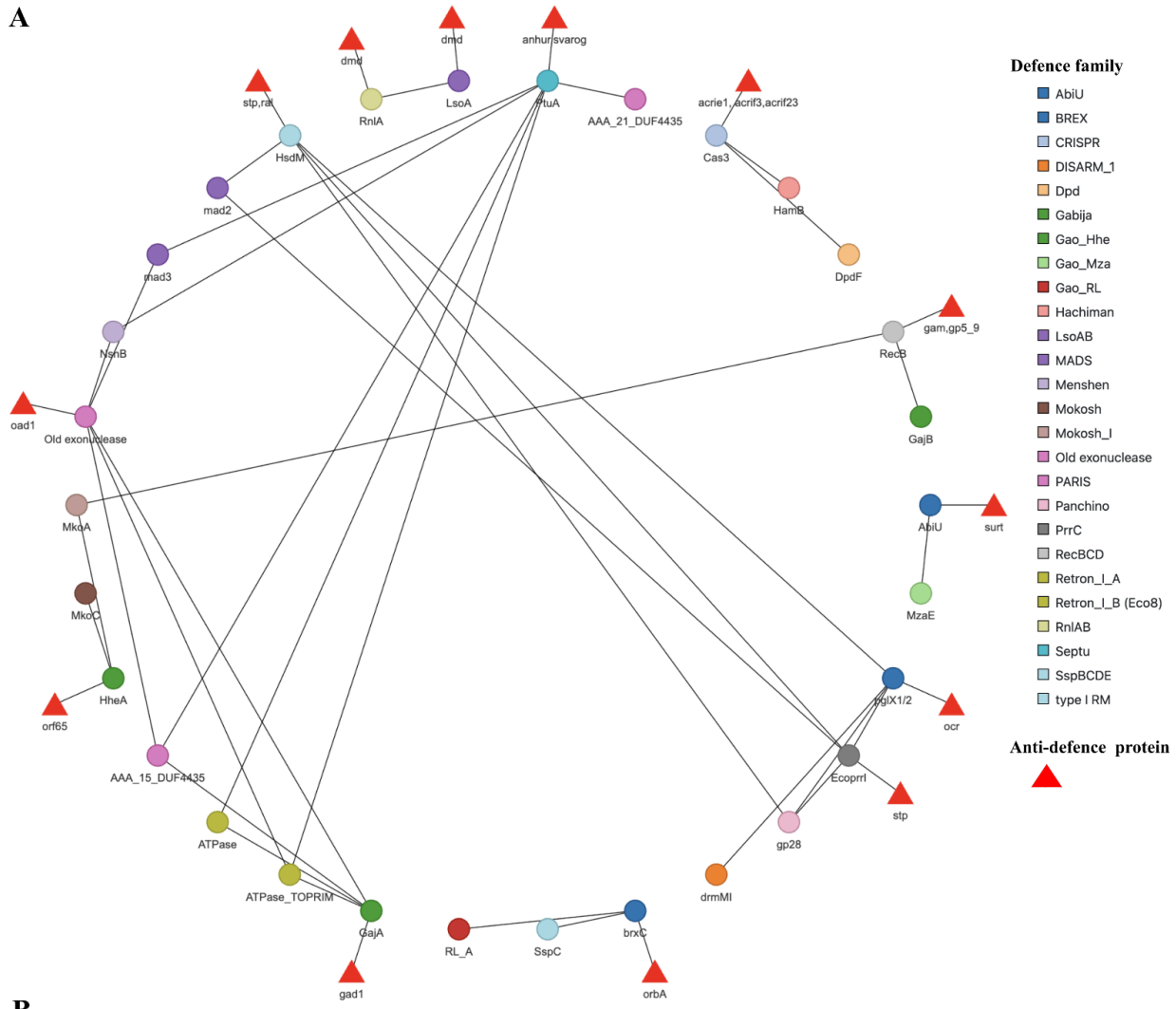
Supplementary Figure 1. Classification of phage ADPs based on mechanism of action.

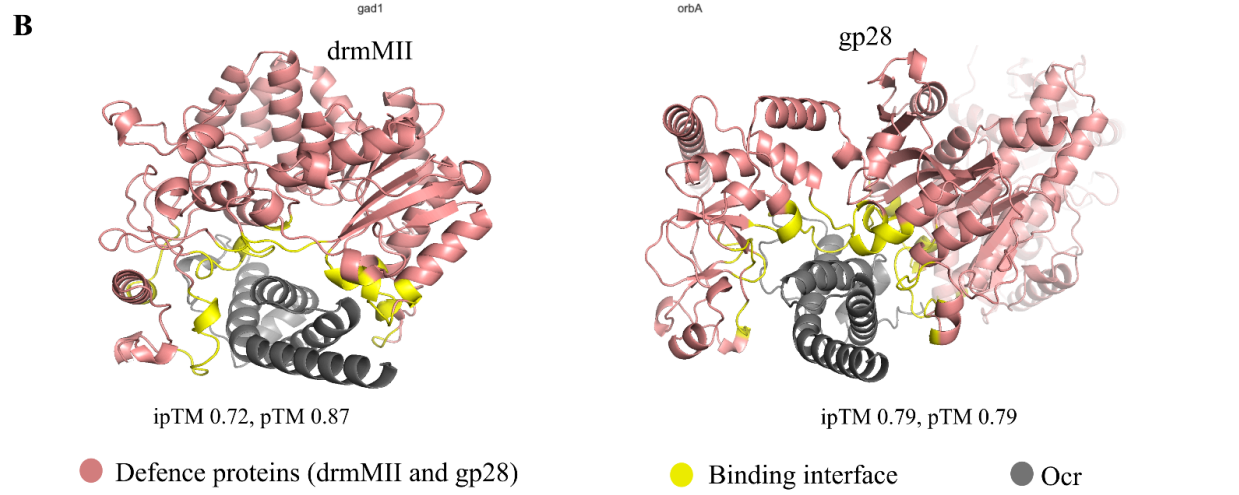
ADPs can be categorised based on their mechanisms of action against bacterial defence systems: (a) ADPs that directly bind and inactivate bacterial defence components (e.g., AcrIF1, AcrIIC2, AcrVA4); (b) ADPs that degrade or sequester signalling molecules or cofactors required by host defences (e.g., Tad1, Acb1, AcrIII-1); (c) ADPs that modify phage components such as DNA to evade recognition (e.g., DNA beta-glucosyltransferase, methylases); (d) ADPs that restore essential molecules depleted by bacterial defences (e.g., Namat, Adps, RNA ligase 1); and (e) ADPs that add covalent modifications to deactivate defence proteins (e.g., AcrIF11, AcrVA5, Alt).



Supplementary Figure 2. Similarity between AcrIIA7 and Tad2.

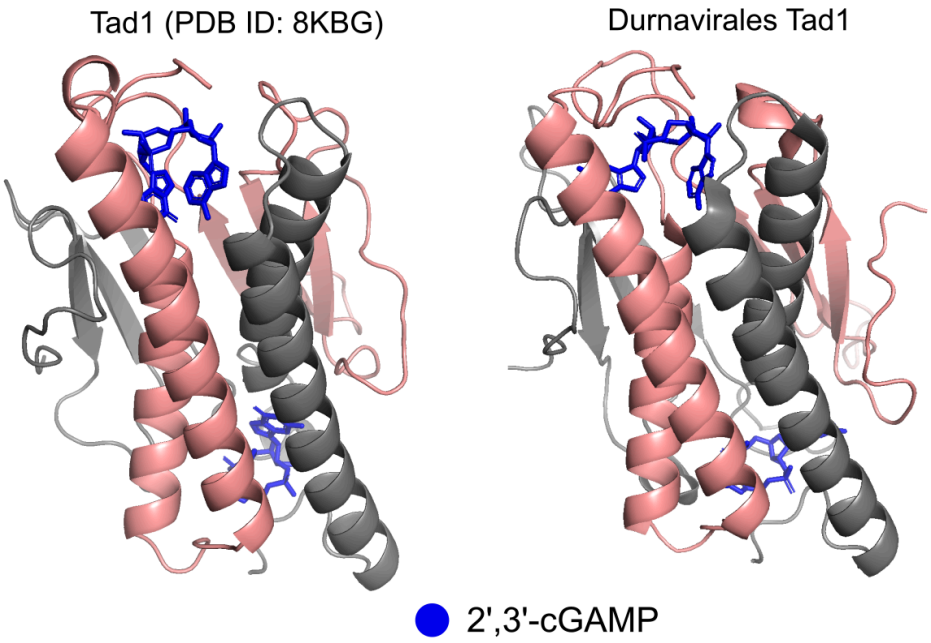
(A) ADP sequence similarity network, ADP node size is proportional to number of proteins an ADP shares similarity to and nodes are coloured based on the defence system(s) they inhibit, edge width is proportional to shared sequence identity. **(B)** Structural alignment between AcrIIA7's structural model (purple) produced by AlphaFold 3 and Tad2 (teal blue, PDB ID: 8SMG) obtained with FATCAT 2.0 (p-value: 2.55^{e-8}). From top to bottom, each view represents a 90° rotation of the previous view around the Y-axis. **(C)** Sequence alignment between AcrIIA7 and Tad (based on structural alignment produced by FATCAT 2.0), residues are coloured based on conservation scores produced by Consurf. Active site residues denoted by asterisk (*).


A
dmd
anhur,svarog
dmd
stp,rat
RnlA
LsoA
PtuA
AAA_21_DUF4435
acrie1,acrif3,acrif23
HsdM
Cas3
mad2
HamB
mad3
DpdF
NshB
gam,gp5_9
RecB
oad1
Old exonuclease
GajB
MkoA
AbiU
surt
MkoC
MzaE
HheA
pglX1/2
orf65
ocr
AAA_15_DUF4435
EcoprrI
stp
ATPase
gp28
ATPase_TOPRIM
drmMI
GajA
RL_A
SspC
brxC
gad1
orbA
Defence family
AbiU
BREX
CRISPR
DISARM_1
Dpd
Gabija
Gao_Hhe
Gao_Mza
Gao_RL
Hachiman
LsoAB
MADS
Menshen
Mokosh
Mokosh_I
Old exonuclease
PARIS
Panchino
PrrC
RecBCD
Retron_I_A
Retron_I_B (Eco8)
RnlAB
Septu
SspBCDE
type I RM
Anti-defence protein


B
drmMII
gp28
ipTM 0.72, pTM 0.87
ipTM 0.79, pTM 0.79
Defence proteins (drmMII and gp28)
Binding interface
Ocr

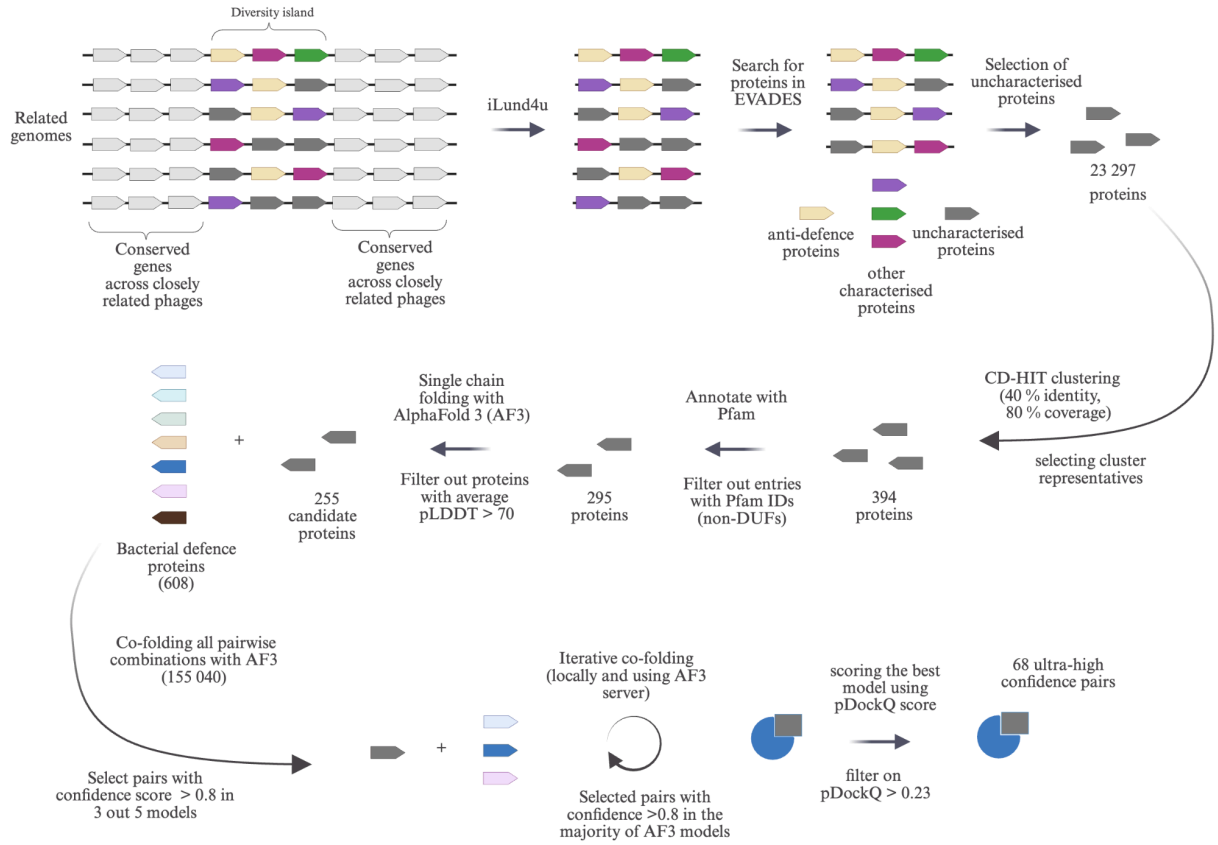
Supplementary Figure 3. Defence protein sharing can mediate multi-defence inhibitory activity of ADPs.

(A) Defence protein (circle nodes, coloured according to the defence system family) similarity network, where edges denote protein pairs that have similar structural folds (FoldSeek e-value = 1.0^{e-8}). ADPs inhibiting the defence system directly or inhibiting single-protein defence systems are denoted as triangular nodes and are connected to the defence protein they target by an edge. **(B)** Complexes Ocr(grey)-drmMII(salmon) and Ocr(grey)-gp28(salmon) predicted by AlphaFold 3. Binding interfaces in drmMII/gp28 are shown in yellow.



Supplementary Figure 4. Genome fragments assigned to order Durnavirales (eukaryotic dsRNA viruses) encode homologs of Tad1.

Experimentally determined structure of the complex between a Tad1 dimer (salmon pink and grey) and two molecules of 2’,3’-cGAMP (blue) is shown on the left. The predicted high-confidence complex (ipTM = 0.90, pTM = 0.91) between a dimer of the Tad1 homolog encoded in Durnavirales (IMG/VR UViG: IMGVR_UViG_3300019224_000070, salmon pink and grey) and two molecules of 2’,3’-cGAMP is shown on the right.



Supplementary Figure 5. Computational pipeline used for predicting proteins binding bacterial defence systems.

The co-folding pipeline selects uncharacterised protein clusters from phage diversity hotspots, filters out those with Pfam annotations or low-confidence models, and applies iterative co-folding to identify interactions with bacterial defences. Phage–bacterial protein pairs were modelled with AlphaFold3, and only those consistently reaching a high confidence score were retained. Finally, interactions were validated with pDockQ, keeping only pairs with scores above 0.23.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Robert D. Finn (rdf@ebi.ac.uk).

Materials availability

Materials generated in this study have been deposited to:
https://ftp.ebi.ac.uk/pub/databases/metagenomics/research-team/projects/evades/

Data and code availability

The database is publicly available at: https://www.ebi.ac.uk/finn-srv/evades

ACKNOWLEDGMENTS

K.M., F.C, and R.D.F. are funded by the European Molecular Biology Laboratory (EMBL). E.K. is an ARISE Fellow and received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 945405. J.W. was a recipient of Marie Skłodowska-Curie Postdoctoral Fellowship, EMBL Interdisciplinary Postdoctoral (EIPOD) Fellowship and EMBO Non-Stipendiary Fellowship (EMBO ALTF 400-2022), and a Junior Research Fellow at Wolfson College, the University of Cambridge, UK.

We would like to thank the MGnify service team members – Mahfouz Shehu, Aleksandar Rajkovic, Martin Beracochea and Sandy Rogers, for their technical assistance and code review, and Anna Lopatina for insightful discussions.

AUTHOR CONTRIBUTIONS

Conceptualization, K.M. and R.D.F.; methodology, K.M., E.K., A.Y., and R.D.F.; Investigation, K.M., A.Y., J.W., and F.C.; writing—original draft, K.M.; writing—review & editing, K.M., F.C., A.Y., A.T., E.K., J.W. A.T. and R.D.F.; funding acquisition, R.D.F.; resources, E.K. and K.M.; supervision, R.D.F.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES

During the preparation of this work, the authors used ChatGPT in order to improve grammar. After using this tool or service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

SUPPLEMENTAL INFORMATION

Figures S1–S5 and Table S1-S10
Table S1. Comparison of EVADES with dbAPIS and Anti-Defense Finder in terms of sequence diversity and defence families
Table S2. Structural similarity between defence proteins
Table S3. Uncharacterised ADPs and proteins from counteracted defence systems used in co-folding
Table S4. Moderate and high-confidence complexes predicted for uncharacterised ADPs
Table S5. Data on protein triggers of bacterial immunity
Table S6. DNA mimics
Table S7. High-confidence complexes predicted between DNA mimics and defence proteins
Table S8. Predicted ipTM/pTM scores for complexes between pairs of NTPs, three ions of Mg^{2+} and YP_009091875 or YP_009031408.
Table S9. High-confidence binders of bacterial defence systems from phage hotspots with anti-defence proteins.
Table S10. Homologs of ADPs in viral and phage fraction of GenBank and IMG/VR 7.1

REFERENCES

1. Georjon, H., and Bernheim, A. (2023). The highly diverse antiphage defence systems of bacteria. Nat. Rev. Microbiol. https://doi.org/10.1038/s41579-023-00934-x.

2. Tal, N., and Sorek, R. (2022). SnapShot: Bacterial immunity. Cell *185*, 578–578.e1.

3. Murtazalieva, K., Mu, A., Petrovskaya, A., and Finn, R.D. (2024). The growing repertoire of phage anti-defence systems. Trends Microbiol. *0*. https://doi.org/10.1016/j.tim.2024.05.005.

4. Spoerel, N., Herrlich, P., and Bickle, T.A. (1979). A novel bacteriophage defence mechanism: the anti-restriction protein. Nature *278*, 30–34.

5. Studier, F.W. (1975). Gene 0.3 of bacteriophage T7 acts to overcome the DNA restriction system of the host. J. Mol. Biol. *94*, 283–295.

6. Pawluk, A., Davidson, A.R., and Maxwell, K.L. (2018). Anti-CRISPR: discovery, mechanism and function. Nat. Rev. Microbiol. *16*, 12–17.

7. Mayo-Muñoz, D., Pinilla-Redondo, R., Camara-Wilpert, S., Birkholz, N., and Fineran, P.C. (2024). Inhibitors of bacterial immune systems: discovery, mechanisms and applications. Nat. Rev. Genet. https://doi.org/10.1038/s41576-023-00676-9.

8. Nadendla, S., Jackson, R., Munro, J., Quaglia, F., Mészáros, B., Olley, D., Hobbs, E.T., Goralski, S.M., Chibucos, M., Mungall, C.J., et al. (2022). ECO: the Evidence and Conclusion Ontology, an update for 2022. Nucleic Acids Res. *50*, D1515–D1521.

9. Bartolec, T., Mitosch, K., Potel, C., Corona, F., Yang, A.L.J., Burtscher, M.L., Koumoutsi, A., Becher, I., Bobonis, J., Karcher, N., et al. (2024). Pervasive phosphorylation by phage T7 kinase disarms bacterial defenses. bioRxiv. https://doi.org/10.1101/2024.12.20.629319.

10. Jiang, S., Chen, C., Huang, W., He, Y., Du, X., Wang, Y., Ou, H., Deng, Z., Xu, C., Jiang, L., et al. (2024). A widespread phage-encoded kinase enables evasion of multiple host antiphage defence systems. Nat. Microbiol. *9*, 3226–3239.

11. Li, D., Xiao, Y., Fedorova, I., Xiong, W., Wang, Y., Liu, X., Huiting, E., Ren, J., Gao, Z., Zhao, X., et al. (2024). Single phage proteins sequester signals from TIR and cGAS-like enzymes. Nature *635*, 719–727.

12. Yirmiya, E., Leavitt, A., Lu, A., Ragucci, A.E., Avraham, C., Osterman, I., Garb, J., Antine, S.P., Mooney, S.E., Hobbs, S.J., et al. (2023). Phages overcome bacterial immunity via diverse anti-defence proteins. Nature. https://doi.org/10.1038/s41586-023-06869-w.

13. Leavitt, A., Yirmiya, E., Amitai, G., Lu, A., Garb, J., Herbst, E., Morehouse, B.R., Hobbs, S.J., Antine, S.P., Sun, Z.-Y.J., et al. (2022). Viruses inhibit TIR gcADPR signalling to overcome bacterial defence. Nature *611*, 326–331.

14. Drobiazko, A., Adams, M.C., Skutel, M., Potekhina, K., Kotovskaya, O., Trofimova, A., Matlashov, M., Yatselenko, D., Maxwell, K.L., Blower, T.R., et al. (2025). Molecular basis of foreign DNA recognition by BREX anti-phage immunity system. Nat. Commun. *16*, 1825.

15. Andriianov, A., Trigüis, S., Drobiazko, A., Sierro, N., Ivanov, N.V., Selmer, M., Severinov, K., and Isaev, A. (2023). Phage T3 overcomes the BREX defense through SAM cleavage and inhibition of SAM synthesis by SAM lyase. Cell Rep. *42*, 112972.

16. Osterman, I., Samra, H., Rousset, F., Loseva, E., Itkin, M., Malitsky, S., Yirmiya, E., Millman, A., and Sorek, R. (2024). Phages reconstitute NAD+ to counter bacterial immunity. Nature *634*, 1160–1167.

17. Ojima, S., Azam, A.H., Kondo, K., Nie, W., Wang, S., Chihara, K., Tamura, A., Yamashita,

W., Nakamura, T., Sugawara, Y., et al. (2024). Systematic discovery of phage genes that inactivate bacterial immune systems. bioRxiv, 2024.04.14.589459. https://doi.org/10.1101/2024.04.14.589459.

18. Lopatina, A., Ferenc, M., Bartlau, N., Wolfram, M., Steensen, K., Muscia, S., Hussain, F., Schnurer, M., Afjehi-Sadat, L., Pollak, S., et al. (2024). Interpretable machine learning reveals a diverse arsenal of anti-defenses in bacterial viruses. bioRxiv, 2024.06.14.598830. https://doi.org/10.1101/2024.06.14.598830.

19. Bernheim, A., Cury, J., and Poirier, E.Z. (2024). The immune modules conserved across the tree of life: Towards a definition of ancestral immunity. PLoS Biol. *22*, e3002717.

20. 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*, 1551–1561.e12.

21. Penner, M., Morad, I., Snyder, L., and Kaufmann, G. (1995). Phage T4-coded Stp: double-edged effector of coupled DNA and tRNA-restriction systems. J. Mol. Biol. *249*, 857–868.

22. Rousset, F., Depardieu, F., Miele, S., Dowding, J., Laval, A.-L., Lieberman, E., Garry, D., Rocha, E.P.C., Bernheim, A., and Bikard, D. (2022). Phages and their satellites encode hotspots of antiviral systems. Cell Host Microbe *30*, 740–753.e5.

23. Bobonis, J., Mitosch, K., Mateus, A., Karcher, N., Kritikos, G., Selkrig, J., Zietek, M., Monzon, V., Pfalz, B., Garcia-Santamarina, S., et al. (2022). Bacterial retrons encode phage-defending tripartite toxin-antitoxin systems. Nature *609*, 144–150.

24. Wein, T., Millman, A., Lange, K., Yirmiya, E., Hadary, R., Garb, J., Melamed, S., Amitai, G., Dym, O., Steinruecke, F., et al. (2025). CARD domains mediate anti-phage defence in bacterial gasdermin systems. Nature, 1–8.

25. Laub, M.T., and Typas, A. (2024). Principles of bacterial innate immunity against viruses. Curr. Opin. Immunol. *89*, 102445.

26. Deep, A., Liang, Q., Enustun, E., Pogliano, J., and Corbett, K.D. (2024). Architecture and activation mechanism of the bacterial PARIS defence system. Nature *634*, 432–439.

27. Wang, Y., Wang, C., Guan, Z., Cao, J., Xu, J., Wang, S., Cui, Y., Wang, Q., Chen, Y., Yin, Y., et al. (2024). DNA methylation activates retron Ec86 filaments for antiphage defense. Cell Rep. *43*, 114857.

28. Stokar-Avihail, A., Fedorenko, T., Hör, J., Garb, J., Leavitt, A., Millman, A., Shulman, G., Wojtania, N., Melamed, S., Amitai, G., et al. (2023). Discovery of phage determinants that confer sensitivity to bacterial immune systems. Cell *186*, 1863–1876.e16.

29. Sullivan, A.E., Nabhani, A., Izrailevsky, D.S., Schinkel, K., Hoffman, C.R.K., Robbins, L.K., Nagy, T.A., Duncan, M.L., Ledvina, H.E., Erbse, A.H., et al. (2025). The Panoptes system uses decoy cyclic nucleotides to defend against phage. Nature, 1–9.

30. Ledvina, H.E., and Whiteley, A.T. (2024). Conservation and similarity of bacterial and eukaryotic innate immunity. Nat. Rev. Microbiol. *22*, 420–434.

31. Wein, T., and Sorek, R. (2022). Bacterial origins of human cell-autonomous innate immune mechanisms. Nat. Rev. Immunol. *22*, 629–638.

32. Gao, L.A., Wilkinson, M.E., Strecker, J., Makarova, K.S., Macrae, R.K., Koonin, E.V., and Zhang, F. (2022). Prokaryotic innate immunity through pattern recognition of conserved viral proteins. Science *377*, eabm4096.

33. Koonin, E.V., Krupovic, M., and Yutin, N. (2015). Evolution of double-stranded DNA viruses of eukaryotes: from bacteriophages to transposons to giant viruses: Evolution of dsDNA viruses of eukaryotes. Ann. N. Y. Acad. Sci. *1341*, 10–24.

34. Koonin, E.V., and Yutin, N. (2010). Origin and evolution of eukaryotic large nucleo-cytoplasmic DNA viruses. Intervirology *53*, 284–292.

35. Iranzo, J., Krupovic, M., and Koonin, E.V. (2016). The double-stranded DNA virosphere as a modular hierarchical network of gene sharing. MBio *7*. https://doi.org/10.1128/mBio.00978-16.

36. Hobbs, S.J., Nomburg, J., Doudna, J.A., and Kranzusch, P.J. (2024). Animal and bacterial viruses share conserved mechanisms of immune evasion. Cell *187*, 5530–5539.e8.

37. Yirmiya, E., Hobbs, S.J., Leavitt, A., Osterman, I., Avraham, C., Hochhauser, D., Madhala, B., Skovorodka, M., Tan, J.M.J., Toyoda, H.C., et al. (2025). Structure-guided discovery of viral proteins that inhibit host immunity. Cell *0*. https://doi.org/10.1016/j.cell.2024.12.035.

38. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature *630*, 493–500.

39. Bondy-Denomy, J., Pawluk, A., Maxwell, K.L., and Davidson, A.R. (2013). Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. Nature *493*, 429–432.

40. León, L.M., Park, A.E., Borges, A.L., Zhang, J.Y., and Bondy-Denomy, J. (2021). Mobile element warfare via CRISPR and anti-CRISPR in Pseudomonas aeruginosa. Nucleic Acids Res. *49*, 2114–2125.

41. Song, G., Zhang, F., Tian, C., Gao, X., Zhu, X., Fan, D., and Tian, Y. (2022). Discovery of potent and versatile CRISPR-Cas9 inhibitors engineered for chemically controllable genome editing. Nucleic Acids Res. *50*, 2836–2853.

42. Pawluk, A., Staals, R.H.J., Taylor, C., Watson, B.N.J., Saha, S., Fineran, P.C., Maxwell, K.L., and Davidson, A.R. (2016). Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species. Nat. Microbiol. *1*, 16085.

43. Pawluk, A., Amrani, N., Zhang, Y., Garcia, B., Hidalgo-Reyes, Y., Lee, J., Edraki, A., Shah, M., Sontheimer, E.J., Maxwell, K.L., et al. (2016). Naturally occurring off-switches for CRISPR-Cas9. Cell *167*, 1829–1838.e9.

44. Marino, N.D., Zhang, J.Y., Borges, A.L., Sousa, A.A., Leon, L.M., Rauch, B.J., Walton, R.T., Berry, J.D., Joung, J.K., Kleinstiver, B.P., et al. (2018). Discovery of widespread type I and type V CRISPR-Cas inhibitors. Science *362*, 240–242.

45. He, F., Bhoobalan-Chitty, Y., Van, L.B., Kjeldsen, A.L., Dedola, M., Makarova, K.S., Koonin,

E.V., Brodersen, D.E., and Peng, X. (2018). Anti-CRISPR proteins encoded by archaeal lytic viruses inhibit subtype I-D immunity. Nat. Microbiol. *3*, 461–469.

46. Pinilla-Redondo, R., Shehreen, S., Marino, N.D., Fagerlund, R.D., Brown, C.M., Sørensen, S.J., Fineran, P.C., and Bondy-Denomy, J. (2020). Discovery of multiple anti-CRISPRs highlights anti-defense gene clustering in mobile genetic elements. Nat. Commun. *11*, 5652.

47. LeGault, K.N., Hays, S.G., Angermeyer, A., McKitterick, A.C., Johura, F.-T., Sultana, M., Ahmed, T., Alam, M., and Seed, K.D. (2021). Temporal shifts in antibiotic resistance elements govern phage-pathogen conflicts. Science *373*. https://doi.org/10.1126/science.abg2166.

48. Silas, S., Carion, H., Makarova, K.S., Laderman, E., Godinez, D.S., Johnson, M., Fossati, A., Swaney, D., Bocek, M., Koonin, E.V., et al. (2023). Parallelized screening of virus accessory genes reveals diverse defense and counter-defense mechanisms. bioRxiv, 2023.04.06.535777. https://doi.org/10.1101/2023.04.06.535777.

49. Hobbs, S.J., Wein, T., Lu, A., Morehouse, B.R., Schnabel, J., Leavitt, A., Yirmiya, E., Sorek, R., and Kranzusch, P.J. (2022). Phage anti-CBASS and anti-Pycsar nucleases subvert bacterial immunity. Nature *605*, 522–526.

50. Huiting, E., Cao, X., Ren, J., Athukoralage, J.S., Luo, Z., Silas, S., An, N., Carion, H., Zhou, Y., Fraser, J.S., et al. (2023). Bacteriophages inhibit and evade cGAS-like immune function in bacteria. Cell *186*, 864–876.e21.

51. Chang, R.B., Toyoda, H.C., Hobbs, S.J., Richmond-Buccola, D., Wein, T., Burger, N., Chouchani, E.T., Sorek, R., and Kranzusch, P.J. (2024). A widespread family of viral sponge proteins reveals specific inhibition of nucleotide signals in anti-phage defense. bioRxiv, 2024.12.30.630793. https://doi.org/10.1101/2024.12.30.630793.

52. Jenson, J.M., Li, T., Du, F., Ea, C.-K., and Chen, Z.J. (2023). Ubiquitin-like conjugation by bacterial cGAS enhances anti-phage defence. Nature *616*, 326–331.

53. Ho, P., Chen, Y., Biswas, S., Canfield, E., Abdolvahabi, A., and Feldman, D.E. (2023). Bacteriophage antidefense genes that neutralize TIR and STING immune responses. Cell Rep. *42*, 112305.

54. Uribe, R.V., van der Helm, E., Misiakou, M.-A., Lee, S.-W., Kol, S., and Sommer, M.O.A. (2019). Discovery and Characterization of Cas9 Inhibitors Disseminated across Seven Bacterial Phyla. Cell Host Microbe *25*, 233–241.e5.

55. Zhang, Z., Pan, S., Liu, T., Li, Y., and Peng, N. (2019). Cas4 Nucleases Can Effect Specific Integration of CRISPR Spacers. J. Bacteriol. *201*. https://doi.org/10.1128/JB.00747-18.

56. Yin, P., Zhang, Y., Yang, L., and Feng, Y. (2023). Non-canonical inhibition strategies and structural basis of anti-CRISPR proteins targeting type I CRISPR-Cas systems. J. Mol. Biol. *435*, 167996.

57. Muzyukina, P., Shkaruta, A., Guzman, N.M., Andreani, J., Borges, A.L., Bondy-Denomy, J., Maikova, A., Semenova, E., Severinov, K., and Soutourina, O. (2023). Identification of an anti-CRISPR protein that inhibits the CRISPR-Cas type I-B system in Clostridioides difficile. mSphere *8*, e0040123.

58. Katz, M.A., Sawyer, E.M., Oriolt, L., Kozlova, A., Williams, M.C., Margolis, S.R., Johnson, M., Bondy-Denomy, J., and Meeske, A.J. (2024). Diverse viral cas genes antagonize CRISPR immunity. Nature *634*, 677–683.

59. Gussow, A.B., Park, A.E., Borges, A.L., Shmakov, S.A., Makarova, K.S., Wolf, Y.I., Bondy-Denomy, J., and Koonin, E.V. (2020). Machine-learning approach expands the repertoire of anti-CRISPR protein families. Nat. Commun. *11*, 3784.

60. Wimmer, F., Englert, F., Wandera, K.G., Alkhnbashi, O.S., Collins, S.P., Backofen, R., and Beisel, C.L. (2024). Interrogating two extensively self-targeting Type I CRISPR-Cas systems in Xanthomonas albilineans reveals distinct anti-CRISPR proteins that block DNA degradation. Nucleic Acids Res. *52*, 769–783.

61. Hu, C., Myers, M.T., Zhou, X., Hou, Z., Lozen, M.L., Nam, K.H., Zhang, Y., and Ke, A. (2024). Exploiting activation and inactivation mechanisms in type I-C CRISPR-Cas3 for genome-editing applications. Mol. Cell *84*, 463–475.e5.

62. Pawluk, A., Bondy-Denomy, J., Cheung, V.H.W., Maxwell, K.L., and Davidson, A.R. (2014). A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of Pseudomonas aeruginosa. MBio *5*, e00896.

63. Koo, J., Lee, G., Park, C., Oh, H., Hong, S.-H., Suh, J.-Y., and Bae, E. (2025). Structural and biochemical insights into the mechanism of the anti-CRISPR protein AcrIE3. Structure *33*, 160–170.e4.

64. Kim, I., Koo, J., An, S.Y., Hong, S., Ka, D., Kim, E.-H., Bae, E., and Suh, J.-Y. (2020). Structural and mechanistic insights into the CRISPR inhibition of AcrIF7. Nucleic Acids Res. *48*, 9959–9968.

65. Rauch, B.J., Silvis, M.R., Hultquist, J.F., Waters, C.S., McGregor, M.J., Krogan, N.J., and Bondy-Denomy, J. (2017). Inhibition of CRISPR-Cas9 with bacteriophage proteins. Cell *168*, 150–158.e10.

66. Forsberg, K.J., Bhatt, I.V., Schmidtke, D.T., Javanmardi, K., Dillard, K.E., Stoddard, B.L., Finkelstein, I.J., Kaiser, B.K., and Malik, H.S. (2019). Functional metagenomics-guided discovery of potent Cas9 inhibitors in the human microbiome. Elife *8*. https://doi.org/10.7554/eLife.46540.

67. Osuna, B.A., Karambelkar, S., Mahendra, C., Christie, K.A., Garcia, B., Davidson, A.R., Kleinstiver, B.P., Kilcher, S., and Bondy-Denomy, J. (2020). Listeria phages induce Cas9 degradation to protect lysogenic genomes. Cell Host Microbe *28*, 31–40.e9.

68. Watters, K.E., Shivram, H., Fellmann, C., Lew, R.J., McMahon, B., and Doudna, J.A. (2020). Potent CRISPR-Cas9 inhibitors from Staphylococcus genomes. Proc. Natl. Acad. Sci. U. S. A. *117*, 6531–6539.

69. Mahendra, C., Christie, K.A., Osuna, B.A., Pinilla-Redondo, R., Kleinstiver, B.P., and Bondy-Denomy, J. (2020). Broad-spectrum anti-CRISPR proteins facilitate horizontal gene transfer. Nat. Microbiol. *5*, 620–629.

70. Forsberg, K.J., Schmidtke, D.T., Werther, R., Uribe, R.V., Hausman, D., Sommer, M.O.A., Stoddard, B.L., Kaiser, B.K., and Malik, H.S. (2021). The novel anti-CRISPR AcrIIA22

relieves DNA torsion in target plasmids and impairs SpyCas9 activity. PLoS Biol. *19*, e3001428.

71. Varble, A., Campisi, E., Euler, C.W., Maguin, P., Kozlova, A., Fyodorova, J., Rostøl, J.T., Fischetti, V.A., and Marraffini, L.A. (2021). Prophage integration into CRISPR loci enables evasion of antiviral immunity in Streptococcus pyogenes. Nat Microbiol *6*, 1516–1525.

72. Song, G., Tian, C., Li, J., Zhang, F., Peng, Y., Gao, X., and Tian, Y. (2023). Rapid characterization of anti-CRISPR proteins and optogenetically engineered variants using a versatile plasmid interference system. Nucleic Acids Res. *51*, 12381–12396.

73. Hynes, A.P., Rousseau, G.M., Lemay, M.-L., Horvath, P., Romero, D.A., Fremaux, C., and Moineau, S. (2017). An anti-CRISPR from a virulent streptococcal phage inhibits Streptococcus pyogenes Cas9. Nat. Microbiol. *2*, 1374–1380.

74. Hynes, A.P., Rousseau, G.M., Agudelo, D., Goulet, A., Amigues, B., Loehr, J., Romero, D.A., Fremaux, C., Horvath, P., Doyon, Y., et al. (2018). Widespread anti-CRISPR proteins in virulent bacteriophages inhibit a range of Cas9 proteins. Nat. Commun. *9*, 2919.

75. Lee, J., Mir, A., Edraki, A., Garcia, B., Amrani, N., Lou, H.E., Gainetdinov, I., Pawluk, A., Ibraheim, R., Gao, X.D., et al. (2018). Potent Cas9 Inhibition in Bacterial and Human Cells by AcrIIC4 and AcrIIC5 Anti-CRISPR Proteins. MBio *9*. https://doi.org/10.1128/mBio.02321-18.

76. Athukoralage, J.S., McMahon, S.A., Zhang, C., Grüschow, S., Graham, S., Krupovic, M., Whitaker, R.J., Gloster, T.M., and White, M.F. (2020). An anti-CRISPR viral ring nuclease subverts type III CRISPR immunity. Nature *577*, 572–575.

77. Bhoobalan-Chitty, Y., Johansen, T.B., Di Cianni, N., and Peng, X. (2019). Inhibition of type III CRISPR-Cas immunity by an Archaeal virus-encoded anti-CRISPR protein. Cell *179*, 448–458.e11.

78. Lin, J., Alfastsen, L., Bhoobalan-Chitty, Y., and Peng, X. (2023). Molecular basis for inhibition of type III-B CRISPR-Cas by an archaeal viral anti-CRISPR protein. Cell Host Microbe *31*, 1837–1849.e5.

79. Duan, N., Hand, E., Pheko, M., Sharma, S., and Emiola, A. (2024). Structure-guided discovery of anti-CRISPR and anti-phage defense proteins. Nat. Commun. *15*, 649.

80. Watters, K.E., Fellmann, C., Bai, H.B., Ren, S.M., and Doudna, J.A. (2018). Systematic discovery of natural CRISPR-Cas12a inhibitors. Science *362*, 236–239.

81. Meeske, A.J., Jia, N., Cassel, A.K., Kozlova, A., Liao, J., Wiedmann, M., Patel, D.J., and Marraffini, L.A. (2020). A phage-encoded anti-CRISPR enables complete evasion of type VI-A CRISPR-Cas immunity. Science *369*, 54–59.

82. Lin, P., Qin, S., Pu, Q., Wang, Z., Wu, Q., Gao, P., Schettler, J., Guo, K., Li, R., Li, G., et al. (2020). CRISPR-Cas13 inhibitors block RNA editing in bacteria and mammalian cells. Mol. Cell *78*, 850–861.e5.

83. Wandera, K.G., Schmelz, S., Migur, A., Kibe, A., Lukat, P., Achmedov, T., Caliskan, N., Blankenfeldt, W., and Beisel, C.L. (2025). AcrVIB1 inhibits CRISPR-Cas13b immunity by

promoting unproductive crRNA binding accessible to RNase attack. Mol. Cell *85*, 1162–1175.e7.

84. Roy, D., Huguet, K.T., Grenier, F., and Burrus, V. (2020). IncC conjugative plasmids and SXT/R391 elements repair double-strand breaks caused by CRISPR-Cas during conjugation. Nucleic Acids Res. *48*, 8815–8827.

85. Cone, R., Bonura, T., and Friedberg, E.C. (1980). Inhibitor of uracil-DNA glycosylase induced by bacteriophage PBS2. Purification and preliminary characterization. J. Biol. Chem. *255*, 10354–10358.

86. Piel, D., Bruto, M., Labreuche, Y., Blanquart, F., Goudenège, D., Barcia-Cruz, R., Chenivesse, S., Le Panse, S., James, A., Dubert, J., et al. (2022). Phage-host coevolution in natural populations. Nat Microbiol *7*, 1075–1086.

87. Garb, J., Lopatina, A., Bernheim, A., Zaremba, M., Siksnys, V., Melamed, S., Leavitt, A., Millman, A., Amitai, G., and Sorek, R. (2022). Multiple phage resistance systems inhibit infection via SIR2-dependent NAD+ depletion. Nat Microbiol *7*, 1849–1856.

88. Antine, S.P., Johnson, A.G., Mooney, S.E., Leavitt, A., Mayer, M.L., Yirmiya, E., Amitai, G., Sorek, R., and Kranzusch, P.J. (2023). Structural basis of Gabija anti-phage defence and viral immune evasion. Nature. https://doi.org/10.1038/s41586-023-06855-2.

89. 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*, 16251.

90. Le, S., Wei, L., Wang, J., Tian, F., Yang, Q., Zhao, J., Zhong, Z., Liu, J., He, X., Zhong, Q., et al. (2024). Bacteriophage protein Dap1 regulates evasion of antiphage immunity and Pseudomonas aeruginosa virulence impacting phage therapy in mice. Nat. Microbiol. *9*, 1828–1841.

91. Zhao, J., Zhu, Y., Wang, C., Tian, F., Deng, J., Liao, J., Zhong, Z., Liu, J., Guo, N., Le, S., et al. (2025). Bacteriophage protein Dap2 inhibits bacterial type III secretion system and synergizes with Dap1 to evade anti-phage immunity. bioRxiv. https://doi.org/10.1101/2025.03.13.642734.

92. Kibby, E.M., Conte, A.N., Burroughs, A.M., Nagy, T.A., Vargas, J.A., Whalen, L.A., Aravind, L., and Whiteley, A.T. (2023). Bacterial NLR-related proteins protect against phage. Cell *186*, 2410–2424.e18.

93. Patel, K.M., and Seed, K.D. (2025). A Class 1 OLD family nuclease encoded by Vibrio cholerae is countered by a vibriophage-encoded direct inhibitor. bioRxivorg. https://doi.org/10.1101/2025.01.06.631583.

94. Amitsur, M., Levitz, R., and Kaufmann, G. (1987). Bacteriophage T4 anticodon nuclease, polynucleotide kinase and RNA ligase reprocess the host lysine tRNA. EMBO J. *6*, 2499–2503.

95. Murphy, K.C., Fenton, A.C., and Poteete, A.R. (1987). Sequence of the bacteriophage P22 anti-recBCD (abc) genes and properties of P22 abc region deletion mutants. Virology *160*, 456–464.

96. Murphy, K.C. (2000). Bacteriophage P22 Abc2 protein binds to RecC increases the 5' strand nicking activity of RecBCD and together with lambda bet, promotes Chi-independent recombination. J. Mol. Biol. *296*, 385–401.

97. Wilkinson, M., Wilkinson, O.J., Feyerherm, C., Fletcher, E.E., Wigley, D.B., and Dillingham, M.S. (2022). Structures of RecBCD in complex with phage-encoded inhibitor proteins reveal distinctive strategies for evasion of a bacterial immunity hub. Elife *11*. https://doi.org/10.7554/eLife.83409.

98. Sakaki, Y., Karu, A.E., Linn, S., and Echols, H. (1973). Purification and Properties of the γ-Protein Specified by Bacteriophage λ: An Inhibitor of the Host RecBC Recombination Enzyme. Proceedings of the National Academy of Sciences *70*, 2215–2219.

99. Zhang, Z., Todeschini, T.C., Wu, Y., Kogay, R., Naji, A., Rodriguez, J.C., Mondi, R., Kaganovich, D., Taylor, D.W., Bravo, J.P.K., et al. (2025). Kiwa is a bacterial membrane-embedded defence supercomplex activated by phage-induced membrane changes. bioRxivorg, 2023.02.26.530102. https://doi.org/10.1101/2023.02.26.530102.

100. Azam, A.H., Chihara, K., Kondo, K., Nakamura, T., Ojima, S., Tamura, A., Yamashita, W., Cui, L., Takahashi, Y., Watashi, K., et al. (2023). Viruses encode tRNA and anti-retron to evade bacterial immunity. bioRxiv, 2023.03.15.532788. https://doi.org/10.1101/2023.03.15.532788.

101. Carabias, A., Camara-Wilpert, S., Mestre, M.R., Lopéz-Méndez, B., Hendriks, I.A., Zhao, R., Pape, T., Fuglsang, A., Luk, S.H.-C., Nielsen, M.L., et al. (2024). Retron-Eco1 assembles NAD+-hydrolyzing filaments that provide immunity against bacteriophages. Mol. Cell *84*, 2185–2202.e12.

102. Shinedling, S., Parma, D., and Gold, L. (1987). Wild-type bacteriophage T4 is restricted by the lambda rex genes. J. Virol. *61*, 3790–3794.

103. Belogurov, A.A., Delver, E.P., and Rodzevich, O.V. (1992). IncN plasmid pKM101 and IncI1 plasmid ColIb-P9 encode homologous antirestriction proteins in their leading regions. J. Bacteriol. *174*, 5079–5085.

104. Serfiotis-Mitsa, D., Herbert, A.P., Roberts, G.A., Soares, D.C., White, J.H., Blakely, G.W., Uhrín, D., and Dryden, D.T.F. (2010). The structure of the KlcA and ArdB proteins reveals a novel fold and antirestriction activity against Type I DNA restriction systems in vivo but not in vitro. Nucleic Acids Res. *38*, 1723–1737.

105. González-Montes, L., Del Campo, I., Garcillán-Barcia, M.P., de la Cruz, F., and Moncalián, G. (2020). ArdC, a ssDNA-binding protein with a metalloprotease domain, overpasses the recipient hsdRMS restriction system broadening conjugation host range. PLoS Genet. *16*, e1008750.

106. Belogurov, A.A., Delver, E.P., and Rodzevich, O.V. (1993). Plasmid pKM101 encodes two nonhomologous antirestriction proteins (ArdA and ArdB) whose expression is controlled by homologous regulatory sequences. J. Bacteriol. *175*, 4843–4850.

107. Meima, R., and Lidstrom, M.E. (2000). Characterization of the Minimal Replicon of a Cryptic Deinococcus radiodurans SARK Plasmid and Development of Versatile Escherichia coli-D. radiodurans Shuttle Vectors. Applied and Environmental Microbiology.

https://doi.org/10.1128/aem.66.9.3856-3867.2000.

108. Dharmalingam, K., and Goldberg, E.B. (1976). Phage-coded protein prevents restriction of unmodified progeny T4 DNA. Nature *260*, 454–456.

109. Ho, C.-H., Wang, H.-C., Ko, T.-P., Chang, Y.-C., and Wang, A.H.-J. (2014). The T4 Phage DNA Mimic Protein Arn Inhibits the DNA Binding Activity of the Bacterial Histone-like Protein H-NS*. J. Biol. Chem. *289*, 27046–27054.

110. Vrielink, A., Rüger, W., Driessen, H.P., and Freemont, P.S. (1994). Crystal structure of the DNA modifying enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. EMBO J. *13*, 3413–3422.

111. Schlagman, S.L., and Hattman, S. (1983). Molecular cloning of a functional dam+ gene coding for phage T4 DNA adenine methylase. Gene *22*, 139–156.

112. Iida, S., Streiff, M.B., Bickle, T.A., and Arber, W. (1987). Two DNA antirestriction systems of bacteriophage P1, darA, and darB: characterization of darA- phages. Virology *157*, 156–166.

113. Song, H.K., Sohn, S.H., and Suh, S.W. (1999). Crystal structure of deoxycytidylate hydroxymethylase from bacteriophage T4, a component of the deoxyribonucleoside triphosphate-synthesizing complex. EMBO J. *18*, 1104–1113.

114. Piya, D., Vara, L., Russell, W.K., Young, R., and Gill, J.J. (2017). The multicomponent antirestriction system of phage P1 is linked to capsid morphogenesis. Mol. Microbiol. *105*, 399–412.

115. Teplyakov, A., Sebastiao, P., Obmolova, G., Perrakis, A., Brush, G.S., Bessman, M.J., and Wilson, K.S. (1996). Crystal structure of bacteriophage T4 deoxynucleotide kinase with its substrates dGMP and ATP. EMBO J. *15*, 3487–3497.

116. Hutinet, G., Kot, W., Cui, L., Hillebrand, R., Balamkundu, S., Gnanakalai, S., Neelakandan, R., Carstens, A.B., Fa Lui, C., Tremblay, D., et al. (2019). 7-Deazaguanine modifications protect phage DNA from host restriction systems. Nat. Commun. *10*, 5442.

117. Drozdz, M., Piekarowicz, A., Bujnicki, J.M., and Radlinska, M. (2012). Novel non-specific DNA adenine methyltransferases. Nucleic Acids Res. *40*, 2119–2130.

118. Rifat, D., Wright, N.T., Varney, K.M., Weber, D.J., and Black, L.W. (2008). Restriction endonuclease inhibitor IPI* of bacteriophage T4: a novel structure for a dedicated target. J. Mol. Biol. *375*, 720–734.

119. Engelberg-Kulka, H., Reches, M., Narasimhan, S., Schoulaker-Schwarz, R., Klemes, Y., Aizenman, E., and Glaser, G. (1998). rexB of bacteriophage lambda is an anti-cell death gene. Proc. Natl. Acad. Sci. U. S. A. *95*, 15481–15486.

120. King, G., and Murray, N.E. (1995). Restriction alleviation and modification enhancement by the Rac prophage of Escherichia coli K-12. Mol. Microbiol. *16*, 769–777.

121. Karambelkar, S., Udupa, S., Gowthami, V.N., Ramachandra, S.G., Swapna, G., and Nagaraja, V. (2020). Emergence of a novel immune-evasion strategy from an ancestral protein fold in bacteriophage Mu. Nucleic Acids Res. *48*, 5294–5305.

122. Kahmann, R., Seiler, A., Wulczyn, F.G., and Pfaff, E. (1985). The mom gene of bacteriophage mu: a unique regulatory scheme to control a lethal function. Gene *39*, 61–70.

123. Toussaint, A. (1976). The DNA modification function of temperate phage Mu-1. Virology *70*, 17–27.

124. Loenen, W.A., and Murray, N.E. (1986). Modification enhancement by the restriction alleviation protein (Ral) of bacteriophage lambda. J. Mol. Biol. *190*, 11–22.

125. Walkinshaw, M.D., Taylor, P., Sturrock, S.S., Atanasiu, C., Berge, T., Henderson, R.M., Edwardson, J.M., and Dryden, D.T.F. (2002). Structure of Ocr from bacteriophage T7, a protein that mimics B-form DNA. Mol. Cell *9*, 187–194.

126. Studier, F.W., and Movva, N.R. (1976). SAMase gene of bacteriophage T3 is responsible for overcoming host restriction. J. Virol. *19*, 136–145.

127. Loh, S., Skurray, R., Célérier, J., Bagdasarian, M., Bailone, A., and Devoret, R. (1990). Nucleotide sequence of the psiA (plasmid SOS inhibition) gene located on the leading region of plasmids F and R6-5. Nucleic Acids Res. *18*, 4597.

128. Petrova, V., Chitteni-Pattu, S., Drees, J.C., Inman, R.B., and Cox, M.M. (2009). An SOS inhibitor that binds to free RecA protein: the PsiB protein. Mol. Cell *36*, 121–130.

129. Bagdasarian, M., Bailone, A., Angulo, J.F., Scholz, P., Bagdasarian, M., and Devoret, R. (1992). PsiB, an anti-SOS protein, is transiently expressed by the F sex factor during its transmission to an Escherichia coli K-12 recipient. Mol. Microbiol. *6*, 885–893.

130. LeRoux, M., Srikant, S., Teodoro, G.I.C., Zhang, T., Littlehale, M.L., Doron, S., Badiee, M., Leung, A.K.L., Sorek, R., and Laub, M.T. (2022). The DarTG toxin-antitoxin system provides phage defence by ADP-ribosylating viral DNA. Nat Microbiol *7*, 1028–1040.

131. Patel, K.M., and Seed, K.D. (2024). Sporadic phage defense in epidemic Vibrio cholerae mediated by the toxin-antitoxin system DarTG is countered by a phage-encoded antitoxin mimic. MBio *15*, e0011124.

132. Johannesman, A., Carlson, N.A., and LeRoux, M. (2024). Phages carry orphan antitoxin-like enzymes to neutralize the DarTG1 toxin-antitoxin defense system. bioRxivorg. https://doi.org/10.1101/2024.07.11.602962.

133. Alawneh, A.M., Qi, D., Yonesaki, T., and Otsuka, Y. (2016). An ADP-ribosyltransferase Alt of bacteriophage T4 negatively regulates the Escherichia coli MazF toxin of a toxin-antitoxin module. Mol. Microbiol. *99*, 188–198.

134. Otsuka, Y., and Yonesaki, T. (2012). Dmd of bacteriophage T4 functions as an antitoxin against Escherichia coli LsoA and RnlA toxins. Mol. Microbiol. *83*, 669–681.

135. 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*, 136–148.

136. Hilliard, J.J., Maurizi, M.R., and Simon, L.D. (1998). Isolation and characterization of the phage T4 PinA protein, an inhibitor of the ATP-dependent lon protease of Escherichia coli. J. Biol. Chem. *273*, 518–523.

137. Srikant, S., Guegler, C.K., and Laub, M.T. (2022). The evolution of a counter-defense mechanism in a virus constrains its host range. Elife *11*. https://doi.org/10.7554/eLife.79549.

138. Azulay, G., Sigal, N., Tzohar Rabinovich, K., Koenigsfeld, D., Stadnyuk, O., Das, A., Lisitzin, P., Biderman, E., Borovok, I., and Herskovits, A.A. (2025). A prophage-encoded anti-phage defense system that prevents phage DNA packaging by targeting the terminase complex. bioRxiv. https://doi.org/10.1101/2025.02.27.640495.

139. Yamashita, W., Chihara, K., Azam, A.H., Kondo, K., Ojima, S., Tamura, A., Imanaka, M., Nobrega, F.L., Takahashi, Y., Watashi, K., et al. (2025). Phage engineering to overcome bacterial Tmn immunity in Dhillonvirus. Commun. Biol. *8*, 290.

140. Azulay, G., Pasechnek, A., Stadnyuk, O., Ran-Sapir, S., Fleisacher, A.M., Borovok, I., Sigal, N., and Herskovits, A.A. (2022). A dual-function phage regulator controls the response of cohabiting phage elements via regulation of the bacterial SOS response. Cell Rep. *39*, 110723.

141. Bhoobalan-Chitty, Y., Xu, S., Martinez-Alvarez, L., Karamycheva, S., Makarova, K.S., Koonin, E.V., and Peng, X. (2024). Regulatory sequence-based discovery of anti-defense genes in archaeal viruses. Nat. Commun. *15*, 3699.

142. Costa, A.R., van den Berg, D.F., Esser, J.Q., van den Bossche, H., Pozhydaieva, N., Kalogeropoulos, K., and Brouns, S.J.J. (2025). Bacteriophage genomes encode both broad and specific counter-defense repertoires to overcome bacterial defense systems. Cell Host Microbe *33*, 1161–1172.e5.

143. Tal, N., Hadari, R., Chang, R.B., Osterman, I., Jacobson, R., Yirmiya, E., Bechon, N., Hochhauser, D., Rivera, M.L., Madhala, B., et al. (2025). Structural modeling reveals viral proteins that manipulate host immune signaling. bioRxivorg, 2025.07.12.664507. https://doi.org/10.1101/2025.07.12.664507.

144. Klemm, B.P., Singh, D., Smith, C.E., Hsu, A.L., Dillard, L.B., Krahn, J.M., London, R.E., Mueller, G.A., Borgnia, M.J., and Schaaper, R.M. (2022). Mechanism by which T7 bacteriophage protein Gp1.2 inhibits Escherichia coli dGTPase. Proc. Natl. Acad. Sci. U. S. A. *119*, e2123092119.

145. Loyo, C.L., and Grossman, A.D. (2025). A phage-encoded counter-defense inhibits an NAD-degrading anti-phage defense system. PLoS Genet. *21*, e1011551.

146. Chi, H., McMahon, S., Daniel-Pedersen, L., Graham, S., Gloster, T.M., and White, M.F. (2025). SAM-AMP lyases in type III CRISPR defence. Nucleic Acids Res. *53*, gkaf655.

147. Tesson, F., Huiting, E., Wei, L., Ren, J., Johnson, M., Planel, R., Cury, J., Feng, Y., Bondy-Denomy, J., and Bernheim, A. (2024). Exploring the diversity of anti-defense systems across prokaryotes, phages and mobile genetic elements. Nucleic Acids Res., gkae1171.

148. Yan, Y., Zheng, J., Zhang, X., and Yin, Y. (2024). dbAPIS: a database of anti-prokaryotic immune system genes. Nucleic Acids Res. *52*, D419–D425.

149. Dong, C., Hao, G.-F., Hua, H.-L., Liu, S., Labena, A.A., Chai, G., Huang, J., Rao, N., and

Guo, F.-B. (2018). Anti-CRISPRdb: a comprehensive online resource for anti-CRISPR proteins. Nucleic Acids Res. *46*, D393–D398.

150. Dong, C., Wang, X., Ma, C., Zeng, Z., Pu, D.-K., Liu, S., Wu, C.-S., Chen, S., Deng, Z., and Guo, F.-B. (2022). Anti-CRISPRdb v2.2: an online repository of anti-CRISPR proteins including information on inhibitory mechanisms, activities and neighbors of curated anti-CRISPR proteins. Database (Oxford) *2022*, baac010.

151. Wang, J., Dai, W., Li, J., Li, Q., Xie, R., Zhang, Y., Stubenrauch, C., and Lithgow, T. (2021). AcrHub: an integrative hub for investigating, predicting and mapping anti-CRISPR proteins. Nucleic Acids Res. *49*, D630–D638.

152. Burley, S.K., Bhikadiya, C., Bi, C., Bittrich, S., Chao, H., Chen, L., Craig, P.A., Crichlow, G.V., Dalenberg, K., Duarte, J.M., et al. (2023). RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. Nucleic Acids Res. *51*, D488–D508.

153. O'Brien, R.E., Santos, I.C., Wrapp, D., Bravo, J.P.K., Schwartz, E.A., Brodbelt, J.S., and Taylor, D.W. (2020). Structural basis for assembly of non-canonical small subunits into type I-C Cascade. Nat. Commun. *11*, 5931.

154. Lu, M., Yu, C., Zhang, Y., Ju, W., Ye, Z., Hua, C., Mao, J., Hu, C., Yang, Z., and Xiao, Y. (2024). Structure and genome editing of type I-B CRISPR-Cas. Nat. Commun. *15*, 4126.

155. Nagy, T.A., Gersabeck, G.W., Conte, A.N., and Whiteley, A.T. (2025). Identifying phage proteins that activate the bacterial innate immune system. bioRxivorg, 2025.07.02.662641. https://doi.org/10.1101/2025.07.02.662641.

156. Zhao, G., Chen, C., Xiong, W., Gao, T., Deng, Z., Wu, G., and He, X. (2016). Structural basis of the substrate preference towards CMP for a thymidylate synthase MilA involved in mildiomycin biosynthesis. Sci. Rep. *6*, 39675.

157. Shu, X., Wang, R., Cheng, F., Li, Z., Wu, A., Xue, Q., Liu, C., Cao, X., Feng, Y., and Li, M. (2024). CRISPR-Cas supervises diverse anti-phage defense systems. bioRxiv, 2024.12.27.630466. https://doi.org/10.1101/2024.12.27.630466.

158. Tuck, O.T., Adler, B.A., Armbruster, E.G., Lahiri, A., Hu, J.J., Zhou, J., Pogliano, J., and Doudna, J.A. (2024). Genome integrity sensing by the broad-spectrum Hachiman antiphage defense complex. Cell *187*, 6914–6928.e20.

159. Zhang, Z., Todeschini, T.C., Wu, Y., Kogay, R., Naji, A., Cardenas Rodriguez, J., Mondi, R., Kaganovich, D., Taylor, D.W., Bravo, J.P.K., et al. (2025). Kiwa is a membrane-embedded defense supercomplex activated at phage attachment sites. Cell *0*. https://doi.org/10.1016/j.cell.2025.07.002.

160. Nomburg, J., Doherty, E.E., Price, N., Bellieny-Rabelo, D., Zhu, Y.K., and Doudna, J.A. (2024). Birth of protein folds and functions in the virome. Nature *633*, 710–717.

161. 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*, 691–695.

162. Bähre, H., Hartwig, C., Munder, A., Wolter, S., Stelzer, T., Schirmer, B., Beckert, U., Frank, D.W., Tümmler, B., Kaever, V., et al. (2015). cCMP and cUMP occur in vivo. Biochem. Biophys. Res. Commun. *460*, 909–914.

163. Hobbs, S.J., and Kranzusch, P.J. (2024). Nucleotide immune signaling in CBASS, Pycsar, Thoeris, and CRISPR antiphage defense. Annu. Rev. Microbiol. *78*. https://doi.org/10.1146/annurev-micro-041222-024843.

164. Nair, P.A., Nandakumar, J., Smith, P., Odell, M., Lima, C.D., and Shuman, S. (2007). Structural basis for nick recognition by a minimal pluripotent DNA ligase. Nat. Struct. Mol. Biol. *14*, 770–778.

165. Sriskanda, V., Schwer, B., Ho, C.K., and Shuman, S. (1999). Mutational analysis of Escherichia coli DNA ligase identifies amino acids required for nick-ligation in vitro and for in vivo complementation of the growth of yeast cells deleted for CDC9 and LIG4. Nucleic Acids Res. *27*, 3953–3963.

166. Ho, C.K., Van Etten, J.L., and Shuman, S. (1997). Characterization of an ATP-dependent DNA ligase encoded by Chlorella virus PBCV-1. J. Virol. *71*, 1931–1937.

167. Chen, Y., Liu, H., Yang, C., Gao, Y., Yu, X., Chen, X., Cui, R., Zheng, L., Li, S., Li, X., et al. (2019). Structure of the error-prone DNA ligase of African swine fever virus identifies critical active site residues. Nat. Commun. *10*, 387.

168. Kerr, S.M., Johnston, L.H., Odell, M., Duncan, S.A., Law, K.M., and Smith, G.L. (1991). Vaccinia DNA ligase complements Saccharomyces cerevisiae cdc9, localizes in cytoplasmic factories and affects virulence and virus sensitivity to DNA damaging agents. EMBO J. *10*, 4343–4350.

169. Zhou, W., Richmond-Buccola, D., Wang, Q., and Kranzusch, P.J. (2022). Structural basis of human TREX1 DNA degradation and autoimmune disease. Nat. Commun. *13*, 4277.

170. Chan, M.P., Onji, M., Fukui, R., Kawane, K., Shibata, T., Saitoh, S.-I., Ohto, U., Shimizu, T., Barber, G.N., and Miyake, K. (2015). DNase II-dependent DNA digestion is required for DNA sensing by TLR9. Nat. Commun. *6*, 5853.

171. Yan, N., Regalado-Magdos, A.D., Stiggelbout, B., Lee-Kirsch, M.A., and Lieberman, J. (2010). The cytosolic exonuclease TREX1 inhibits the innate immune response to human immunodeficiency virus type 1. Nat. Immunol. *11*, 1005–1013.

172. Brzostek-Racine, S., Gordon, C., Van Scoy, S., and Reich, N.C. (2011). The DNA damage response induces IFN. J. Immunol. *187*, 5336–5345.

173. Morales, A.J., Carrero, J.A., Hung, P.J., Tubbs, A.T., Andrews, J.M., Edelson, B.T., Calderon, B., Innes, C.L., Paules, R.S., Payton, J.E., et al. (2017). A type I IFN-dependent DNA damage response regulates the genetic program and inflammasome activation in macrophages. Elife *6*. https://doi.org/10.7554/eLife.24655.

174. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res. *28*, 235–242.

175. Seo, H., Kim, S., Sagong, H.-Y., Son, H.F., Jin, K.S., Kim, I.-K., and Kim, K.-J. (2016).

Structural basis for cytokinin production by LOG from Corynebacterium glutamicum. Sci. Rep. *6*, 31390.

176. Quevillon-Cheruel, S., Campo, N., Mirouze, N., Mortier-Barrière, I., Brooks, M.A., Boudes, M., Durand, D., Soulet, A.-L., Lisboa, J., Noirot, P., et al. (2012). Structure-function analysis of pneumococcal DprA protein reveals that dimerization is crucial for loading RecA recombinase onto DNA during transformation. Proc. Natl. Acad. Sci. U. S. A. *109*, E2466–E2475.

177. Govande, A.A., Duncan-Lowey, B., Eaglesham, J.B., Whiteley, A.T., and Kranzusch, P.J. (2021). Molecular basis of CD-NTase nucleotide selection in CBASS anti-phage defense. Cell Rep. *35*, 109206.

178. Kranzusch, P.J., Wilson, S.C., Lee, A.S.Y., Berger, J.M., Doudna, J.A., and Vance, R.E. (2015). Ancient origin of cGAS-STING reveals mechanism of universal 2',3' cGAMP signaling. Mol. Cell *59*, 891–903.

179. Egorov, A.A., Hauryliuk, V., and Atkinson, G.C. (2024). Systematic annotation of hyper-variability hotspots in phage genomes and plasmids. bioRxiv, 2024.10.15.618418. https://doi.org/10.1101/2024.10.15.618418.

180. Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., et al. (2021). Protein complex prediction with AlphaFold-Multimer. bioRxiv, 2021.10.04.463034. https://doi.org/10.1101/2021.10.04.463034.

181. Wu, Y., Garushyants, S.K., van den Hurk, A., Aparicio-Maldonado, C., Kushwaha, S.K., King, C.M., Ou, Y., Todeschini, T.C., Clokie, M.R.J., Millard, A.D., et al. (2024). Bacterial defense systems exhibit synergistic anti-phage activity. Cell Host Microbe *32*, 557–572.e6.

182. Chong, C.E., Weimann, A., Agapov, A., Fothergill, J.L., Brockhurst, M.A., Parkhill, J., Floto, R.A., Szczelkun, M.D., Westra, E.R., Multi-Defence Consortium, et al. (2025). Defence systems drive accessory genome interactions in *Pseudomonas aeruginosa*. bioRxiv, 2025.05.06.652404. https://doi.org/10.1101/2025.05.06.652404.

183. Eddy, S.R. (1998). Profile hidden Markov models. Bioinformatics *14*, 755–763.

184. Paysan-Lafosse, T., Andreeva, A., Blum, M., Chuguransky, S.R., Grego, T., Pinto, B.L., Salazar, G.A., Bileschi, M.L., Llinares-López, F., Meng-Papaxanthos, L., et al. (2025). The Pfam protein families database: embracing AI/ML. Nucleic Acids Res. *53*, D523–D534.

185. van Kempen, M., Kim, S.S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C.L.M., Söding, J., and Steinegger, M. (2023). Fast and accurate protein structure search with Foldseek. Nat. Biotechnol. https://doi.org/10.1038/s41587-023-01773-0.

186. Moffat, L., and Jones, D.T. (2021). Increasing the accuracy of single sequence prediction methods using a deep semi-supervised learning framework. Bioinformatics *37*, 3744–3751.

187. Sayers, E.W., Cavanaugh, M., Clark, K., Ostell, J., Pruitt, K.D., and Karsch-Mizrachi, I. (2020). GenBank. Nucleic Acids Res. *48*, D84–D86.

188. Camargo, A.P., Nayfach, S., Chen, I.-M.A., Palaniappan, K., Ratner, A., Chu, K., Ritter, S.J., Reddy, T.B.K., Mukherjee, S., Schulz, F., et al. (2023). IMG/VR v4: an expanded

database of uncultivated virus genomes within a framework of extensive functional, taxonomic, and ecological metadata. Nucleic Acids Res. *51*, D733–D743.

189. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

190. Steinegger, M., and Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat. Biotechnol. *35*, 1026–1028.

191. Steinegger, M., and Söding, J. (2018). Clustering huge protein sequence sets in linear time. Nat. Commun. *9*, 2542.

192. Di Tommaso, P., Chatzou, M., Floden, E.W., Barja, P.P., Palumbo, E., and Notredame, C. (2017). Nextflow enables reproducible computational workflows. Nat. Biotechnol. *35*, 316–319.

193. Salazar, G.A., Luciani, A., Watkins, X., Kandasaamy, S., Rice, D.L., Blum, M., Bateman, A., and Martin, M. (2023). Nightingale: web components for protein feature visualization. Bioinform. Adv. *3*, vbad064.

194. Boratyn, G.M., Camacho, C., Cooper, P.S., Coulouris, G., Fong, A., Ma, N., Madden, T.L., Matten, W.T., McGinnis, S.D., Merezhuk, Y., et al. (2013). BLAST: a more efficient report with usability improvements. Nucleic Acids Res. *41*, W29–W33.

195. Lee, Y.-J., Dai, N., Walsh, S.E., Müller, S., Fraser, M.E., Kauffman, K.M., Guan, C., Corrêa, I.R., Jr, and Weigele, P.R. (2018). Identification and biosynthesis of thymidine hypermodifications in the genomic DNA of widespread bacterial viruses. Proc. Natl. Acad. Sci. U. S. A. *115*, E3116–E3125.

196. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods *17*, 261–272.

197. Azam, A.H., Kondo, K., Chihara, K., Nakamura, T., Ojima, S., Nie, W., Tamura, A., Yamashita, W., Sugawara, Y., Sugai, M., et al. (2024). Evasion of antiviral bacterial immunity by phage tRNAs. Nat. Commun. *15*, 9586.

198. Tesson, F., Planel, R., Egorov, A.A., Georjon, H., Vaysset, H., Brancotte, B., Néron, B., Mordret, E., Atkinson, G.C., Bernheim, A., et al. (2024). A comprehensive resource for exploring antiphage defense: DefenseFinder webservice,wiki and databases. Peer Community J. *4*. https://doi.org/10.24072/pcjournal.470.

199. Bryant, P., Pozzati, G., and Elofsson, A. (2022). Improved prediction of protein-protein interactions using AlphaFold2. Nat. Commun. *13*, 1265.

200. Tesson, F., Hervé, A., Mordret, E., Touchon, M., d'Humières, C., Cury, J., and Bernheim, A. (2022). Systematic and quantitative view of the antiviral arsenal of prokaryotes. Nat. Commun. *13*, 2561.

201. Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics *25*, 1422–1423.

202. Goldfarb, T., Kodali, V.K., Pujar, S., Brover, V., Robbertse, B., Farrell, C.M., Oh, D.-H., Astashyn, A., Ermolaeva, O., Haddad, D., et al. (2025). NCBI RefSeq: reference sequence standards through 25 years of curation and annotation. Nucleic Acids Res. *53*, D243–D257.

203. Couvin, D., Bernheim, A., Toffano-Nioche, C., Touchon, M., Michalik, J., Néron, B., Rocha, E.P.C., Vergnaud, G., Gautheret, D., and Pourcel, C. (2018). CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. Nucleic Acids Res. *46*, W246–W251.

204. UniProt Consortium (2025). UniProt: The universal protein knowledgebase in 2025. Nucleic Acids Res. *53*, D609–D617.

205. Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L.E., Brookes, D.H., Wilson, L., Chen, J., Liles, K., et al. (2018). Improvements to the APBS biomolecular solvation software suite. Protein Sci. *27*, 112–128.

206. Huo, Y., Nam, K.H., Ding, F., Lee, H., Wu, L., Xiao, Y., Farchione, M.D., Jr, Zhou, S., Rajashankar, K., Kurinov, I., et al. (2014). Structures of CRISPR Cas3 offer mechanistic insights into Cascade-activated DNA unwinding and degradation. Nat. Struct. Mol. Biol. *21*, 771–777.

207. Wang, S., Gabel, C., Siddique, R., Klose, T., and Chang, L. (2023). Molecular mechanism for Tn7-like transposon recruitment by a type I-B CRISPR effector. Cell *186*, 4204–4215.e19.

208. Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell *156*, 935–949.

209. Seo, H., and Kim, K.-J. (2017). Structural basis for a novel type of cytokinin-activating protein. Sci. Rep. *7*, 45985.

210. Li, Z., Jaroszewski, L., Iyer, M., Sedova, M., and Godzik, A. (2020). FATCAT 2.0: towards a better understanding of the structural diversity of proteins. Nucleic Acids Res. *48*, W60–W64.

211. Bouras, G., Nepal, R., Houtak, G., Psaltis, A.J., Wormald, P.-J., and Vreugde, S. (2023). Pharokka: a fast scalable bacteriophage annotation tool. Bioinformatics *39*, btac776.

212. Terzian, P., Olo Ndela, E., Galiez, C., Lossouarn, J., Pérez Bucio, R.E., Mom, R., Toussaint, A., Petit, M.-A., and Enault, F. (2021). PHROG: families of prokaryotic virus proteins clustered using remote homology. NAR Genom. Bioinform. *3*, lqab067.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT OR RESOURCE	SOURCE	IDENTIFIER
Deposited Data		

FTP with data	This study	https://ftp.ebi.ac.uk/pub/databases/metagenomics/research-team/projects/evades/
EVADES database	This study	https://www.ebi.ac.uk/finn-srv/evades
Code with analysis	This study	https://github.com/Khalimat/EVADES
Software and Algorithms		
AlphaFold 3 v3.0.1	Abramson et al., 2024	https://github.com/google-deepmind/alphafold3/
FoldSeek (release 10-941cd33)	Kempen et al., 2023	https://github.com/steineggerlab/foldseek
Consurf	Yariv et al., 2023	https://consurf.tau.ac.il/consurf_index.php
PyMOL 3.1.4.1	Schrödinger, LLC.	https://www.PyMOL.org/
Blast 2.12.0+	Altschul et al., 1990	https://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/
S4PRED	Moffat et al., 2021	https://github.com/psipred/s4pred
HMMER 3.4	Eddy, 1990	http://hmmer.org/
Python 3.8.17		
NextFlow	Tommaso et al., 2017	https://www.nextflow.io/
Django		https://github.com/django/django
APBS 3.4.1	Jurrus et al., 2017	https://www.poissonboltzmann.org/
BioPython 1.79	Cock et al., 2009	https://github.com/biopython/biopython
SQLite	_	https://sqlite.org/
jquery.dataTables	_	https://datatables.net/
DefenseFinder v2.0.0	Tesson et al. 2024	https://github.com/mdmparis/defense-finder
iLund4u	Egorov et al., 2024	https://art-egorov.github.io/iLund4u/
FATCAT 2.0 (local version)	Li et al., 2020	https://github.com/GodzikLab/FATCAT-dist
nightingale		https://github.com/open-contracting/nightingale

feature-viewer v1.1.0	Paladin et al., 2020	https://github.com/calipho-sib/feature-viewer
MAFFT v7.520	Katoh et al., 2013	https://mafft.cbrc.jp/alignment/software/
MMseqs 13-45111	Steinegger et al., 2013	https://github.com/soedinglab/MMseqs2
matplotlib 3.4.3	_	https://matplotlib.org/
pdbe-molstar v3.1.0	_	https://github.com/molstar/pdbe-molstar
ete3 (v.3.1.3)	Huerta-Cepas et al., 2016	https://github.com/etetoolkit/ete
pyvis 0.3.2	_	https://github.com/WestHealth/pyvis
Other		
PDB	Burley et al., 2023	https://www.rcsb.org/
GenBank	Sayers et al., 2020	https://www.ncbi.nlm.nih.gov/genbank/
RefSeq	Goldfarb et al., 2025	https://www.ncbi.nlm.nih.gov/refseq/
UniProt (Release 2025_02)	The UniProt Consortium, 2023	https://www.uniprot.org/
Pfam 37.2	Paysan-Lafosse et al., 2024	https://www.ebi.ac.uk/interpro/download/Pfam/
SciPy 1.10.1	Virtanen et al., 2020	https://github.com/scipy/scipy
IMG/VR 7.1	Camargo et al., 2023	https://img.jgi.doe.gov/vr/

METHOD DETAILS

Data collection

We analysed research literature reporting phage ADPs and extracted information on the defence systems they counteract, MoAs, PDB accessions, data on their potential to trigger other defence systems, and protein sequences. We exclude entries where protein sequences were not provided in the articles, supplementary material or as identifiers in protein or gene databases. For ADPs that function as part of multicomponent systems, we collected and cross-referenced relevant entries accordingly. Additionally, we assigned Evidence and Conclusion Ontology (ECO)[8] labels to describe evidence supporting entries in our database. We categorised entries in our database based on MoA.

We ran hmmer v3.4[183] (with option --cut_ga) with Pfam 37.2[184] to identify existing profile HMMs in Pfam corresponding to ADPs. We produced structural models using AlphaFold 3[38] for ADPs without a corresponding entry in the PDB. We used AlphaFold 3 v3.0.1 and generated four seeds (1, 2, 3 and 4) through this work. For ADPs without characterised MoA, we searched for structural homolog using FoldSeek[185] in the pregenerated target databases bundled with FoldSeek (PDB, Alphafold/UniProt and ESMAtlas30, e-value < 0.00001). We used S4PRED to predict protein secondary structure based on the ADP sequences[186]. We used FoldSeek release 10-941cd33 and the following command for the identification of homologous structures through this work: *foldseek easy-search [query_structure] [target_database] [output_file]*.

We downloaded three databases to identify homologs of ADPs: (i) the GenBank database of phage genomes (release from 30-04-2025)[187], (ii) the GenBank database of viral genomes (release from 30-04-2025), and (iii) IMG/VR version 7.1[188]. We used BLASTp (v2.12.0+)[189] searches to identify homologs, applying the following criteria to determine significance: e-value < 1^{e-5}, query coverage ≥ 80%, and identity ≥ 30%. Additionally, we extracted available phylogenetic information about genomes encoding homologs of ADPs using python package ete3 (v.3.1.3) and from metadata provided in IMG/VR.

We combined the hits obtained from these three databases, added the initial query sequence, de-replicated the sequences using CD-HIT version 4.8.1 at 90% identity and 90% coverage, aligned them using MAFFT version 7.520, manually trimmed the MSAs using Belvu version 4.44.1, and created profile HMMs using *hmmbuild* from the HMMER package. The resulting profile HMM library is available via FTP.

Comparison with other resources

To compare ADP coverage in EVADES, dbAPIS[148], and Anti-Defense Finder[147], we combined all sequences from these databases, dereplicated them using MMseqs2 (v13.45111, --cov-mode 0)[190,191] at 90% identity and 90% coverage, and assessed the presence of sequences or closely related homologs (defined as sequences sharing 90% identity over 90% of the query length) in all three databases. Additionally, we calculated the number of ADPs per bacterial defense family and aggregated the results to assess how well EVADES, dbAPIS, and Anti-Defense Finder represent the diversity of ADPs.

Creation of the web resource

The web resource was developed using the Python Django framework, and the presented data and metadata are stored in an SQLite database. We created a Nextflow[192] pipeline to automate SQLite database generation from a json-file with the database. For the visualisations, the following JavaScript libraries were utilised: pdbe-molstar v3.1.0 for the 3D structure viewer, feature-viewer@v1.1.0 for the predicted secondary structure viewer, jQuery DataTables for Pfam annotations, and the Nightingale library[193] for the interactive Pfam annotations viewer.

Sequence similarity comparison

We generated a BLASTp database from all ADP sequences, we conducted homology searches using individual ADP sequences as queries and an e-value threshold of 0.001, and filtered out self-matches. All sequence pairs which passed the e-value threshold, except for two (Lidtsur 6 - Forsur 7 and AcrIIA1 - AcrIB5), exceeded the empirical threshold of 25% identity used to

establish homology. To further investigate these exceptions, we used FATCAT 2.0 to assess structural similarity. We used the local version of FATCAT 2.0 and the following command for all pairwise structure similarity comparisons throughout this work: *FATCAT -p1 [pdb_file_1] -p2 [pdb_file_1] -o [pdb_file_1-pdb_file_2] -m -ac -t.* The resulting p-values were: 4.65^{e-12} for Lidtsur 6 - Forsur 7 and 4.38^{e-07} for AcrIIA1 - AcrIB5, so we included them as pairs sharing homology.

Similarity with the defence triggers

We compared ADP sequences from our collection with the triggers of bacterial defence systems[28] using BLASTp (e-value < 0.001)[194] and analysed the results in order to find new connections to ADPs. We identified multiple matches between the mga47 protein, which inhibits DarTG defence, and several putative Borvo triggers, showing strong sequence similarity. The highest-scoring match exhibited 100% identity across the full length of both sequences. We also observed similarities between promiscuous kinases from phages T7 and JSS1 and putative ShosTA triggers, with the T7 kinase being identical to one of the putative ShosTA triggers. Then we explored the Avs type I trigger (NCBI GenPept accession: WP_317906625.1) from *Bacillus* phage SBSphiC, as it annotated in the original paper as dUMP hydroxymethylase and was not reported as an ADP, but we speculated it may function as a protein involved in DNA hypermodification. We predicted its 3D structure using AlphaFold 3 using the same parameters as previously specified[38], and searched for homologs in PDB[174] using FoldSeek[185]. The closest homolog identified was CMP hydroxymethylase MilA (PDB accession: 5JNH). We aligned the structure of MilA with the Avs type I trigger using the local version of FATCAT 2.0 (p-value 0.00^{e+00}). Residues K133 and A176 define MilA's substrate specificity; it has been shown that the mutations K133R and A176S change substrate specificity from CMP to dCMP[156]. We analysed the alignment between the Avs type I trigger and CMP hydroxymethylase MilA in PyMOL (version 3.1.4.1, used throughout this work), identifying the closest residues in the Avs type I trigger that align with those determining substrate specificity in MilA. It has previously *Bacillus* phage SP10, which is closely related to SBSphiC, encodes a dUMP hydroxymethylase involved in DNA hypermodification[195] (NCBI GenPept accession: YP_007003445.1), which shares 97% sequence identity with the Avs type I trigger. This suggested that dUMP may also serve as a substrate of the type I Avs trigger. To explore this, we co-folded the Avs type I trigger with dUMP, dCMP, and CMP using AlphaFold 3[38], as ipTM scores for complexes can be indicative of the substrate preference.

We also compared the structures in EVADES to triggers of bacterial defence systems reported by Nagy *et al.* [155] using Foldseek. We created query and target databases with the command *foldseek createdb*, performed the comparison using *foldseek search*, converted the output to a TSV file with *foldseek convertalis,* and filtered the results using an e-value threshold of 0.01. However, we did not identify any matches apart from Gam protein, which was previously reported as a trigger and.

We calculated the number of homologs per protein for ADPs known to trigger bacterial defense systems and for non-triggers in the viral and phage fractions of GenBank and IMG/VR (see the Data Collection section). We then compared the two samples using the Mann–Whitney U test from the SciPy package (version 1.10.1)[196] and calculated the means using the *statistics.mean* function in Python.

Prediction of ADPs inhibiting multiple defence systems based on protein sharing network for defence systems

We compiled information from research articles on defence proteins known to interact directly with ADPs. Although direct interactions between anti-Septu proteins (Anhur and Svarog[18]) and the Septu proteins have not been reported, we included PtuA and PtuB in our analysis due to their presence and conservation between Septu defence and type 1A retrons (Ec78, Ec83, Vc95)[197]. Some ADPs with uncharacterised MoA inhibit defence systems composed of a single protein, suggesting that those defence proteins are likely to be targeted—directly or indirectly—by ADPs. Therefore, we also included these defence proteins in our analysis. A full list of the defence proteins used in this experiment is provided in Supplementary Table 2.

We retrieved the structures of these defence proteins from DefenseFinder[198] (monomer structures) and the PDB, or predicted them if not available in either source. All monomer structures available in DefenseFinder[198] were included and supplemented with structures of defence proteins from the previous experiment (see Supplementary Table 2) to build a reference database of defence proteins. We used the command *foldseek createdb*[185] to construct the target database. Structural homologues of ADP-binding proteins were identified using *foldseek easy-search*[185], applying a stringent e-value cut-off ($<1^{e-8}$) to prioritise functionally relevant structural matches. We then merged the resulting tables, removed redundancy (e.g. multiple entries for the same protein across different defence system subtypes), and visualised the results using the Pyvis 0.3.2 library in Python.

To validate selected structural matches, we used AlphaFold 3 to predict complex formation, with a particular focus on interactions between PglX homologs and Ocr. First, we predicted the structure of the PglX-Ocr complex and assessed whether the ipTM score exceeded the recommended threshold of 0.6. We also predicted interactions between Ocr and four other defence proteins: gp28 (Panchino), drmMII (DISARM), Ecoprrl (RM), and HsdM (RM). While predictions for Ecoprrl (ipTM = 0.42, pTM = 0.75) and HsdM (ipTM = 0.3, pTM = 0.72) did not meet the threshold, AlphaFold 3 produced moderate-confidence predictions for the Ocr–gp28 complex (ipTM = 0.79, pTM = 0.79, pDockQ = 0.56) and the Ocr–drmMII complex (ipTM = 0.72, pTM = 0.87, pDockQ = 0.401). To calculate the pDockQ score, we implemented a custom script capable of processing cif-files generated by AlphaFold 3. The pDockQ metric was proposed by Bryant *et al.*[199] as a measure of protein-protein complex prediction quality, with a recommended threshold of pDockQ $\geq$ 0.23. We visualised both predicted complexes in PyMOL, highlighting residues in gp28 and drmMII that interact with Ocr, defined as residues within 8 Å of any Ocr residue.

Prediction of direct interactions between ADPs with uncharacterised mechanisms of action and corresponding defence proteins

First, we obtained the sequences of the defence systems which were inhibited by uncharacterised proteins. We downloaded all monomer PDB-files provided by the DefenseFinder database[198,200], converted them to fasta-format using BioPython 1.79[201], and removed duplicate entries. As sequences for CRISPR and RM proteins were absent from the DefenseFinder database, we searched for them in RefSeq[202], CRISPRCasFinder[203], UniProt[204] and PDB[152]. We looked for defence protein sequences with taxonomic annotations corresponding to species specified in the original papers reporting ADPs. We excluded the O-antigen-based barrier, as the polysaccharide chains are highly variable and cannot be reliably modelled. We also excluded ADPs with uncharacterised MoA shown to act indirectly—such as ArdB and KlcA—or are hypothesised to act as negative regulators of anti-defence systems (e.g.

ArdK, ArdR, and DdrB). The complete table of sequences for co-folding experiments is provided in Supplementary Table 3. Secondly, we run the AlphaFold 3[38] v3.0.1 data pipeline (*--run_inference=false*) on individual defence and anti-defence protein sequences to generate MSA and structural templates. This approach allowed reuse of MSA/template data for proteins appearing multiple times. Third, we created JSON files combining the MSAs and templates from the previous step for each corresponding defence and anti-defence protein pair. We then ran AlphaFold 3[38] inference with 4 seeds (1, 2, 3, 4), producing 20 models and focused only on those complexes where both the average ipTM and the best model's ipTM were greater than 0.6. We then computed the pDockQ scores for the best models. Predicted complexes were classified into high-confidence (best model ipTM ≥ 0.8) and moderate-confidence (0.6 < best model ipTM < 0.8). All results are available in Supplementary Table 4.

Putative MoA for AcrIC5

We predicted a high-confidence (ipTM = 0.92, pTM = 0.85, pDockQ = 0.56) complex between AcrIC5 and *Pseudomonas aeruginosa* Cas8c (*Pa*Cas8c, NCBI GenPept accession: UEM35121.1). To gain mechanistic insights from the predicted complex between AcrIC5 and *Pa*Cas8c, we sought to annotate the region in Cas8c that binds AcrIC5. We looked for the closest homolog in the PDB using FoldSeek[185] and found a Cascade complex containing a Cas8c homolog from *Desulfovibrio vulgaris* (PDB accession: 7KHA)[153]. We aligned it using FATCAT 2.0 with the structure of *Pa*Cas8c predicted with AlphaFold 3[38] and found that the binding site for AcrIC5 is extensive and includes the putative PAM-binding site (residues: GVDAKGK). We visualised the results in PyMOL, colouring AcrIC5 based on electrostatic potential (using Adaptive Poisson-Boltzmann Solver, APBS, Electrostatics module[205]).

Putative MoA for AcrIC3

We predicted a high-confidence complex (ipTM = 0.91, pTM = 0.90, pDockQ = 0.633) between AcrIC3 and *Pseudomonas aeruginosa* Cas3 (*Pa*Cas3, NCBI GenPept: UEM35119.1). To annotate *Pa*Cas3 protein, we took its structure predicted by AlphaFold 3[38] and looked for structural homologs in PDB using FoldSeek[185] . We identified Cas3 from *Thermobifida fusca* YX (PDB accession: 4QQZ) as a closest homolog (e-value = 1.74^{e-18})[206]. Then we aligned the structures using FATCAT 2.0, and relying on the domain coordinates reported for *Thermobifida fusca'*s Cas, annotated domains (1-216 HD, 217-430 RecA1, 431-602 RecA2, 603-649 Linker, and 650-744 CTD) in *Pa*Cas3. We visualised the complex between AcrIC3 and *Pa*Cas3 in PyMOL. We used the BioPython 1.79[201] module NeighborSearch to identify residues in Cas3 likely interacting with residues in AcrIC3 (defined as residues in 8Å or less). We used matplotlib (v.3.4.3) to visualise the domain architecture, plotting grey dots on residues in *Pa*Cas3 interacting with AcrIC3.

Putative MoA for AcrIB4

We predicted a high-confidence complex (ipTM = 0.91, pTM = 0.76, pDockQ = 0.424) between AcrIB4 and *Listeria seeligeri* Cas8 (*Ls*Cas8b1) with its N-terminal residues (504–530) forming the binding interface. To annotate the *Ls*Cas8b1 binding interface, we took its AlphaFold 3[38] predicted structure and looked for structural homologs in the PDB using FoldSeek[185] . However, as we did not identify any matches, we manually looked for type I-B Cascade complexes containing Cas8 in the PDB and found a complex (PDB accession: 8FCJ) where Cas8 protein

shares high similarity to *Ls*Cas8b1[207]. We found that the Cas8 residues corresponding to the AcrIB4 binding interface in *Ls*Cas8b1 are involved in binding Cas11 in the type I-B Cascade. We visualised the complexes between AcrIB4-*Ls*Cas8b1 and Cas8-Cas11 in PyMOL.

Putative MoA for AcrIIA26

We predicted a high-confidence complex (ipTM = 0.87, pTM = 0.79, pDockQ = 0.634) between AcrIIA26 and *Spy*Cas9 (NCBI GenPept accession: QHB64844.1). We wanted to explore the positioning of AcrIIA26 relative to sgRNA and target DNA, so we took sgRNA and target DNA sequences from a PDB entry with accession 4OO8[208], co-folded them with *Spy*Cas9 using AlphaFold 3[38] and predicted a very high-confidence complex (ipTM = 0.94, pTM = 0.92). We positioned the AcrIIA26-*Spy*Cas9 complex and the complex between sgRNA, target DNA and *Spy*Cas9 similarly in space using PyMOL, which allowed us to infer that AcrIIA26 occludes the DNA binding site. We calculated electrostatic potential on AcrIIA26 using APBS Electrostatics module in PyMOL and visualised its surface interacting with *Spy*Cas9.

Structure based search for anti-defence proteins in eukaryotic viral protein database

We conducted structure-based homology search for phage ADPs in a eukaryotic viral protein database[160] using FoldSeek[185] (considering the size of DB, 68K entries, we filtered on e-value < 0.00001). Then, we analysed the results and focused on matches with ADPs counteracting defence systems shared between bacteria and eukaryotes, namely anti-Avs proteins ORF83 and ORF55[17] and NTase[53] inhibiting TIR and STING immune response.

ORF55 functional annotation and homologs in eukaryotic viruses

We looked for homologs of ORF55 using its structural model in PDB[152] with FoldSeek[185] and found that the closest match is the DNA ligase (PDB accession: 2Q2U) from *Chlorella* virus (eukaryotic DNA viruses). We superimposed ORF55 and the DNA ligase from *Chlorella* virus using FATCAT 2.0 (p-value = 0.00^{e+00}) to obtain a more accurate estimate of similarity. We analysed the paper reporting the structure of the DNA ligase from the *Chlorella* virus and retrieved active site residues responsible for its catalytic activity. Then we used the superimposed complex of ORF55 with the *Chlorella* virus DNA ligase obtained using FATCAT, analysed the alignment in PyMOL, and retrieved the residues in ORF55 aligning with the active site residues in the *Chlorella* virus DNA ligase.

We modelled interaction between ORF55 and the DNA fragment from the *Chlorella* virus DNA ligase using AlphaFold 3[38] and predicted a high confidence complex. Next we used the same DNA fragment, but introduced a nick, and one- and two-nucleotide gapes and co-folded with ORF55 and ATP, where similar to the *Chlorella* virus DNA ligase, ORF55 showed preference for nicked (ipTM = 0.95, pTM = 0.95) and one-nucleotide gapped DNA (ipTM = 0.95, pTM = 0.95) compared to two-nucleotide gapped DNA (ipTM = 0.87, pTM = 0.91).

ORF83 functional annotation and homologs in eukaryotic viruses

We looked for structural homologs of ORF83 in the PDB database using FoldSeek[185] , aligned the top four matches using FATCAT 2.0 and identified *Cg*1261 (PDB ID: 5WQ3, 4.39^{e-7}) – the type II cytokine activating protein from LOG family (The Lonely Guy) as the highest scoring match. We analysed the paper with functional characterisation of *Cg*1261 and found it was reported to hydrolyze AMP into an adenine base and ribose 5-phosphate. So, we modelled

interaction between ORF83 and AMP using AlphaFold 3[38]. Then we aligned the ORF83-AMP complex with *Cg*1261 using FATCAT 2.0, selected active site residues in *Cg*1261 from the paper[209] and found closest aligning residues in ORF83. Then we predicted the structure of ORF83's homolog from eukaryotic dsDNA viruses (NCBI GenPept accession: YP_009506880) in complex with AMP, aligned it with ORF83 using FATCAT 2.0, and visualised the results in PyMOL.

NTase

We aligned the structural model of *Bacillus* phage Bcp1 NTase with its homolog we identified in *Erinnyis ello* granulovirus (NCBI GenPept accession: YP_009091875.1) using FATCAT 2.0 (p-value = 2.74^{e-10})[210] and visualised the results in PyMOL. Next, we calculated the number of positions in the phage NTase and its homolog from *Erinnyis ello* granulovirus matching the catalytic motif of phage NTases and visualised the results.

We aimed to profile substrate specificity of phage and viral NTases, so we modelled[38] the complexes between one of the NTases and pairs of NTPs (we only used ATP, GTP, or UTP, as they are typically utilised by CD-NTases) and 3 ions of Mg^{2+}, which have been reported as cofactors stabilising donor and acceptor NTPs[177]. We visualised the ipTM and pTM scores as heatmaps. We used PyMOL to visualise the hydrogen bonds formed between *Erinnyis ello* granulovirus NTase, two molecules of UTP and three ions of Mg^{2+}.

Prediction of putative ADPs from diversity hotspots

Diversity hotspots are genomic regions characterised by elevated variability, typically flanked by conserved genes, and frequently observed in phage and plasmid genomes. These regions arise through mechanisms such as mobile genetic element insertions, recombination, and hypermutation, and often carry accessory genes involved in defence, anti-defence, virulence, and antimicrobial resistance[179]. We ran iLund4u to identify ADPs encoded such diversity hotspots in phages (*iLund4u protein -fa [adp].fa -db phages -o [adp]_phages*) using the default parameters. iLund4u uses a tool named Pharokka[211] to annotate proteins from diversity hotspots replying on profile HMMs derived from the database Prokaryotic Virus Remote Homologous Groups (PHROG)[212]. We selected all uncharacterised proteins based on iLund4u results from phage diversity hotspots that encode ADPs in this analysis. As the iLund4u results also contain information about cluster membership for proteins, we selected protein clusters containing more than 50 members. These clusters were further refined by re-clustering the proteins at 40% sequence identity over 80% of their length. To exclude proteins with other known functional annotations, we used HMMER with Pfam version 37.2 and the *--cut_ga* option on cluster representatives. These steps enabled us to focus on uncharacterised proteins from diversity hotspots for structural modelling. We first modelled individual phage proteins using a local version of AlphaFold 3**,** generating four models with seeds 1, 2, 3, and 4. Proteins with an average predicted local Distance Difference Test (pLDDT) score below 70 were filtered out.

Next, we performed a three-step co-folding process to identify potential interactions between phage proteins and bacterial defence proteins.We aimed to select pairs of proteins that consistently showed high confidence scores across different seed numbers, so we randomly generated a different seed number for each step:

1. Initial Co-folding:
   We co-folded all possible phage–bacterial protein pairs (255 × 608 = 155,040 combinations) using local AlphaFold 3 with seed 203490234, generating five models per pair. We retained pairs where at least three of the five models had a predicted

confidence score (0.8 × ipTM + 0.2 × pTM) above 0.8.

2. Server-based Co-folding:
   The shortlisted pairs were co-folded again using the AlphaFold 3 server (between 27 May and 20 June 2025) with seed 349024. We applied the same filtering criterion: a confidence score above 0.8 in at least three out of five models.

3. Final Co-folding with Multiple Seeds:
   Pairs passing the second step were modelled again locally with AlphaFold 3 using five different seeds (3294, 938934343, 778343, 343, and 53343). We selected those with an average confidence score above 0.8 in at least 17 out of the 25 total models.

Finally, for each selected protein pair, we calculated the pDockQ score and retained only those with a pDockQ score ≥ 0.23.

QUANTIFICATION AND STATISTICAL ANALYSIS

The p-values and e-values produced by tools are described in the text with references to original articles.

We used the Mann-Whitney U-test to compare the distributions in Figure 3B, and the Fisher exact test to calculate the Odds ratio and p-value in Figure 3C.

We calculated correlation between ipTM scores for phage and viral NTase using Pearson correlation coefficient.