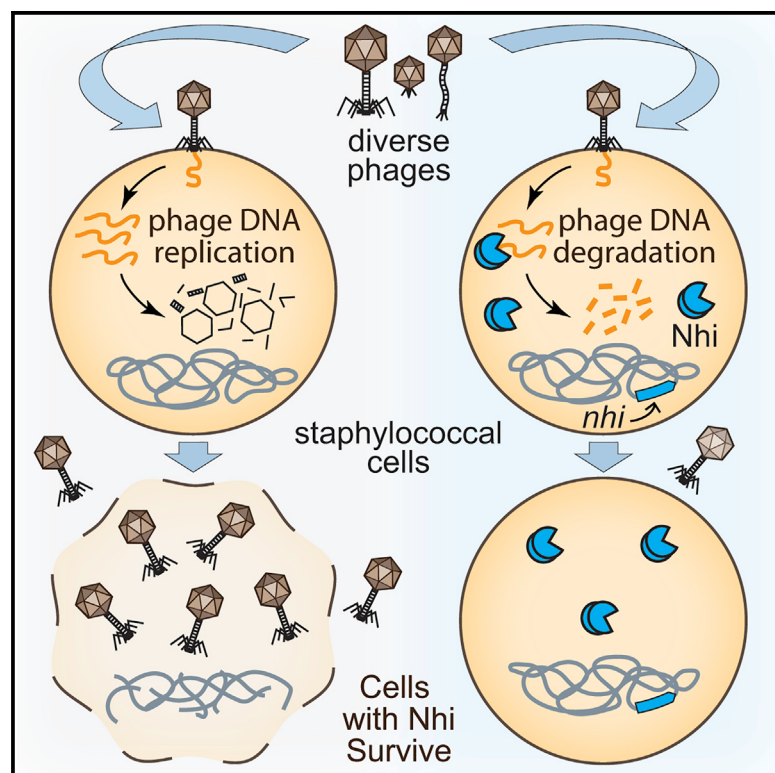


Cell Host & Microbe

A unique mode of nucleic acid immunity performed by a multifunctional bacterial enzyme

Graphical abstract



Authors

S.M. Nayeemul Bari,
Lucy Chou-Zheng, Olivia Howell, ...,
Alexander Thomas, Barbaros Aslan,
Asma Hatoum-Aslan

Correspondence

ahatoum@illinois.edu

In brief

Bacteria have evolved a variety of immune systems that use multiple components to detect and destroy the nucleic acids of their viral parasites (i.e., phages). Bari et al. report the discovery of a unique mode of immunity mediated by a single enzyme called NHI, which targets and degrades phage DNA.

Highlights

- NHI is a nuclease-helicase that confers immunity against diverse phages
- NHI blocks phage DNA accumulation without causing abortive infection
- Phage replication machinery plays a critical role in NHI specificity
- NHI homologs from diverse bacteria exhibit functional conservation



Article

A unique mode of nucleic acid immunity performed by a multifunctional bacterial enzyme

S.M. Nayeemul Bari,^{1,3} Lucy Chou-Zheng,^{1,3} Olivia Howell,¹ Motaher Hossain,¹ Courtney M. Hill,¹ Tori A. Boyle,¹ Katie Cater,² Vidya Sree Dandu,² Alexander Thomas,² Barbaros Aslan,¹ and Asma Hatoum-Aslan^{1,4,*}

¹Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, IL 61821, USA

²Department of Biological Sciences, University of Alabama, Tuscaloosa, AL 35487, USA

³These authors contributed equally

⁴Lead contact

*Correspondence: ahatoum@illinois.edu

<https://doi.org/10.1016/j.chom.2022.03.001>

SUMMARY

The perpetual arms race between bacteria and their viruses (phages) has given rise to diverse immune systems, including restriction-modification and CRISPR-Cas, which sense and degrade phage-derived nucleic acids. These complex systems rely upon production and maintenance of multiple components to achieve antiphage defense. However, the prevalence and effectiveness of minimal, single-component systems that cleave DNA remain unknown. Here, we describe a unique mode of nucleic acid immunity mediated by a single enzyme with nuclease and helicase activities, herein referred to as Nhi (nuclease-helicase immunity). This enzyme provides robust protection against diverse staphylococcal phages and prevents phage DNA accumulation in cells stripped of all other known defenses. Our observations support a model in which Nhi targets and degrades phage-specific replication intermediates. Importantly, Nhi homologs are distributed in diverse bacteria and exhibit functional conservation, highlighting the versatility of such compact weapons as major players in antiphage defense.

INTRODUCTION

Phages are the most abundant entities in the biosphere (Bergh *et al.*, 1989), and as such, they impose a tremendous selective pressure upon their bacterial hosts. Phages attach to a specific host, inject their genetic material, and utilize the host's enzymes and energy stores to replicate exponentially in a process that typically leads to cell lysis and death. In response to this constant threat, bacteria have evolved an impressive collection of immune systems that undermine nearly every step of the phage infection cycle (Hampton *et al.*, 2020). Such systems may block phage genome entry, interfere with phage DNA replication/expression, and/or, as a last resort, precipitate programmed cell death, a process known as abortive infection (Abi), to prevent phages from spreading to neighboring bacteria in the population (Lopatina *et al.*, 2020). Abi can be achieved through a variety of mechanisms and constitutes a remarkably common defense strategy. Indeed, recent years have witnessed a surge of reports on new bacterial immune systems (Cohen *et al.*, 2019; Doron *et al.*, 2018; Gao *et al.*, 2020; Kronheim *et al.*, 2018; Millman *et al.*, 2020), and many of these ultimately cause cell death. Notable examples utilize RNA-modifying enzymes (Gao *et al.*, 2020), retrons (Gao *et al.*, 2020; Millman *et al.*, 2020), and small molecules (Cohen *et al.*, 2019) as the basis for defense.

As a more direct and perhaps more effective approach to stemming a phage infection, bacteria employ defenses that

sense and destroy phage genetic material. Such systems exhibit a range of complexities, from the simpler restriction-modification (RM) to the more sophisticated adaptive immune systems that rely upon clusters of regularly interspaced short palindromic repeats (CRISPRs) and CRISPR-associated (Cas) genes (Hampton *et al.*, 2020). RM systems are innate immune systems that require at least two components for defense—a nuclease that cleaves specific DNA sequences and a DNA methyltransferase that modifies and protects the host genome from cleavage. Further, more complex RM-like systems have recently been described, which use methyltransferases in conjunction with a variety of proteins such as proteases, phosphatases, and phospholipases to provide other necessary functionalities for defense (Goldfarb *et al.*, 2015; Gordeeva *et al.*, 2019; Hoskisson *et al.*, 2015; Ofir *et al.*, 2018; Sumby and Smith, 2002). CRISPR-Cas systems are even more elaborate—they integrate short stretches of phage-derived nucleic acids into the CRISPR locus, which in turn are used to generate small RNAs that combine with Cas nucleases to identify and eliminate complementary phage sequences (Hille *et al.*, 2018). CRISPR-Cas systems are remarkably diverse (Makarova *et al.*, 2020), and different types have been shown to work together (Deng *et al.*, 2013; Hoikkala *et al.*, 2021; Pinilla-Redondo *et al.*, 2020; Silas *et al.*, 2017) and even synergize with RM systems (Dupuis *et al.*, 2013) to ensure a more effective defense. Such added layers have obvious advantages in protecting against diverse and evolving phage



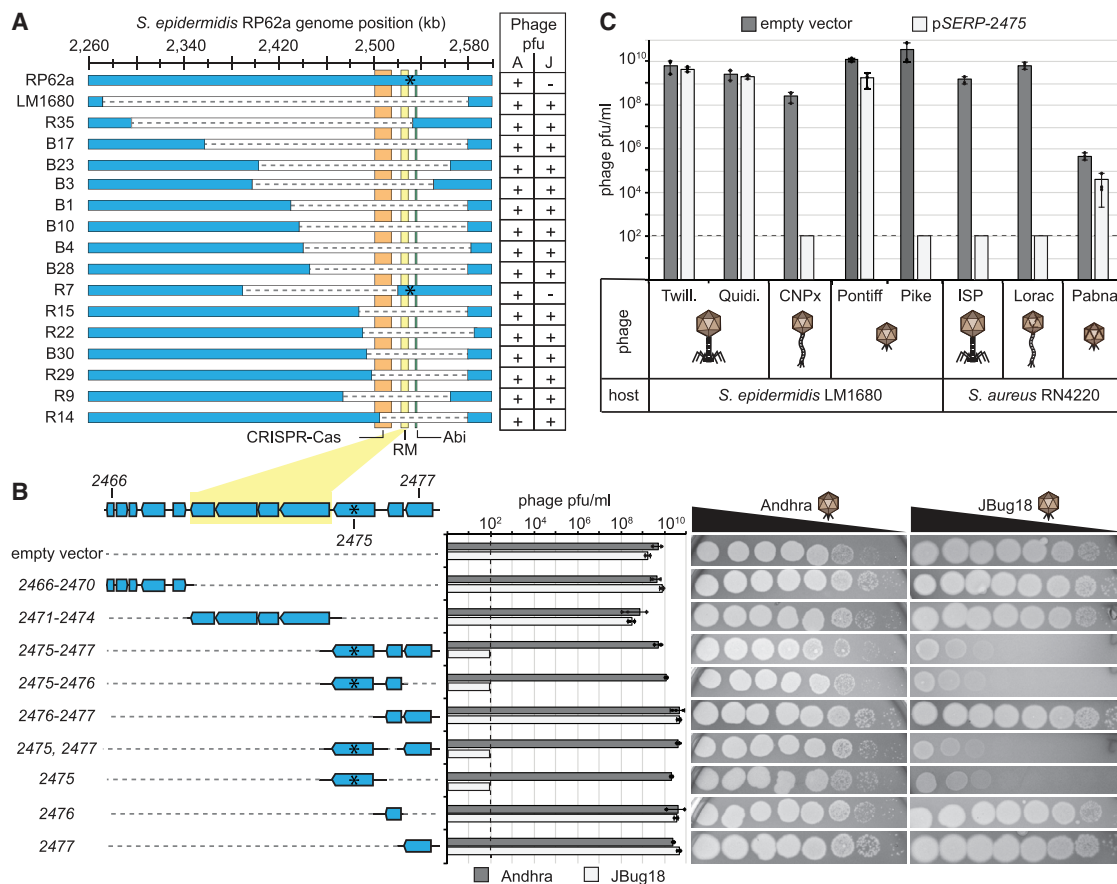


Figure 1. *SERP2475* provides robust immunity against diverse staphylococcal phages

(A) A segment of the *S. epidermidis* RP62a genome and deletion mutants. Dashes indicate deleted regions, and regions encoding CRISPR-Cas, RM, and Abi systems are highlighted. Strains were challenged with Andhra (A) and JBug18 (J), and resulting plaque-forming units per milliliter (pfu/mL) are indicated: +, $\sim 1 \times 10^9$ pfu/mL; -, 0 pfu/mL. Asterisks mark the location of *SERP2475*.

(B) Magnified view of the genomic region responsible for immunity and corresponding plasmids (pSERP-) containing them. *S. epidermidis* LM1680 strains harboring indicated plasmids were challenged with 10-fold dilutions of Andhra and JBug18 (1×10^0 to 1×10^{-7}), and resulting pfu/mL are shown as an average of triplicate measurements ($\pm$ SD). Representative plate images are also shown.

(C) *S. epidermidis* and *S. aureus* strains were challenged with indicated phages, and resulting pfu/mL are shown. Dotted line indicates the limit of detection, and short bars underneath indicate 0 pfu/mL. See also Figure S1.

predators; however, this strategy comes with the energetic cost of producing and maintaining multiple components as well as the risk that damage or loss of a single part may render the system inactive. Thus, it is reasonable to speculate that bacteria also employ minimal, single-component systems that degrade phage nucleic acids; however, the prevalence and effectiveness of such systems remain poorly understood.

We undertook this study with an aim to uncover new mechanisms of immunity in the commensal opportunistic pathogen *Staphylococcus epidermidis* RP62a (Christensen et al., 1987), and our efforts fortuitously led to the discovery of a unique mode of immunity mediated by a single enzyme, *SERP2475*, herein referred to as Nhi (nuclease-helicase immunity). Nhi provides full (10^6 - to 10^8 -fold) protection against diverse staphylococcal phages, and it is sufficient to prevent phage DNA accumulation in a strain devoid of all other known defenses. Biochemical characterization of Nhi combined with genetic analyses of Nhi-resistant phage hybrids and “escapers” from diverse families support a model in which Nhi targets and de-

grades phage-specific replication intermediates. Importantly, Nhi homologs can be found in diverse bacterial phyla, and we provide evidence that some are also involved in immunity. Altogether, our findings highlight the versatility of such compact systems as powerful weapons in antiphage defense.

RESULTS

SERP2475 protects against diverse staphylococcal phages

S. epidermidis RP62a harbors a Type III-A CRISPR-Cas system (Marraffini and Sontheimer, 2008), an Abi mechanism (Depardieu et al., 2016), and a putative Type I RM system, all of which are encoded within $\sim 30,000$ nt of each other (Figure 1A). Their close proximity is consistent with recent reports that show prokaryotic immune systems typically cluster together within discrete genomic loci known as defense islands (Doron et al., 2018; Gao et al., 2020; Makarova et al., 2011). Importantly, key insights into CRISPR-Cas and Abi in this organism were revealed by studying their molecular

interactions with temperate phages Φ NM1 and CNP_x, respectively, related phages that belong to the family *Siphoviridae* (Depardieu et al., 2016; Goldberg et al., 2014). Indeed, siphophages are the most common members in staphylococcal phage collections (Oliveira et al., 2019), and we reasoned that the identification of new immunity mechanisms might necessarily require the examination of more diverse members. Toward that end, we isolated and characterized four lytic *S. epidermidis* phages belonging to the family *Podoviridae*: Andhra, JBug18, Pontiff, and Pike (Cater et al., 2017; Culbertson et al., 2019). These phages share over 95% sequence identity and the same 20 genes (Figure S1); however, we noticed that they have distinct host ranges—although Andhra and Pontiff can infect wild-type (WT) RP62a, JBug18 and Pike can only infect a mutant variant, LM1680 (Jiang et al., 2013), which has a large (~300-k nt) deletion encompassing the defense island (Figure 1A; Culbertson et al., 2019). These observations led to the hypothesis that JBug18 and Pike are sensitive to genetic element(s) within the defense island.

To test this, we used a set of *S. epidermidis* RP62a mutants that were originally identified in plasmid transfer experiments as being defective in CRISPR immunity and later found to bear deletions of varying extents across the defense island and beyond (Jiang et al., 2013; Figure 1A). These strains were challenged with Andhra and JBug18, representatives with resistant and sensitive phenotypes, respectively. The resulting zones of bacterial growth inhibition (plaques) were enumerated, revealing that only one of the mutants (R7) encodes the gene(s) required for full protection against JBug18. These observations narrowed the protective genetic element(s) to a stretch of ~12,000 nt containing 12 genes (designated as SERP2466–SERP2477) that incidentally encompasses the RM system (Figure 1B). To determine which gene(s) are responsible for immunity, they were inserted into a derivative of plasmid pC194 (Ehrlich, 1977) (herein referred to as pSERP-), introduced into LM1680, and resulting strains were challenged with Andhra and JBug18. Through this analysis, we found that a single gene of unknown function, *SERP2475* (new locus tag *SERP_RS12125*), is sufficient to protect against JBug18 (Figure 1B). A repeat of this assay with phages Pontiff and Pike showed results similar to those observed with Andhra and JBug18, respectively—although *SERP2475* has little/no effect on Pontiff, it completely protects against Pike (>10⁸-fold, Figure 1C).

To further understand the breadth of protection this gene affords, we first challenged LM1680/pSERP-2475 with additional phages from our collection from different morphological families—*Herelleviridae* (Barylski et al., 2020) (formerly *Myoviridae*) and *Siphoviridae*. Importantly, staphylococcal phages from different morphological families are genetically very distinct and share little/no sequence homology (Oliveira et al., 2019). We observed that although *SERP2475* has no noticeable impact on the lytic myophages Twillingate and Quidividi (Freeman et al., 2019), it fully protects against siphophage CNP_x (>10⁶-fold, Figure 1C). We also tested the effectiveness of *SERP2475* against *S. aureus* phages representing all three families of tailed phages and found that it affords full (>10⁷-fold) protection against myophage ISP (Vandersteegen et al., 2011) and siphophage Lorac (Marc et al., 2019). Taken together, these data demonstrate that *SERP2475* is sufficient to provide robust protection against diverse staphylococcal phages from all three morphological families.

SERP2475 homologs are distributed in diverse bacteria and exhibit antiphage activity

We next assessed the distribution of *SERP2475* homologs and the extent to which they are functionally conserved. We used tBLASTn to query NCBI databases with fully assembled microbial genomes and identified 302 homologs in distinct genetic backgrounds (Table S1). Although homologs were present in <1% of the genomes surveyed, representatives could be found in three bacterial phyla: *Firmicutes*, *Bacteroidetes*, and *Proteobacteria*. Additionally, two of the homologs were found in two free *Streptococcus* phages. To better understand their relationships to one another, we generated a phylogenetic tree from 100 selected representatives that encompass the phylogenetic diversity of the group (Figures 2A and S2; Data S1). This analysis revealed that the homologs cluster into three distinct clades that are somewhat incongruent with host phylogeny. Although clade III contains homologs strictly found in *Proteobacteria*, clades I and II contain homologs originating from *Firmicutes* and one additional phylum. This observation suggests that *SERP2475* and its homologs are likely disseminated through horizontal gene transfer. Supporting this, many of the homologs are encoded on plasmids (Figure 2A; Table S1). Further, of the five homologs represented by WP_115261955 in the tree (which share >95% amino acid identity), four originate from related *Streptococcus* phages or prophages, whereas one resides in the genome of *S. dysgalactiae* (CP033163.1) in an entirely different genetic context. To gain insight into their level of functional conservation, we first checked for proximity to genes with known defense functions. Amino acid sequences of the thirty proteins encoded upstream and downstream of each homolog were searched for identifiable protein domains using hmmer v.3.3.2 (hmmer.org). The predicted protein families (pfams) of these flanking proteins were then searched against 306 pfams with known functions in antiphage defense (Table S2). Figure 2B shows the fraction of homologs that have at least one and up to ten defense-related neighbors encoded within expanding windows of 10, 20, and 30 genes. These data revealed that a significant fraction of *SERP2475* homologs are indeed encoded proximal to known defenses, supporting potential roles for these homologs in immunity. For instance, 72% of homologs have at least one defense neighbor encoded within 20 genes (Figure 2B; Table S3). The most frequently encountered defense neighbors within 20 genes include proteins with TOPRIM (topoisomerase-primase, PF01751), ATPase (PF13304, PF00176, and PF00004), and methyltransferase (PF02384 and PF01420) domains (Figure 2C; Table S4).

To further explore their functional conservation, four representative homologs from the three clades were tested for antiphage activity (Figures 2A and S2; Table S1). Importantly, these homologs share minimal (32%–36%) amino acid sequence similarity when compared with *SERP2475*. The coding sequences for *SERP2475* and selected homologs were inserted into a plasmid (herein referred to as pTET-) downstream of an anhydrotetracycline (aTc) inducible promoter and then introduced into *S. aureus* RN4220. The resulting strains were challenged with phage Lorac in the presence and absence of inducer. As expected, *SERP2475* (WP_002489608) completely protects against Lorac when the cells were grown in the presence of the maximum concentration of aTc (Figure 2D). Although expression of the clade II homolog (WP_013870910

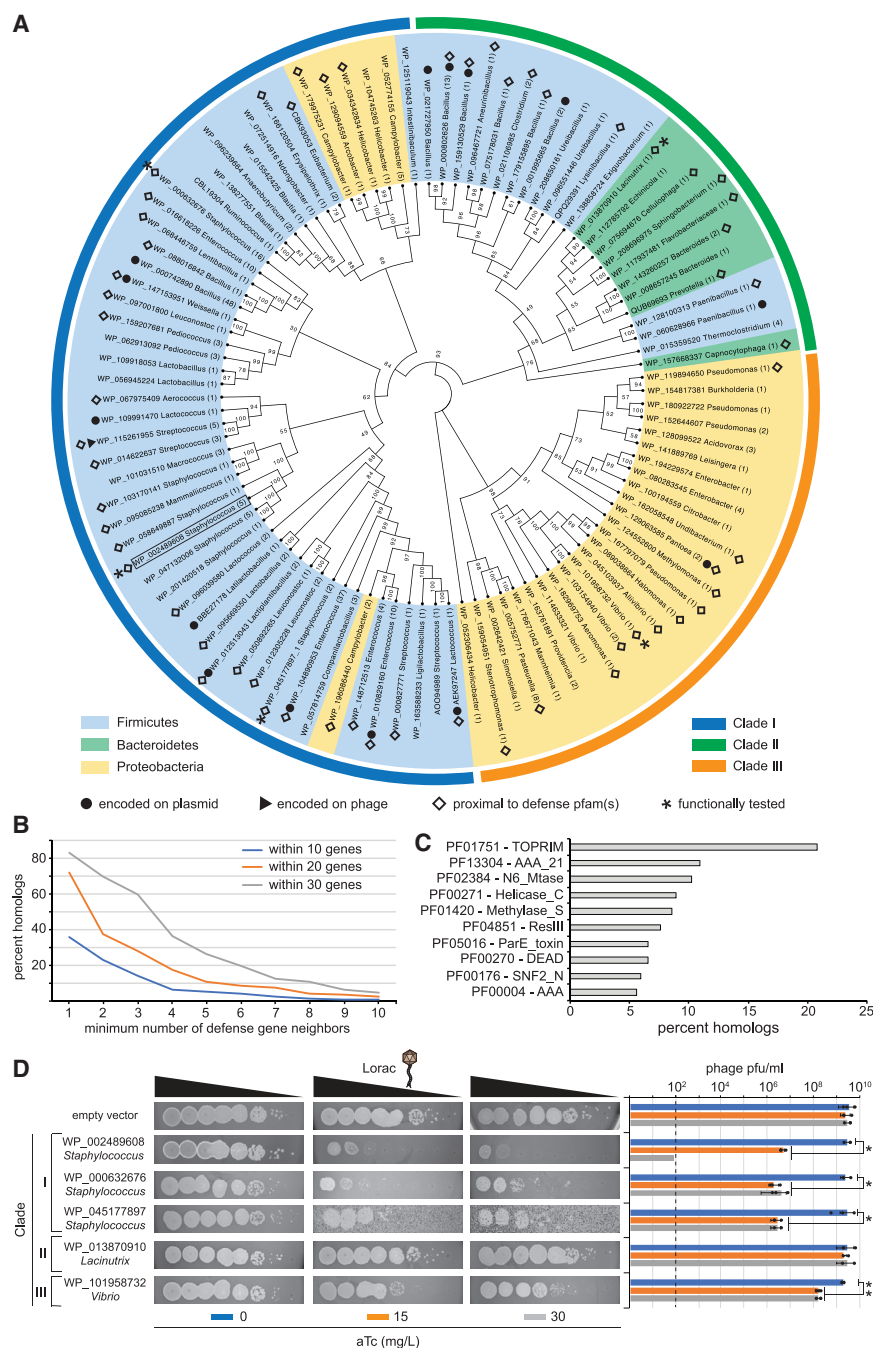


Figure 2. SERP2475 homologs exhibit functional conservation

(A) A cladogram showing the phylogenetic distribution of SERP2475 and selected homologs. Branch labels indicate bootstrap values and tip labels show the NCBI RefSeq ID number, genus from which the homolog originated, and number of distinct species/strains within the same genus that harbor a closely related homolog (>90% amino acid identity). Inner ring highlighting indicates the phylum, and the outer ring coloration indicates the clade (I–III). Representatives encoded on plasmids (closed circles), phages (closed triangle), and within 20 genes of at least one neighbor with a predicted defense function (open diamonds) are indicated. Asterisks mark homologs that are functionally characterized in this study, and SERP2475 is additionally enclosed with a box. (B) Neighborhood analysis showing the fraction of homologs (including SERP2475, $n = 303$) that have at least x number (1–10) gene neighbors involved in antiphage defense within 10, 20, and 30 flanking genes. The x axis values are cumulative. (C) Top ten defense-related protein families (pfams) encoded within 20 genes of SERP2475 homologs. (D) Coding sequences of indicated homologs were inserted into a plasmid downstream of an anhydrotetracycline (aTc) inducible promoter and subsequently introduced into *S. aureus* RN4220. Images show 10-fold dilutions of Lorac (10^0 – 10^{-7}) spotted atop lawns of indicated strains grown in the absence or presence of aTc (15 and 30 mg/L). An average of triplicate measurements of pfu/mL ($\pm$ SD) is shown as a representative of three independent trials. Asterisks indicate p values <0.05 (*) and <0.0005 (**) in a two-tailed t test. Dotted line indicates the limit of detection, and short bars underneath indicate 0 pfu/mL. See also Figure S2.

from *Bacteroidetes*) resulted in no detectable decrease in plaque size or number, the two *Staphylococcus* homologs from clade I (WP_000632676 and WP_045177897) and the *Vibrio* homolog from clade III (WP_101958732 from *Proteobacteria*) caused significant reduction in plaque numbers (1,000-fold and 10-fold, respectively). A noticeable decrease in plaque size was also observed in the latter. Interestingly, one of the staphylococcal homologs (WP_045177897) also appeared to be toxic to the cells, as evidenced by the “unhealthy” appearance of the lawn when cells harboring the homolog are grown in the presence of inducer (Figure 2D). Altogether, these data

provide evidence that distant homologs of SERP2475 are also involved in antiphage defense.

SERP2475 limits phage DNA accumulation

We next sought to investigate the mechanism of immunity and began by assessing which stage of the phage infection cycle SERP2475 targets. As previously mentioned, common strategies that bacteria employ in antiphage defense include masking/modifying the cell surface to prevent phage attachment, targeting and degrading phage-derived nucleic acids, and causing programmed cell death through various Abi mechanisms

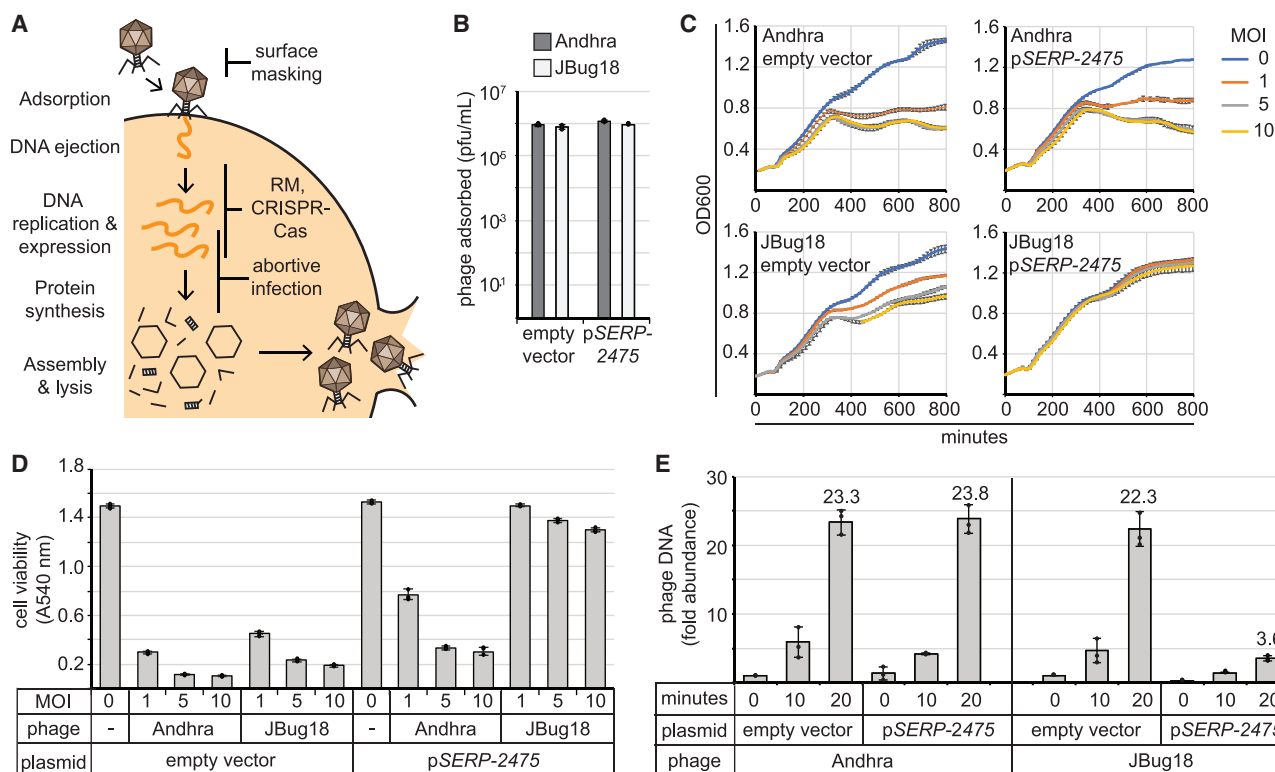


Figure 3. SERP2475 impairs phage DNA accumulation

(A) Illustration of the lytic phage replication cycle and common defenses that interfere with each step. (B–D) Results of an adsorption assay (B), cell growth assay (C), and cell viability assay (D) following challenge of *S. epidermidis* LM1680 cells bearing indicated plasmids with Andhra and JBug18. Multiplicity of infection (MOI), ratio of bacteria:phage; OD600, optical density at 600 nm; and A540, absorbance at 540 nm. (E) Relative abundance of phage DNA at various time points following phage infection as measured by qPCR. For all experiments, the mean $\pm$ SD of triplicate measurements are shown as a representative of at least two independent trials.

(Figure 3A). To test if *SERP2475* hinders phage attachment, we conducted an adsorption assay in which LM1680/p*SERP-2475* was combined with a defined number of phage particles for 10 min—just enough time for phages to attach to cells but not long enough for these phages to complete their replication cycle (Cater et al., 2017). Phages remaining in suspension were then enumerated to determine the number that adsorbed to cells. This assay revealed that JBug18 attaches to LM1680/p*SERP-2475* just as efficiently as it attaches to LM1680 cells bearing the empty vector (Figure 3B), thus ruling out an adsorption-blocking mechanism.

Cell growth and viability assays were next performed to test for Abi. The prediction is that if programmed cell death accompanies immunity, then challenge with a high proportion of phages to cells ($\geq 1:1$) would lead to significant decline in cell growth and viability similar to that observed in the absence of immune protection (Goldfarb et al., 2015; Ofir et al., 2018). To test this, phages were combined with LM1680/p*SERP-2475* in liquid media at ratios of 1:1, 5:1, or 10:1. Cell growth was tracked by taking optical density measurements at 600 nm (OD600) every 15 min over 800 min (~ 13 h). As expected, both phages caused a significant decline in the OD600 of LM1680 cells harboring the empty vector after about 5 h of growth (Figure 3C). However, although LM1680/p*SERP-2475* remained sensitive to Andhra, the strain grew normally in the presence of JBug18 at all phage:bacteria

ratios tested. We also quantified cell viability with an assay that uses the enzymatic reduction of the colorless 2,3,5-triphenyl tetrazolium chloride (TTC) reagent by living cells and concomitant generation of a red product as a proxy for viable cell count (Tengerdy et al., 1967). We found that although Andhra causes significant death of LM1680/p*SERP-2475*, JBug18 elicits only a minor decrease in the viability of this strain, even when phages outnumber bacteria 10:1 (Figure 3D). Since the cytotoxic effects of Abi should become apparent at a phage:bacteria ratio of $\sim 1:1$, these observations suggest that cell suicide is unlikely to be the mechanism by which *SERP2475* affords protection.

Finally, we tested whether *SERP2475* impacts phage DNA levels in the cell. Quantitative PCR (qPCR) was used to track the accumulation of phage DNA at various time points following infection with both phages. The results showed that although Andhra's DNA accumulates to ~ 20 -fold by 20 min postadsorption in LM1680/p*SERP-2475*, JBug18's DNA accumulates to less than 4-fold in the same time period (Figure 3E). These observations support a hypothesis whereby *SERP2475* protects against phages by interfering with phage DNA replication.

SERP2475 relies upon nuclease and helicase activities to perform immunity

To begin to understand the catalytic function of this protein, we first conducted *in silico* analyses. *SERP2475* encodes a

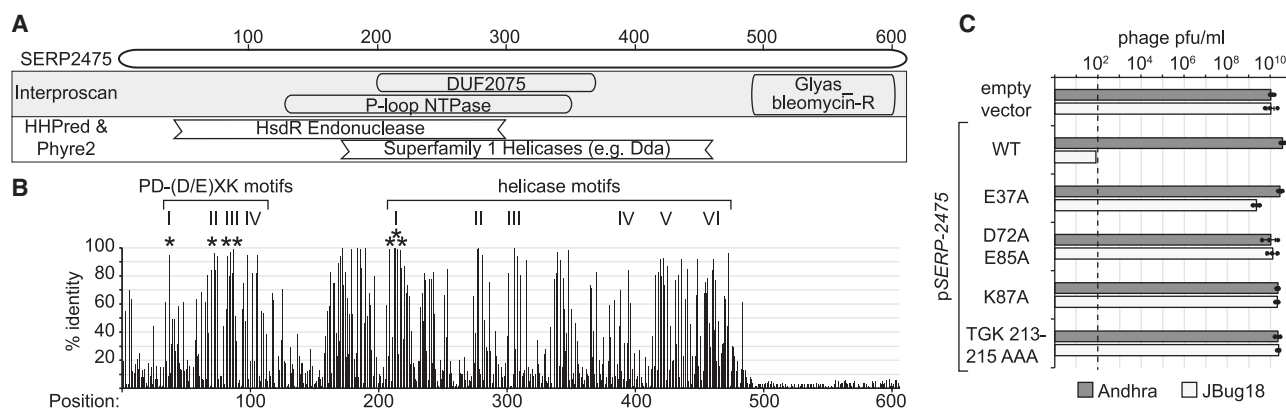


Figure 4. Conserved nuclease and helicase domains are required for immunity *in vivo*

(A) Predicted domains and structural homologs of SERP2475.

(B) The fraction of 99 homologs that possess amino acids identical to those of SERP2475 at each position in their multiple sequence alignment. Putative nuclease (PD-(D/E)XK) and helicase motifs are labeled, and asterisks mark positions that were subjected to mutational analysis in this study.

(C) *S. epidermidis* LM1680 strains bearing plasmids with the WT and mutant versions of SERP2475 were challenged with Andhra and JBug18, and resulting pfu/mL were enumerated. Shown are an average of triplicate measurements ($\pm$ SD) as a representative of at least three independent trials. Dotted line indicates the limit of detection, and short bars underneath indicate 0 pfu/mL.

606-amino acid protein, and according to Interproscan (Blum et al., 2021), a web-based tool that predicts protein domains and other important functional sites, the most prominent features of SERP2475 are a central domain of unknown function (DUF2075, protein family PF09848) and an overlapping P-loop NTPase domain (Figure 4A). Interproscan also identified a putative domain involved in bleomycin resistance on the C terminus (discussed later). Additionally, the structural homology search tools HHPred (Zimmermann et al., 2018) and Phyre2 (Kelley et al., 2015) identified a handful of superfamily 1 helicases as close homologs, including T4 phage Dda helicase (He et al., 2012) (E value 1.2×10^{-21}). These tools also predicted structural similarities between the N terminus of SERP2475 and a putative HsdR restriction endonuclease from *Vibrio vulnificus* (Uyen et al., 2009) (E value 1.1×10^{-6}). Although SERP2475 shares minimal amino acid sequence similarity in pairwise comparisons with full-length HsdR_Vv and Dda (19% and 22%, respectively), we located conserved residues in key motifs corresponding to a putative PD-(D/E)XK nuclease domain on its N terminus (Figure S3A) and putative helicase domain spanning the central portion of the protein (Figure S3B). Further, we observed a striking conservation of these critical residues in the sequence alignment with SERP2475 and the 99 representative homologs (Figure 4B), suggesting that these domains are likely important for protein function. To confirm, we constructed mutant versions of pSERP-2475 that encode alanine substitutions in the putative nuclease and helicase domains (asterisks in Figures 4B and S3). The strains were then challenged with Andhra and JBug18, and the results showed a restoration of JBug18 replication in all mutant strains (Figure 4C). These observations support the hypothesis that SERP2475 and its homologs use nuclease and helicase activities to achieve immunity.

To test for these enzymatic functions directly, we took a biochemical approach. SERP2475 was introduced in the pET28b expression vector downstream of an N-terminal tag, overexpressed in *E. coli*, and subjected to a three-step purification process (Figure S4A). During the first two steps, we

noticed that SERP2475 copurifies with a smaller species, and following digestion of the tag, both exhibit a reduction in size of ~ 14 kDa (Figure S4B). This suggested that the smaller species is also tagged and thus comprises an N-terminal fragment of the full-length (FL) protein. To confirm, we excised the bands in the SDS-PAGE gel corresponding to the FL (~ 72 kDa) and putative truncated variant (~ 23 kDa) and subjected the proteins to mass spectrometry analysis. The results indicated that SERP2475 is indeed the most abundant protein in both bands (Tables S5 and S6), with the truncated version showing dense peptide coverage over only the first ~ 230 amino acids (AAs) (Figure S4C). The subsequent size exclusion chromatography step successfully separated the FL and truncated versions and also revealed that the FL species forms a dimer in solution, as evidenced by the presence of two adjacent peaks in the chromatogram that contain the pure FL protein (Figure S4D).

We next conducted *in vitro* functional analyses. Nuclease assays were first performed with the fractionated protein and revealed both 3'-5' exonuclease and plasmid nicking activities (Figures S4E and S4F) stemming from fractions containing the N-terminal fragment. Interestingly, these activities were nearly absent in fractions containing the FL protein, suggesting a possible mechanism of autoinhibition. To confirm that the N-terminal domain is sufficient to produce both activities, we introduced the region encoding the first 230 AAs of SERP2475 into pET28b and purified this truncated version using the same purification protocol (Figure S4A). We observed that the N-terminal fragment elutes as a monomer (Figure S4G), indicating that the dimerization domain is likely located elsewhere on the protein. Importantly, nuclease assays confirmed that the N-terminal domain is sufficient to produce robust exonuclease and nickase activities (Figures 5A, 5B, S4H, and S4I). Further, the introduction of point mutations in conserved residues, specifically E37A and K87A, cause loss of exonuclease activity (Figure 5A) and significant reduction in nickase activity (Figure 5B) *in vitro*. The WT enzyme is unable to further degrade the nicked or linearized

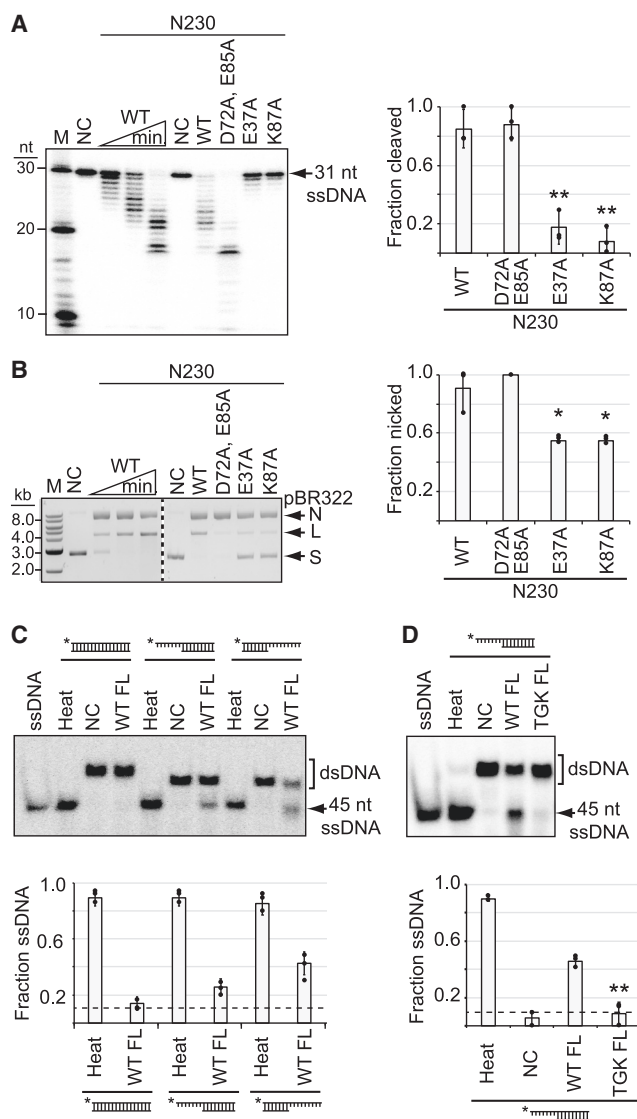


Figure 5. SERP2475 has nuclease and helicase activities

(A) Exonuclease assays are shown in which a radiolabeled single-stranded (ssDNA) substrate was combined with the N-terminal nuclease domain (N230) of WT SERP2475 for 5, 10, and 20 min or indicated mutants for 20 min.

(B) Nickase assays are shown in which plasmid pBR322 was combined with the N-terminal domain of WT SERP2475 for 15, 30, and 60 min or indicated mutants for 30 min. Representative gel images (left) and average of three independent trials ($\pm$ SD, right) are shown. For nickase assays, nicked (N), linear (L), and supercoiled (S) forms of the plasmid are labeled.

(C and D) Helicase assays are shown in which radiolabeled double-stranded DNA (dsDNA) substrates with blunt ends or overhangs were incubated for 1 h with full-length (FL) SERP2475 (WT or TGK 213–215 AAA triple mutant) and resolved using native PAGE. As a positive control, substrates were heated to 95°C for 2 min. Representative gel images (top) and averages of three independent trials ($\pm$ SD, bottom) are shown. Asterisks indicate p values <0.05 (*) and <0.005 (**) in a two-tailed t test.

M, molecular weight ladder; kb, kilobase; nt, nucleotide; NC, negative control (no protein). Table S7 lists all substrates used for these assays. See also Figures S4 and S5.

double-stranded DNA (dsDNA) products when given 1–4 h of incubation (Figure 5B and not shown), suggesting that supercoiling in plasmids is essential for cleaving dsDNA. Both activities require either Mg^{2+} (preferred) or Mn^{2+} and can be observed using different DNA substrates that are devoid of phage-derived sequences (Figure S5; Table S7), providing evidence that these activities are likely sequence nonspecific.

We also tested for helicase activity. As mentioned earlier, the central portion of SERP2475 has conserved helicase motifs, and many of its predicted structural homologs are superfamily 1 helicases (Figure 4). These enzymes function as a monomer or dimer to unwind double-stranded substrates using energy from ATP (Fairman-Williams et al., 2010). To test for this activity, the FL fractions containing the monomer and dimer peaks (Figure S4D) were pooled, concentrated, and incubated with various DNA substrates (Table S7). The results showed that SERP2475 can indeed unwind dsDNA when offered a 5'- or 3'-single-stranded overhang (Figure 5C). Further, point mutations in conserved residues comprising motif I of the helicase domain, which constitutes the ATP binding site (T213A, G214A, and K215A), caused complete loss of helicase activity (Figure 5D). Altogether, these results demonstrate that SERP2475 possesses exonuclease and nickase activities stemming from its N terminus, as well as bidirectional helicase activity from a central helicase domain. On the basis of these activities, we renamed SERP2475 to Nhi and refer to it as such here onward.

Phage-encoded single-stranded DNA-binding proteins mediate Nhi immunity

To further refine our understanding of Nhi's targeting specificity, we sought to determine how diverse phages evolve to resist immunity. Andhra and JBug18 encode the same 20 proteins (Figure S6A), and a pairwise alignment of their coding regions show that they differ at only 705 positions by either a single-nucleotide polymorphism (SNP) or a gap (Data S2). To narrow down which SNPs and/or gaps in Andhra are important for resistance to immunity, we first attempted to isolate naturally evolved JBug18 mutants that can escape immunity by plating concentrated phage preparations with LM1680/pSERP-2475. After several failed attempts at recovering plaques, one attempt yielded resistant phages, which upon further inspection were found to possess hybrid genomes that contain a patchwork of Andhra and JBug18 sequences. These hybrids necessarily arose through the inadvertent mixing of the two phages and propagation on the same host strain. Nonetheless, this fortuitous accident proved invaluable in helping to pinpoint the region required for immune resistance—since all hybrids can escape immunity, we reasoned that they must share Andhra-derived sequences in the region(s) required for resistance. To test this, we purified and sequenced eight such hybrids and determined the fraction that possess Andhra identity at each of the 705 differing positions across their coding regions. We found that all hybrids share Andhra identity at positions 891–2,117 in the alignment (Figure 6A; Data S2). This region overlaps gene products (*gp03-06* in the phage genomes and encompasses 69 SNPs and gaps, of which 64 occur within *gp03* and *gp04* (Figure 6B). Accordingly, we speculated that one or both of the latter genes are responsible for resistance.

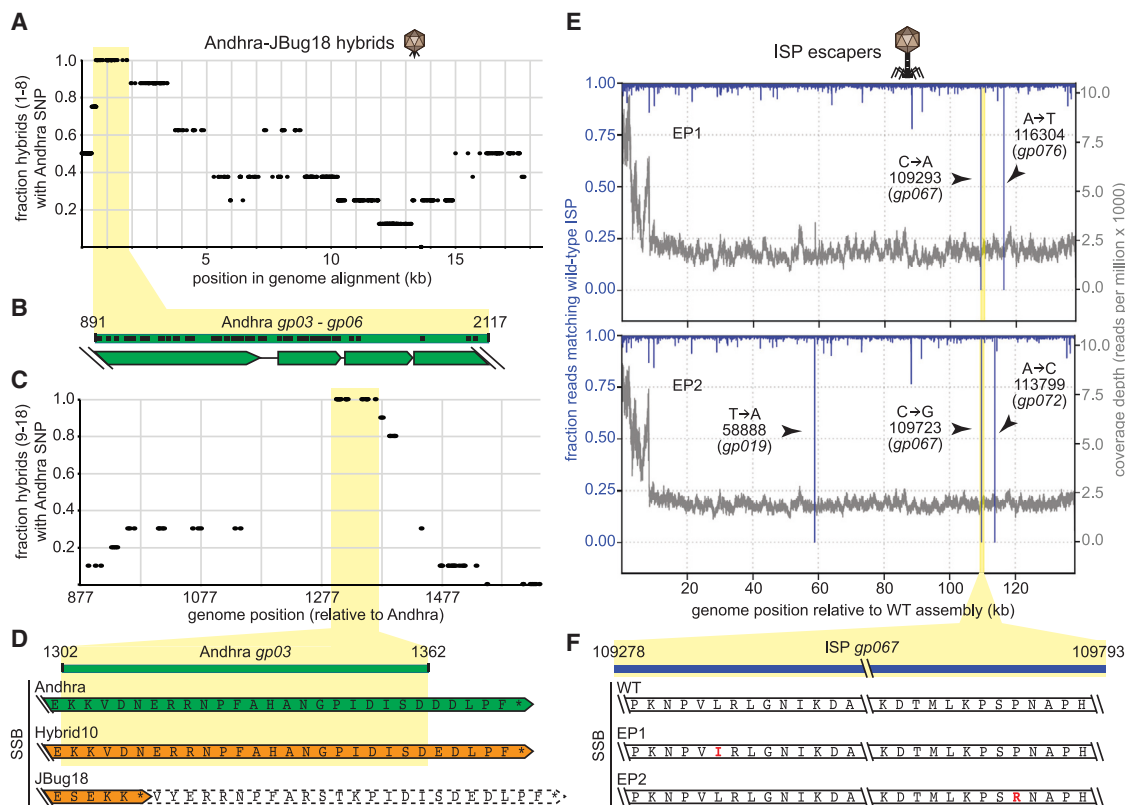


Figure 6. Phage-encoded single-stranded DNA-binding proteins mediate Nhi immunity

(A) Fractions of JBug18-Andhra hybrids 1-8 that harbor Andhra identity at positions where the parental phage genomes differ. The segment common to all hybrids is highlighted yellow.

(B) Expanded view of highlighted region in (A). Black dots show relative positions where SNPs and/or gaps occur in an alignment between parental phages.

(C) Fractions of JBug18-Andhra hybrids 9-18 that harbor Andhra identity in the *gp03/04* region. See also Figure S6.

(D) Expanded view of highlighted region in (C). Green and orange arrows delimit phage-encoded SSBs, and white arrow with dotted border delimits a truncated portion. See also Data S2 and S3.

(E) Illumina sequencing reads from two ISP escaper phages aligned with a *de novo* assembly of the ancestral WT phage. The fraction of nucleotides matching the WT sequence at each position (blue, left axis) and coverage depth (gray, right axis) are shown. Positions of SNPs are marked with arrows, and their identities are indicated. The yellow strip highlights the position of the SSB (*gp067*).

(F) Expanded view of ISP's SSB showing the amino acid changes caused by the SNPs.

To narrow down the protective region further, a second set of resistant Andhra-JBug18 hybrids were generated that bear Andhra-derived sequences in *gp03* and *gp04*. This was accomplished by introducing Andhra's *gp03* and/or *gp04* coding regions into *S. epidermidis* LM1680 on plasmids and then propagating JBug18 on these strains to allow for recombination with the Andhra-derived sequences (Figure S6B). The resulting phages were then plated on LM1680/pSERP-2475 to select for resistant phage recombinants. Ten such hybrids (9-18) were purified, and sequencing across *gp03* and *gp04* revealed that they had all acquired a 60-nt stretch spanning positions 1,302-1,362 in Andhra's genome (Figure 6C; Data S2). This region overlaps *gp03*, which encodes a single-stranded DNA-binding protein (SSB, Figure 6D; Data S3). Importantly, JBug18 has a 5-nt insertion in this region and consequently harbors a truncated variant of the SSB (Figures 6D and S6C). However, by acquiring the 60-nt stretch from Andhra, all ten hybrids had restored the reading frame and hence encode a FL SSB, suggesting that the C terminus of the SSB in *Podoviridae* phages plays a critical

role in escape from immunity. In agreement with these observations, phages Pontiff and Pike possess the expected *gp03* genotypes: While the resistant Pontiff encodes a FL SSB, the sensitive Pike encodes a truncated version, this time due to a single-nucleotide deletion (Figure S6D).

Finally, we attempted to isolate "escaper" phages from members of the remaining two families that exhibit sensitivity to Nhi-*Siphoviridae* phage CNP_x and *Herelleviridae* phage ISP (Figure 1D). After several trials of plating concentrated lysates of these phages on LM1680/pSERP-2475, we were unable to recover resistant CNP_x mutants; however, this approach yielded Nhi-resistant ISP variants. Two of these escapers (EP1 and EP2) were purified, and their DNA was extracted and subjected to Illumina sequencing along with the ancestral WT phage. The WT ISP genome was then assembled *de novo*, and the reads from escapers were aligned to the WT assembly to identify positions that differed between them. We found that the escapers harbor just two (EP1) and three (EP2) SNPs at different positions, as evidenced by an abrupt

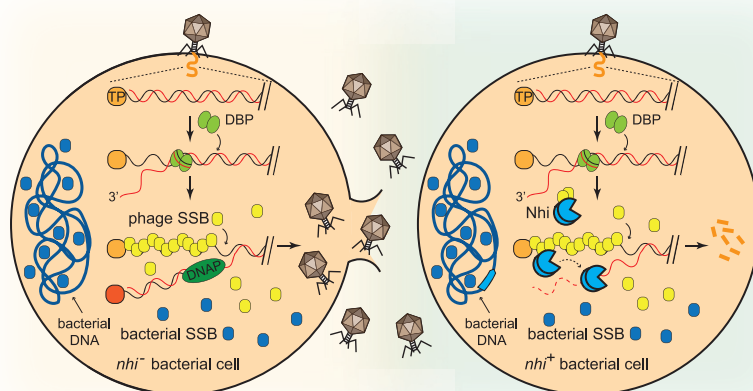


Figure 7. Proposed mechanism for Nhi defense against *Podoviridae* phages

Each 5' end of phage DNA is covalently linked to a terminal protein (TP). During replication initiation, phage-encoded double-stranded DNA-binding proteins (DBPs) unwind DNA ends. In the absence of Nhi (left), the phage DNA polymerase (DNAP) in complex with a second TP loads onto the free 3' end (red strand), adds the first base to a 3'-OH group on the TP, and continues to build a complementary DNA strand (black). Phage-encoded single-stranded DNA-binding proteins (SSBs, yellow) protect free single-stranded DNA and coordinate the replication process. In the presence of Nhi (cyan, right), specific structural elements in phage-encoded SSBs recruit Nhi to the phage DNA. Bacterial SSBs (dark blue) lack the Nhi recognition site, thus allowing the bacterial genome to remain unharmed.

drop to zero in the fraction of reads that match the WT sequence (Figure 6E). Remarkably, of the genes bearing mutations in these escapers, one is common to both-*gp067*, a predicted replicative SSB (HHPred, E value 3.9×10^{-20} , 99.85% probability). Altogether, these observations suggest that phage-encoded SSBs play a central role in determining the final outcome of Nhi immunity.

DISCUSSION

Here, we describe a unique mode of nucleic acid immunity mediated by Nhi. One attribute that sets Nhi apart from other innate DNA-targeting immune systems (such as RM) is that it does not appear to rely upon recognition of specific DNA sequences (Figures 5 and S5) and yet abrogates phage DNA accumulation without causing appreciable cell death (Figure 3). These observations beg the question: What is the basis for phage specificity? Our finding that resistance to Nhi in diverse phages correlates with variations in their respective SSBs (Figure 6) supports a preliminary model in which Nhi is recruited to phage-specific replication intermediates by phage-derived SSBs (Figure 7). Importantly, Nhi affords protection against members from all three staphylococcal phage families (Figure 1), and although these phages are genetically distinct, they all rely upon their own SSBs to coordinate genome replication.

Staphylococcal *Podoviridae* phages use a protein-priming mechanism of replication (Salas et al., 2016; Vybiral et al., 2003), in which terminal proteins (TPs) covalently linked to the 5' ends of their linear dsDNA provide hydroxyl groups upon which to initiate replication (Figure 7). The initial stages of replication involve local unwinding of DNA ends by double-stranded DBPs and subsequent release of the free 3' end. Under normal circumstances, the 3' end is captured by the phage DNA polymerase (DNAP) in complex with a second TP, whereupon replication ensues. We speculate that Nhi is recruited to the phage DNA by binding unique features of the phage-encoded SSB. From there, it may compete with DNAP for the free 3' end and use its 3' exonuclease and helicase activities to processively degrade the phage genome. SSBs are known to bind and protect DNA and coordinate the replication machinery, particularly through interactions with their C terminus

(Shereda et al., 2008). The precise mechanism by which the C terminus of the *Podoviridae* phages' SSB protects against Nhi's effects remains to be determined. As one possibility, the C terminus may obscure Nhi's recognition site, either on its own or through recruitment of accessory factor(s). We were also able to isolate Nhi-resistant variants of the Twort-like phage ISP (Figure 6), and although their replication mechanism remains poorly understood (Klumpp et al., 2010), the DNAP, SSB, and other components of the replication machinery can be readily identified through *in silico* analyses. In light of the fact that staphylococcal *Podoviridae* and *Herelleviridae* phages are completely devoid of recognizable sequence homology (Oliveira et al., 2019), it is striking that Nhi-resistant ISP variants also harbor mutations in their SSB (Figure 6), which presumably results in the loss of Nhi's recognition site.

Remarkably, siphophages Lorac and CNP_x, which exhibit a distinct mechanism of replication, also succumb to Nhi (Figure 1C). Such lambda-like phages undergo several rounds of theta replication, followed by rolling-circle replication (Casjens and Hendrix, 2015; Narajczyk et al., 2007). Once recruited, Nhi has the capacity to nick supercoiled theta replication intermediates. Since the FL Nhi forms a dimer in solution (Figure 5D), it may introduce two adjacent nicks on opposing strands and cause double-stranded breaks in the phage genome. Indeed, even the monomeric 230 AA fragment caused some plasmid linearization (Figure 5B), presumably through simultaneous nicking on opposite strands that colocalized by chance. Since FL Nhi is devoid of nuclease activity *in vitro* (Figure S4), we speculate the phage-encoded SSBs may also play a role in Nhi activation *in vivo*. Supporting this notion, the eukaryotic SSB known as Rpa binds and modulates the enzymatic activities of the nuclease-helicase DNA2 (Zhou et al., 2015), which was identified as a close structural homolog of Nhi according to our *in silico* analyses. SSBs from the three families of staphylococcal phages share little/no sequence identity at the amino acid level; however they possess common structural elements (such as the OB fold and linker region) that coordinate SSB interactions with multiple binding partners and facilitate their different functions (Bianco, 2021). In order to prevent autoimmunity, Nhi likely recognizes structural features of phage-derived SSBs that are absent in the SSB of the bacterial host. Future work will explore

the physical and functional interactions between Nhi and phage-encoded SSBs.

It is unclear how the FL version of Nhi is truncated in *E. coli* prior to purification (Figures S4B and S4C), and whether a truncated form plays role(s) in the native host remains unknown. Nonetheless, its unexpected appearance allowed us to reconstitute the nuclease activities stemming from Nhi's N-terminal domain. Nhi homologs share conserved residues on their N terminus corresponding to a PD-(D/E)XK domain (Figure 4) in which the acidic residues (D and E) coordinate metal ion(s) and the basic residue (K) stabilizes the transition state. These domains are found in a broad superfamily of nucleases involved in diverse functions, including DNA restriction, recombination, and repair (Steczkiewicz et al., 2012). Members of this group include restriction enzymes, holliday junction resolvases, herpesvirus exonucleases, and many others from all kingdoms of life. Nhi's helicase motifs comprise a separate domain that overlap with the DUF2075 (Figure 4). In addition to the phage T4 Dda helicase, close predicted structural homologs include the human helicase Upf1 (E value 2.5×10^{-23}) and mouse nuclease-helicase DNA2 (E value 1.2×10^{-21}), which are involved in nonsense-mediated mRNA decay and DNA replication/repair, respectively (Chakrabarti et al., 2011; Zhou et al., 2015). All three are superfamily 1B helicases, which unwind double-stranded substrates in the 5'-3' direction. Interestingly, Nhi's helicase activity is bipolar (Figure 5C), which undoubtedly allows it to act on more diverse substrates. Importantly, mutations in the predicted nuclease and helicase motifs of Nhi eliminate antiphage immunity *in vivo* (Figure 4), and most of the mutant variants showed reduction/elimination of these activities in a purified system (Figure 5). The one exception is the D72A, E85A double mutant that retains both nuclease activities *in vitro*. This may be explained by the presence of additional acidic residues that play a redundant role in metal ion coordination. Supporting this possibility, a recent study showed that the exonuclease and nickase activities of OLD (overcoming lysogenization defect) family nucleases harbor two metal binding sites that comprise 3 AAs each, and removal of all three from one site is required to eliminate exonuclease activity *in vitro* (Schiltz et al., 2019). Interestingly, nicking activity in OLD family nucleases is extremely robust and persists in the presence of multiple active site mutations. The latter is consistent with our observations of Nhi's behavior and highlights the need for structural analyses to glean more detailed mechanistic insights. We also noted the presence of a C-terminal domain (Glyas_Bleomycin-R) that is poorly conserved across Nhi homologs (Figures 4A and 4B). Such domains are found in a group of metalloenzymes that perform a variety of activities, including isomerizations and epimerizations (Armstrong, 2000). Whether and how this domain coordinates with the others to achieve immunity are subjects of ongoing work.

In contrast to the *S. epidermidis* Nhi, other homologs tested showed more modest antiphage activity (Figure 2). This could be explained by incompatibility with the foreign phage target or heterologous host background, which together might lessen their apparent effectiveness. Supporting this, the homologs that were most effective against *S. aureus* RN4220 phage Lorac originated from two different *S. aureus* strains. Nonetheless, the observation that the *V. vulnificus* homolog could still afford protection is remarkable and underscores the versatility of this mode of pro-

karyotic immunity. Interestingly, one of the *S. aureus* homologs abrogated cell growth while also blocking phage replication (Figure 2), suggesting that it may cleave both phage and host DNA, and as a consequence, slow cell division and/or cause cell death.

Finally, it bears mentioning that the Nhi homologs identified in this study represent but a small subset of DUF2075 domain-containing proteins. These constitute a large family of conserved proteins with over 7,000 members that can be found in organisms spanning all domains of life. Although the majority of members are encoded in bacteria, many are found in eukaryotes, a handful of which are in humans. Of these, the Schlafen (slfn) family proteins Slfn5, Slfn11, and Slfn13 have been shown to restrict the replication of diverse human viruses (Kim et al., 2021; Li et al., 2012; Valdez et al., 2019; Yang et al., 2018). Notably, Slfn13 also relies upon endonuclease activity for antiviral defense (Yang et al., 2018). Such functional conservation across phylogenetic boundaries has become a recurring theme in recently described defense systems (Bernheim et al., 2021; Cohen et al., 2019; Kazlauskienė et al., 2017; Niewoehner et al., 2017), and we anticipate that continued investigation of prokaryotic DUF2075 proteins has the potential to seed new insights into human immunity.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Bacterial strains and growth conditions
 - Phage propagation and enumeration
- METHOD DETAILS
 - Plate-based phage infection assays
 - Constructing pC194, pT181, and pTET-based plasmids
 - Electroporation into staphylococci
 - Homolog identification
 - Phylogenetic tree generation
 - Homolog neighborhood analysis
 - Phage adsorption assay
 - Cell growth and viability assays
 - Phage infection time course and quantitative PCR
 - Construction of pET28b-10His-Smt3-based plasmids
 - Transformation of *E. coli*
 - Purification of recombinant SERP2475
 - Nuclease assays
 - Helicase Assays
 - Mass spectrometry analysis
 - Phage hybrid generation and sequencing
 - Hybrid phage genome sequence analysis
 - Isolation and amplification of ISP escaper phages
 - Genomic DNA extraction of ISP escaper phages
 - Escaper phage DNA sequencing and analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.chom.2022.03.001>.

ACKNOWLEDGMENTS

A.H.-A. would like to acknowledge funding for this project from the National Science Foundation CAREER award (MCB/2054755) and an Investigators in the Pathogenesis of Infectious Disease award from the Burroughs Wellcome Fund (awarded to A.H.-A.).

AUTHOR CONTRIBUTIONS

Conceptualization, A.H.-A.; methodology, S.M.N.B., L.C.-Z., B.A., and A.H.-A.; software, B.A.; investigation, S.M.N.B., L.C.-Z., O.H., M.H., C.M.H., T.A.B., K.C., V.S.D., A.T., B.A., and A.H.-A.; writing—original draft, A.H.-A.; writing—review & editing, S.M.N.B., L.C.-Z., and O.H.; supervision and funding acquisition, A.H.-A.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 23, 2021

Revised: December 10, 2021

Accepted: March 1, 2022

Published: April 13, 2022

REFERENCES

- Armstrong, R.N. (2000). Mechanistic diversity in a metalloenzyme superfamily. *Biochemistry* 39, 13625–13632.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Pribelski, A.D., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477.
- Bari, S.M.N., Walker, F.C., Cater, K., Aslan, B., and Hatoum-Aslan, A. (2017). Strategies for editing virulent staphylococcal phages using CRISPR-Cas10. *ACS Synth. Biol.* 6, 2316–2325.
- Barylski, J., Kropinski, A.M., Alikhan, N.-F., and Adriaenssens, E.M.; Ictv Report Consortium (2020). ICTV virus taxonomy profile: *Herelleviridae*. *J. Gen. Virol.* 101, 362–363.
- Bergh, O., Børshheim, K.Y., Bratbak, G., and Heldal, M. (1989). High abundance of viruses found in aquatic environments. *Nature* 340, 467–468.
- Bernheim, A., Millman, A., Ofir, G., Meitav, G., Avraham, C., Shomar, H., Rosenberg, M.M., Tal, N., Melamed, S., Amitai, G., and Sorek, R. (2021). Prokaryotic viperins produce diverse antiviral molecules. *Nature* 589, 120–124.
- Bianco, P.R. (2021). The mechanism of action of the SSB interactome reveals it is the first OB-fold family of genome guardians in prokaryotes. *Protein Sci.* 30, 1757–1775.
- Blum, M., Chang, H.-Y., Chuguransky, S., Grego, T., Kandasamy, S., Mitchell, A., Nuka, G., Paysan-Lafosse, T., Qureshi, M., Raj, S., et al. (2021). The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res.* 49, D344–D354.
- Casjens, S.R., and Hendrix, R.W. (2015). Bacteriophage lambda: early pioneer and still relevant. *Virology* 479–480, 310–330.
- Cater, K., Dandu, V.S., Bari, S.M.N., Lackey, K., Everett, G.F.K., and Hatoum-Aslan, A. (2017). A novel *Staphylococcus* Podophage encodes a unique lysin with unusual modular design. *mSphere* 2, e00040–e00017.
- Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H., and Conti, E. (2011). Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. *Mol. Cell* 41, 693–703.
- Chou-Zheng, L., and Hatoum-Aslan, A. (2019). A type III-A CRISPR-Cas system employs degradosome nucleases to ensure robust immunity. *Elife* 8, e45393.
- Christensen, G.D., Baddour, L.M., and Simpson, W.A. (1987). Phenotypic variation of *Staphylococcus epidermidis* slime production *in vitro* and *in vivo*. *Infect. Immun.* 55, 2870–2877.
- 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., and de Hoon, M.J. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25, 1422–1423.
- 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.
- Culbertson, E.K., Bari, S.M.N., Dandu, V.S., Kriznik, J.M., Scopel, S.E., Stanley, S.P., Lackey, K., Hernandez, A.C., and Hatoum-Aslan, A. (2019). Draft genome sequences of *Staphylococcus* Podophages JBug18, Pike, Pontiff, and Pabna. *Microbiol. Resour. Announc.* 8, e00054–e00019.
- Deng, L., Garrett, R.A., Shah, S.A., Peng, X., and She, Q. (2013). A novel interference mechanism by a type IIIB CRISPR-Cmr module in *Sulfolobus*. *Mol. Microbiol.* 87, 1088–1099.
- Depardieu, F., Didier, J.P., Bernheim, A., Sherlock, A., Molina, H., Duclos, B., and Bikard, D. (2016). A eukaryotic-like serine/threonine kinase protects staphylococci against phages. *Cell Host Microbe* 20, 471–481.
- Doron, S., Melamed, S., Ofir, G., Leavitt, A., Lopatina, A., Keren, M., Amitai, G., and Sorek, R. (2018). Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* 359, eaar4120.
- Dupuis, M.-É., Villion, M., Magadán, A.H., and Moineau, S. (2013). CRISPR-Cas and restriction-modification systems are compatible and increase phage resistance. *Nat. Commun.* 4, 2087.
- Ehrlich, S.D. (1977). Replication and expression of plasmids from *Staphylococcus aureus* in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* 74, 1680–1682.
- Fairman-Williams, M.E., Guenther, U.-P., and Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. *Curr. Opin. Struct. Biol.* 20, 313–324.
- Freeman, M.E., Kenny, S.E., Lanier, A., Cater, K., Wilhite, M.C., Gamble, P., O'Leary, C.J., Hatoum-Aslan, A., Young, R.F.I., and Liu, M. (2019). Complete genome sequences of *Staphylococcus epidermidis* Myophages Quidivi, Terranova, and Twillingate. *Microbiol. Resour. Announc.* 8, e00598–e00519.
- Gao, L., Altae-Tran, H., Böhning, F., Makarova, K.S., Segel, M., Schmid-Burgk, J.L., Koob, J., Wolf, Y.I., Koonin, E.V., and Zhang, F. (2020). Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science* 369, 1077–1084.
- Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345.
- Goldberg, G.W., Jiang, W., Bikard, D., and Marraffini, L.A. (2014). Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. *Nature* 514, 633–637.
- Goldfarb, T., Sberro, H., Weinstock, E., Cohen, O., Doron, S., Charpak-Amikam, Y., Afik, S., Ofir, G., and Sorek, R. (2015). BREX is a novel phage resistance system widespread in microbial genomes. *EMBO J.* 34, 169–183.
- Gordeeva, J., Morozova, N., Sierro, N., Isaev, A., Sinkunas, T., Tsvetkova, K., Matlashov, M., Truncaite, L., Morgan, R.D., Ivanov, N.V., et al. (2019). BREX system of *Escherichia coli* distinguishes self from non-self by methylation of a specific DNA site. *Nucleic Acids Res.* 47, 253–265.
- Hampton, H.G., Watson, B.N.J., and Fineran, P.C. (2020). The arms race between bacteria and their phage foes. *Nature* 577, 327–336.
- Hatoum-Aslan, A., Maniv, I., and Marraffini, L.A. (2011). Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site. *Proc. Natl. Acad. Sci. USA* 108, 21218–21222.
- He, X., Byrd, A.K., Yun, M.-K., Pemble, C.W., Harrison, D., Yeruva, L., Dahl, C., Kreuzer, K.N., Raney, K.D., and White, S.W. (2012). The T4 phage SF1B

- helicase Dda is structurally optimized to perform DNA strand separation. *Structure* 20, 1189–1200.
- Hille, F., Richter, H., Wong, S.P., Bratović, M., Ressel, S., and Charpentier, E. (2018). The biology of CRISPR-Cas: backward and forward. *Cell* 172, 1239–1259.
- Hoikkala, V., Ravaniti, J., Díez-Villaseñor, C., Tirola, M., Conrad, R.A., McBride, M.J., Moineau, S., and Sundberg, L.-R. (2021). Cooperation between different CRISPR-Cas types enables adaptation in an RNA-targeting system. *mBio* 12, e03338–e03320.
- Hoskisson, P.A., Sumby, P., and Smith, M.C.M. (2015). The phage growth limitation system in *Streptomyces coelicolor* A(3)2 is a toxin/antitoxin system, comprising enzymes with DNA methyltransferase, protein kinase and ATPase activity. *Virology* 477, 100–109.
- Jiang, W., Maniv, I., Arain, F., Wang, Y., Levin, B.R., and Marraffini, L.A. (2013). Dealing with the evolutionary downside of CRISPR immunity: bacteria and beneficial plasmids. *PLoS Genet.* 9, e1003844.
- Kazlauskienė, M., Kostiuk, G., Venclovas, Č., Tamulaitis, G., and Siksnys, V. (2017). A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science* 357, 605–609.
- Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858.
- Khan, S.A., Carleton, S.M., and Novick, R.P. (1981). Replication of plasmid pT181 DNA *in vitro*: requirement for a plasmid-encoded product. *Proc. Natl. Acad. Sci. USA* 78, 4902–4906.
- Kim, E.T., Dybas, J.M., Kulej, K., Reyes, E.D., Price, A.M., Akhtar, L.N., Orr, A., Garcia, B.A., Boutell, C., and Weitzman, M.D. (2021). Comparative proteomics identifies Schlafen 5 (SLFN5) as a herpes simplex virus restriction factor that suppresses viral transcription. *Nat. Microbiol.* 6, 234–245.
- Klump, J., Lavigne, R., Loessner, M.J., and Ackermann, H.-W. (2010). The SPO1-related bacteriophages. *Arch. Virol.* 155, 1547–1561.
- Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., and Stamatakis, A. (2019). RAXML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 35, 4453–4455.
- Kronheim, S., Daniel-Ivad, M., Duan, Z., Hwang, S., Wong, A.I., Mantel, I., Nodwell, J.R., and Maxwell, K.L. (2018). A chemical defence against phage infection. *Nature* 564, 283–286.
- Kwong, S.M., Ramsay, J.P., Jensen, S.O., and Firth, N. (2017). Replication of staphylococcal resistance plasmids. *Front. Microbiol.* 8, 2279.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.
- Li, M., Kao, E., Gao, X., Sandig, H., Limmer, K., Pavon-Eternod, M., Jones, T.E., Landry, S., Pan, T., Weitzman, M.D., and David, M. (2012). Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11. *Nature* 491, 125–128.
- Lopatina, A., Tal, N., and Sorek, R. (2020). Abortive infection: bacterial suicide as an antiviral immune strategy. *Annu. Rev. Virol.* 7, 371–384.
- Ludwig, M.R., Kojima, K., Bowersock, G.J., Chen, D., Jhala, N.C., Buchsbaum, D.J., Grizzle, W.E., Klug, C.A., and Mobley, J.A. (2016). Surveying the serologic proteome in a tissue-specific kras(G12D) knockin mouse model of pancreatic cancer. *Proteomics* 16, 516–531.
- Makarova, K.S., Wolf, Y.I., Iranzo, J., Shmakov, S.A., Alkhnbashi, O.S., Brouns, S.J.J., Charpentier, E., Cheng, D., Haft, D.H., Horvath, P., et al. (2020). Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* 18, 67–83.
- Makarova, K.S., Wolf, Y.I., Snir, S., and Koonin, E.V. (2011). Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. *J. Bacteriol.* 193, 6039–6056.
- Marc, A., Cater, K., Kongari, R., Hatoum-Aslan, A., Young, R.F., and Liu, M. (2019). Complete genome sequence of *Staphylococcus aureus* Siphophage Lorac. *Microbiol. Resour. Announc.* 8, e00586–e00519.
- Marraffini, L.A., and Sontheimer, E.J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322, 1843–1845.
- McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y.M., Buso, N., Cowley, A.P., and Lopez, R. (2013). Analysis tool web services from the EMBL-EBI. *Nucleic Acids Res.* 41, W597–W600.
- 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.
- Monk, I.R., Shah, I.M., Xu, M., Tan, M.W., and Foster, T.J. (2012). Transforming the untransformable: application of direct transformation to manipulate genetically *Staphylococcus aureus* and *Staphylococcus epidermidis*. *mBio* 3, e00277–11.
- Nair, D., Memmi, G., Hernandez, D., Bard, J., Beaume, M., Gill, S., Francois, P., and Cheung, A.L. (2011). Whole-genome sequencing of *Staphylococcus aureus* strain RN4220, a key laboratory strain used in virulence research, identifies mutations that affect not only virulence factors but also the fitness of the strain. *J. Bacteriol.* 193, 2332–2335.
- Narajczyk, M., Barańska, S., Węgrzyn, G. (2007). Switch from θ to σ replication of bacteriophage λ DNA: factors involved in the process and a model for its regulation. *Mol. Genet. Genomics* 278, 65–74.
- Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274.
- Niewoehner, O., Garcia-Doval, C., Rostol, J.T., Berk, C., Schwede, F., Bigler, L., Hall, J., Marraffini, L.A., and Jinek, M. (2017). Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature* 548, 543–548.
- Ofir, G., Melamed, S., Sberro, H., Mukamel, Z., Silverman, S., Yaakov, G., Doron, S., and Sorek, R. (2018). DISARM is a widespread bacterial defence system with broad anti-phage activities. *Nat. Microbiol.* 3, 90–98.
- Oliveira, H., Sampaio, M., Melo, L.D.R., Dias, O., Pope, W.H., Hatfull, G.F., and Azeredo, J. (2019). Staphylococci phages display vast genomic diversity and evolutionary relationships. *BMC Genomics* 20, 357.
- Pinilla-Redondo, R., Mayo-Muñoz, D., Russel, J., Garrett, R.A., Randau, L., Sørensen, S.J., and Shah, S.A. (2020). Type IV CRISPR-Cas systems are highly diverse and involved in competition between plasmids. *Nucleic Acids Res.* 48, 2000–2012.
- Potter, S.C., Luciani, A., Eddy, S.R., Park, Y., Lopez, R., and Finn, R.D. (2018). HMMER web server: 2018 update. *Nucleic Acids Res.* 46, W200–W204.
- Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. *Nat. Biotechnol.* 29, 24–26.
- Salas, M., Holguera, I., Redrejo-Rodríguez, M., and de Vega, M. (2016). DNA-binding proteins essential for protein-primed bacteriophage Φ 29 DNA replication. *Front. Mol. Biosci.* 3, 37.
- Samai, P., Pyenson, N., Jiang, W., Goldberg, G.W., Hatoum-Aslan, A., and Marraffini, L.A. (2015). Co-transcriptional DNA and RNA cleavage during Type III CRISPR-Cas immunity. *Cell* 161, 1164–1174.
- Sayyari, E., and Mirarab, S. (2018). Testing for polytomies in phylogenetic species trees using quartet frequencies. *Genes (Basel)* 9, 1–22.
- Schiltz, C.J., Lee, A., Partlow, E.A., Hosford, C.J., and Chappie, J.S. (2019). Structural characterization of Class 2 OLD family nucleases supports a two-metal catalysis mechanism for cleavage. *Nucleic Acids Res.* 47, 9448–9463.
- Shereda, R.D., Kozlov, A.G., Lohman, T.M., Shereda, R.D., and Kozlov, A.G. (2008). SSB as an organizer/mobilizer of genome maintenance complexes. *Crit. Rev. Biochem. Mol. Biol.* 43, 289–318.
- Silas, S., Lucas-Elio, P., Jackson, S.A., Aroca-Crevillén, A., Hansen, L.L., Fineran, P.C., Fire, A.Z., and Sánchez-Amat, A. (2017). Type III CRISPR-Cas systems can provide redundancy to counteract viral escape from type I systems. *Elife* 6, 1–27.

- Steczkiwicz, K., Muszewska, A., Knizewski, L., Rychlewski, L., and Ginalski, K. (2012). Sequence, structure and functional diversity of PD-(D/E)XK phosphodiesterase superfamily. *Nucleic Acids Res.* **40**, 7016–7045.
- Sumby, P., and Smith, M.C.M. (2002). Genetics of the phage growth limitation (Pgl) system of *Streptomyces coelicolor* A3 (2). *Mol. Microbiol.* **44**, 489–500.
- Tengerdy, R.P., Nagy, J.G., and Martin, B. (1967). Quantitative measurement of bacterial growth by the reduction of tetrazolium salts. *Appl. Microbiol.* **15**, 954–955.
- Uyen, N.T., Park, S.-Y., Choi, J.-W., Lee, H.-J., Nishi, K., and Kim, J.-S. (2009). The fragment structure of a putative HsdR subunit of a type I restriction enzyme from *Vibrio vulnificus* YJ016: implications for DNA restriction and translocation activity. *Nucleic Acids Res.* **37**, 6960–6969.
- Valdez, F., Salvador, J., Palermo, P.M., Mohl, J.E., Hanley, K.A., Watts, D., and Llano, M. (2019). Schlafen 11 restricts Flavivirus replication. *J. Virol.* **93**, e00104–e00119.
- Vandersteegen, K., Mattheus, W., Ceysens, P.J., Bilocq, F., Vos, D. De, Pimay, J.P., Noben, J.P., Merabishvili, M., Lipinska, U., Hermans, K., and Lavigne, R. (2011). Microbiological and molecular assessment of bacteriophage ISP for the control of *Staphylococcus aureus*. *PLoS One* **6**, e24418.
- Vybiral, D., Takáč, M., Loessner, M., Witte, A., Von Ahsen, U., and Bläsi, U. (2003). Complete nucleotide sequence and molecular characterization of two lytic *Staphylococcus aureus* phages: 44AHJD and P68. *FEMS Microbiol. Lett.* **219**, 275–283.
- Weisblum, B., Graham, M.Y., Gryczan, T., and Dubnau, D. (1979). Plasmid copy number control: isolation and characterization of high-copy-number mutants of plasmid pE194. *J. Bacteriol.* **137**, 635–643.
- Wick, R.R., Schultz, M.B., Zobel, J., and Holt, K.E. (2015). Genome analysis Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* **31**, 3350–3352.
- Yang, J.-Y., Deng, X.-Y., Li, Y.-S., Ma, X.-C., Feng, J.-X., Yu, B., Chen, Y., Luo, Y.-L., Wang, X., Chen, M.-L., et al. (2018). Structure of Schlafen13 reveals a new class of tRNA/rRNA-targeting RNase engaged in translational control. *Nat. Commun.* **9**, 1165.
- Zhou, C., Pourmal, S., and Pavletich, N.P. (2015). Dna2 nuclease-helicase structure, mechanism and regulation by Rpa. *Elife* **4**, e09832.
- Zimmermann, L., Stephens, A., Nam, S.Z., Rau, D., Kübler, J., Lozajic, M., Gabler, F., Söding, J., Lupas, A.N., and Alva, V. (2018). A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* **430**, 2237–2243.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
<i>Staphylococcus epidermidis</i> RP62a	ATCC (Christensen et al., 1987)	Genbank: NC_002976.3
<i>Staphylococcus epidermidis</i> LM1680 and other deletion mutants of RP62a	Luciano Marraffini (Jiang et al., 2013)	N/A
<i>Staphylococcus aureus</i> RN4220	Luciano Marraffini (Nair et al., 2011)	N/A
<i>Escherichia coli</i> DH5a	ATCC	Genbank: CP080399.1
<i>Escherichia coli</i> Rosetta2 (DE3)	Novagen	Genbank: NZ_CP083274.1
Bacteriophage Andhra	Previous study (Cater et al., 2017)	Genbank: KY442063
Bacteriophage JBug18	Previous study (Culbertson et al., 2019)	Genbank: MH972263
Bacteriophage Pontiff	Previous study (Culbertson et al., 2019)	Genbank: MH972262
Bacteriophage Pike	Previous study (Culbertson et al., 2019)	Genbank: MH972261
Bacteriophage CNP _x	Luciano Marraffini (Depardieu et al., 2016)	Genbank: NC_031241.1
Bacteriophage Twillingate	Previous study (Freeman et al., 2019)	Genbank: MH321491.1
Bacteriophage Quidividi	Previous study (Freeman et al., 2019)	Genbank: MH321490.1
Bacteriophage ISP	Luciano Marraffini (Vandersteegen et al., 2011)	Genbank: NC_047720.1
Bacteriophage Lorac	Previous study (Marc et al., 2019)	Genbank: MH321492.1
Bacteriophage Pabna	Previous study (Culbertson et al., 2019)	Genbank: NC_048107.1
Chemicals, peptides, and recombinant proteins		
g-32P-ATP	Perkin Elmer	Cat. # SP-100
2,3,5 triphenyltetrazolium chloride (TTC)	Fisher Scientific	Cat. # T052025G
SUMO protease	MCLAB, http://www.mclab.com/SUMO-Protease.html	Cat. # SP-100
T4 polynucleotide kinase	New England Biolabs	Cat. # M0201L
DpnI	New England Biolabs	Cat. # R0176S
HisPur™ Ni-NTA Resin	Thermo Fisher	Cat. # M0201L
SERP2475 (Nhi)	This manuscript	Genbank: NC_002976.3_SERP2475
Critical commercial assays		
EZNA Cycle Pure Kit	Omega Bio-tek via VWR	Cat. # 101318-892
EZNA Plasmid DNA Mini Kit	Omega Bio-tek via VWR	Cat. # 101318-898
Wizard® Genomic DNA Purification Kit	Promega Corporation via VWR	Cat. # A1120
PerfeCTa® SYBR® Green SuperMix	Quanta Biosciences via VWR	Cat. # 101414-150
Deposited data		
Sanger sequencing reads for JBug18-Andhra hybrids	This manuscript	Figshare: 10.6084/m9.figshare.9598040
Illumina reads for ISP and escaper mutants	This manuscript	BioProject: PRJNA786381
Python code library used for homolog and neighborhood analysis	This manuscript	GitHub: https://github.com/ahatoum/Nhi
Python code written to analyze phage hybrid data	This manuscript	GitHub: https://github.com/ahatoum/Hybrid-phage-genome-sequence-analysis
Oligonucleotides		
DNA oligonucleotides for PCR and Sanger sequencing (multiple)	Eurofins MWG Operon	Table S8
DNA oligonucleotides (PAGE purified) for biochemical assays	Eurofins MWG Operon	Table S7

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Recombinant DNA		
pBR322	NEB	Cat. # N3033S
pUC19	NEB	Cat. # N3041S
pC194	ATCC (Ehrlich, 1977)	Genbank: NC_002013.1
pC194-based constructs (i.e. pSERP-, multiple)	(Hatoum-Aslan et al., 2011) and this manuscript	Table S8
pT181	(Khan et al., 1981)	Genbank: J01764.1
pT181-based constructs (multiple)	This manuscript	Table S8
pE194	ATCC (Weisblum et al., 1979)	Genbank: M17811.1
pE194-based constructs (i.e. pTET-, multiple)	(Samai et al., 2015) and This manuscript	Table S8
pET28b	Novagen	Cat. # 69865
pET28b-based constructs (multiple)	This manuscript	Table S8
Software and algorithms		
ImageQuant TL v. 8.1	GE Healthcare/Life Sciences	RRID: SCR_014246
Biopython v. 1.7.8	(Cock et al., 2009)	https://pypi.org/project/biopython/
IQ-TREE21 multi-core v. 1.6.12	(Nguyen et al., 2015)	http://www.iqtree.org/release/v1.6.12
RAxML-NG v. 1.0.2	(Kozlov et al., 2019)	https://github.com/amkozlov/raxml-ng
Fig Tree v. 1.4.4	Rambaut Lab	http://tree.bio.ed.ac.uk/software/figtree/
hmmer 3.3.2	(Potter et al., 2018)	hmmer.org
Clustal Omega Multiple Sequence Alignment tool	(McWilliam et al., 2013)	https://www.ebi.ac.uk/Tools/msa/clustalo/
bcl2fastq v. 1.8.4	Illumina	https://support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software/downloads.html
FastQC v. 0.11.9	Babraham Bioinformatics	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
SPAdes v. 3.15.3	(Bankevich et al., 2012)	https://cab.spbu.ru/software/spades/
Bandage v. 0.8.1	(Wick et al., 2015)	https://rrwick.github.io/Bandage/
bowtie2 v. 2.4.4	(Langmead and Salzberg, 2012)	https://github.com/BenLangmead/bowtie2
samtools v. 1.13	(Li et al., 2009)	https://github.com/samtools/samtools/releases/
igvtools v. 2.11.1	(Robinson et al., 2011)	https://software.broadinstitute.org/software/igv/2.11.x

RESOURCE AVAILABILITY

Lead contact

Requests for further information and/or reagents and resources should be directed to and will be fulfilled by the lead contact, Asma Hatoum-Aslan (ahatoum@illinois.edu).

Materials availability

Phages, mutant derivatives, and constructs generated in this study will be made available upon written request to the [lead contact](#).

Data and code availability

- Raw Illumina sequencing reads for phage ISP and mutant variants have been deposited at NCBI Sequence Read Archive (SRA) under BioProject ID [PRJNA786381](#) and are publicly available as of the date of publication. Accession numbers are listed in the [key resources table](#). The raw sequence reads for JBug18-Andhra hybrid phages have been deposited at Figshare and are publicly available. DOIs are listed in the [key resources table](#). This paper analyzes existing, publicly available data for draft genome sequences of Andhra, JBug18, Pontiff, and Pike. Accession numbers are listed in the [key resources table](#). All other data reported in this paper will be shared by the [lead contact](#) upon request.
- All original code has been deposited at GitHub and is publicly available as of the date of publication. DOIs are listed in the [key resources table](#).

- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Bacterial strains and growth conditions

S. epidermidis RP62a ([Christensen et al., 1987](#)) and mutant variants were a generous gift from Luciano Marraffini. *S. epidermidis* strains were grown in Brain Heart Infusion (BHI, BD Diagnostics), *S. aureus* RN4220 ([Nair et al., 2011](#)) was grown in Tryptic Soy Broth (TSB, BD Diagnostics), *E. coli* DH5 α was grown in Luria Bertani (LB) broth (VWR), and *E. coli* Rosetta2 (DE3) was grown in Terrific Broth (TB, VWR) for protein purification. Growth media was supplemented with the following: 10 μ g/ml chloramphenicol (to select for pC194-based plasmids), 10 μ g/ml tetracycline (to select pT181-based plasmids), 10 μ g/ml erythromycin (to select for pTET-based plasmids), 15 μ g/ml neomycin (to select for *S. epidermidis* cells), 30 μ g/ml chloramphenicol (to select for *E. coli* Rosetta2 plasmids) and 50 μ g/ml kanamycin (to select for pET28b-10HisSmt3-based plasmids). All bacterial strains were grown at 37°C unless otherwise indicated. Liquid cultures were propagated with agitation in an orbital shaker set to 180–200 rpm. Strains were routinely authenticated via PCR amplification and sequencing of genomic regions unique to each strain.

Phage propagation and enumeration

S. epidermidis phages (Andhra, JBug18, Pontiff, Pike, Quidividi, and Twillingate) and *S. aureus* phages (ISP, Lorac, and Pabna) were propagated on their respective host strains (LM1680 and RN4220, respectively). Concentrated phage stocks were prepared by combining 1–5 purified phage plaques into 500 μ l of TSB and vortexing for 30 sec. Each suspension was then subjected to centrifugation at $\sim$ 15,000 $\times g$ for 2 min to pellet agar and cells. The resulting phage lysate (i.e. supernatant) was passed through a 0.45 μ m syringe filter and then combined with overnight host culture (diluted 1:100) in 7 ml of Heart Infusion Agar (HIA, Hardy Diagnostics, prepared at 0.3 $\times$ concentration) supplemented with 5 mM CaCl₂. Phage-host mixtures were then poured atop a plate containing a solid layer of TSA supplemented with 5 mM CaCl₂, and the top agar layer was allowed to solidify $\sim$ 10 min at room temperature. Following overnight incubation at 37°C, the entire top agar layer was harvested and resuspended in 20 mL of fresh TSB. The suspension was vortexed for 5 min to release phages from the agar, and then subjected to centrifugation at 10,000 $\times g$ for 10 min to remove agar and cell debris. The resulting concentrated phage lysate was passed through a 0.45 μ m bottle filter, and phage concentrations were determined using the double-agar overlay method as described in [Cater et al. \(2017\)](#). Briefly, HIA (prepared at 0.5 $\times$ concentration) was equilibrated to 55°C. Equilibrated HIA supplemented with 5 mM CaCl₂ was combined with an overnight culture of the bacterial host strain (at a 1:100 final dilution). 4 ml of this mixture was overlaid atop TSA plates containing 5 mM CaCl₂ and allowed to solidify on the benchtop for 10 min. In the meantime, ten-fold dilutions (10^0 – 10^{-7}) of concentrated phage lysate were prepared and 10 μ l spots were dropped atop the semisolid HIA layer, allowed to air dry, and incubated overnight at 37°C. The phage concentration (i.e. titer) in plaque-forming units per ml (pfu/ml) was determined by the following formula: ((number of plaques counted on the most diluted spot)/(dilution factor))*100. Phage stocks were stored at 4°C. Phages were routinely authenticated via PCR amplification and sequencing of genomic regions unique to each phage.

METHOD DETAILS

Plate-based phage infection assays

For all plate infection assays, 10-fold dilutions of concentrated phage stocks were spotted atop a lawn of cells and enumerated using the protocol described in the section above titled “phage propagation and enumeration”. For assays using the anhydrotetracycline-(aTc-) inducible system ([Figure 2](#)), plates and top agar were supplemented with 15 or 30 mg/L aTc, as indicated. All graphs show an average of triplicate measurements ($\pm$ S.D.) as a representative of at least three independent trials.

Constructing pC194, pT181, and pTET-based plasmids

All pC194-, pT181-, and pTET-based plasmids were constructed using either inverse PCR or Gibson Assembly ([Gibson et al., 2009](#)) with the primers listed in [Table S8](#). pAH011 ([Hatoum-Aslan et al., 2011](#)), a derivative of pC194 ([Ehrlich, 1977](#)), was used as the backbone for plasmids designated as pSERP- in this study. Plasmid pT181 ([Khan et al., 1981](#)) was used as backbone for pT181-*gp03* and pT181-*gp0304*. pTarget ([Samai et al., 2015](#)), a derivative of pE194 ([Weisblum et al., 1979](#)), was used as the backbone for all plasmids designated as pTET- in this study. All base plasmids (pC194, pT181, and pE194) are extrachromosomal/ectopic plasmids that maintain a high copy number (10–20 per cell) in staphylococci ([Khan et al., 1981](#); [Kwong et al., 2017](#); [Weisblum et al., 1979](#)). All assembled plasmids were first introduced into *S. aureus* RN4220 (pC194- and pE194-based plasmids) or OS2 (pT181-based plasmids) via electroporation (described in the section below), and inserted sequences were confirmed by PCR amplification and Sanger sequencing (performed by Eurofins MWG Operon) using primers shown in [Table S8](#). At least two transformants were confirmed by sequencing and at least one of each construct was purified using EZNA Plasmid Mini Kit (Omega Bio-tek), and where indicated, introduced into *S. epidermidis* LM1680 via electroporation.

Electroporation into staphylococci

Electrocompetent cells were prepared as described in ([Monk et al., 2012](#)). Briefly, 10 mL of overnight culture was diluted to OD600 = 0.5 using fresh media. Diluted cultures were incubated for 30 min at 37°C, then ice-shocked for 10 min. All subsequent steps were

performed at 4°C or on ice. Cells were pelleted at 4000 x *g* for 10 min, then washed twice with equal volume with ice-cold water. Pelleted cells were further washed twice with ice-cold 10% glycerol using 1/20- and 1/25- the volume of culture, respectively. Lastly, cells were resuspended in 1/200 the volume of initial culture and saved at -80°C in 50 µl aliquots. For electroporation, ligated constructs, Gibson assembled constructs, or purified plasmids, were dialyzed against sterile water on a 0.022 µm filter for 20 minutes. Meanwhile, competent cells were thawed on ice for 5 min, and then left at room temperature for another 5 min. Cells were then pelleted via centrifugation at 5,000 x *g* for 1 min. The pellet was resuspended in 50 µl of sterilized 10% glycerol containing 500 mM sucrose and the entire amount of the dialyzed plasmid was added into the cell suspension. The mixture was then transferred into a 2 mm electroporation cuvette (VWR) and pulsed at 29 kV/cm, 100W, and 25 mF with a GenePulser Xcell instrument (Bio-Rad). Cells were then allowed to recover in 1 ml of sterile TSB containing 500 mM sucrose at 37°C with agitation for 2 hr. Recovered cells (200 µl) were plated on TSA or BHI supplemented with appropriate antibiotics and incubated at 37°C. Transformants were recovered on the following day.

Homolog identification

To identify homologs, the amino acid sequence of SERP2475 was independently queried against three databases (prok_complete_genomes, refseq_genomes, and nt) using NCBI's tBLASTn webserver and hits were downloaded in xml format. Using a Python script utilizing Biopython v. 1.7.8 functions (Cock et al., 2009), xml files were parsed and combined into a unique set of hits. Fully annotated complete genomes containing the homologs from the combined list of hits were downloaded from NCBI in genbank (gbk) format (~3GB). Each genome was parsed and coding sequence (CDS) features of corresponding BLAST hits were extracted. Proteins with unique accession numbers and unique sequences (excluding pseudogenes) were retained and combined into a fasta file. Hits shorter than 200 amino acids were eliminated from the list to obtain the final set of homologs in distinct genetic backgrounds (Table S1).

Phylogenetic tree generation

An iterative process was used to select the final set of 100 homologs and build the tree. The fasta file with all identified homologs was first submitted to the MAFFT webserver (used May 8th, 2021, <https://mafft.cbrc.jp/alignment/server/>) to obtain a multiple sequence alignment (MSA) with the E-INS-I option selected along with the remainder parameters set to their default values. Upon inspection of the MSA, low-scoring homologs that were also observed to be from adequately represented genera/species but introducing large gaps in the alignment were removed from consideration. The MSA computation was then repeated as above. Using this second MSA, a preliminary phylogenetic tree was generated with IQ-TREE21 multi-core v. 1.6.12 (Nguyen et al., 2015) with the optimal substitution model LG+R6 that provided the lowest Bayesian information criterion. One thousand ultra-fast bootstraps were performed to evaluate node support (options -bb 1000 -wbt). Upon inspection of the resulting tree, hard polytomies (Sayyari and Mirarab, 2018) that resulted in many extremely short branches were identified and, to preserve the phylogenetic diversity, only one representative homolog for each were retained to perform the final MAFFT alignment. A final MSA was computed with the remaining homologs (n=100 including SERP2475) and the final phylogenetic tree was computed using IQ-TREE with the same parameters noted above. RAXML-NG v. 1.0.2 (Kozlov et al., 2019) was used to confirm the tree with similar corresponding parameters (100 bootstraps). Fig Tree v. 1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>) was used for tree visualization, and Adobe Illustrator was used to overlay highlights and markings relevant to the study.

Homolog neighborhood analysis

For the neighborhood analysis, the genomes containing the original 303 homologs (including SERP2475) were parsed and the amino acid sequences of neighboring proteins encoded on either side of the Blast hit (window sizes of 10, 20, and 30) were extracted into individual fasta files using a Python script. Pfam 34.0 database was downloaded (Pfam-A.hmm) and hmmpress (hmmer 3.3.2, (Potter et al., 2018)) was used to index it. For each of the neighborhood fasta files, a Python script utilizing hmmsearch was used to obtain and generate a new set of files that included protein family (pfam) predictions of its contents. Predicted pfams were then searched against a set of 306 pfams with known defense-related functions (Table S3)—this list was compiled from the old and newly-identified defense pfams cited in Gao et al. (2020). A Python script was then developed and utilized to determine the defense related neighbors for each homolog and analyzed to generate the plots for each neighbor window size.

Phage adsorption assay

Overnight cultures of *S. epidermidis* LM1680 bearing pSERP-2475 or the empty vector were diluted 1:100 in fresh BHI supplemented with antibiotics and 5 mM CaCl₂, and incubated at 37°C with agitation for one hour. Andhra or JBug18 were then added to cultures (0.01:1 phage:cell ratio) and incubated at 37°C for 10 min. Cells along with adsorbed phages were pelleted at 8000 x *g* for 5 min at 4°C and resulting supernatants were passed through 0.45 µm syringe filter. The number of free phages in the supernatants were enumerated by the double-agar overlay method as described in the section above ("Plate-based phage infection assays" (Cater et al., 2017). The number of adsorbed phages were determined by subtracting the number of phages in suspension from the number that was initially added. Triplicate samples were prepared for each treatment, and two independent trials were conducted.

Cell growth and viability assays

For cell growth and viability assays, 200 µl of the bacterial cultures were distributed into a 96-well microtiter plate (into triplicate wells for each treatment), and phages were added to cells at ratios of 1:1, 5:1, or 10:1. To generate growth curves, plates were incubated at

37°C with agitation in a Tecan Infinite 200 Pro and OD600 measurements were taken every 15 minutes for a period of 800 minutes. For the cell viability assay, bacteria-phage mixtures were incubated at 37°C with agitation for five hours, enough time for several phage replication cycles (latent period ~30 min). 25 µl of 0.1% (w/v) 2,3,5 triphenyltetrazolium chloride (TTC, Fisher Scientific) was added into each well and the microtiter plate was incubated at 37°C for an additional 30 mins to allow the colorless TTC to become enzymatically reduced to the red 1,3,5-triphenylformazan product by actively growing bacterial cells. The relative numbers of viable cells were then determined by measuring the absorbance at 540 nm. Triplicate measurements were taken in each trial, and three independent trials were conducted.

Phage infection time course and quantitative PCR

Phage infection time course assays were conducted in liquid media as previously described (Chou-Zheng and Hatoum-Aslan, 2019). Briefly, *S. epidermidis* LM1680 mid-log cells bearing *pSERP-2475* or the empty vector were infected with Andhra or JBug18 (phage:cell ratio of 0.5:1), cells were harvested at 0-, 10-, or 20- minutes post-infection, and their total DNA was extracted. Each qPCR reaction (25 µl) contained 500 ng of total DNA as template, 0.4 nM of phage-specific primers (N233 and N234) or host-specific primers (S001 and S002) (Table S8), and 1X PerfeCTa SYBR Green SuperMix (Quanta Biosciences). Separate standard reactions containing 10^2 – 10^9 DNA molecules were also prepared using purified Andhra phage DNA extract, JBug phage DNA extract, or bacterial genomic DNA extract. A CFX Connect Real-Time PCR Detection System (Bio-Rad) was used to amplify the DNA templates as follows: one cycle, 95°C for 3 min; 40 cycles, 95°C for 10 sec and 55°C for 30 sec. Phage DNA copy number was normalized against host values, using the glyceraldehyde-3-phosphate dehydrogenase (*gap*) gene, and the normalized value for the 0 min time point was set to one to obtain the relative DNA abundance for the rest of the time points as described previously (Chou-Zheng and Hatoum-Aslan, 2019). Briefly, relative DNA abundance (i.e. fold difference) was determined using the following equation from the “Real-Time PCR Handbook” (ThermoFisher Scientific): Fold Difference = $(E_{\text{target}})^{Ct_{\text{target}}} / (E_{\text{normalizer}})^{\Delta Ct_{\text{normalizer}}}$, where $E = 10^{(-1/\text{slope})}$, $Ct_{\text{target}} = Ct_{\text{target-calibrator}} - Ct_{\text{target-samples}}$, and $\Delta Ct_{\text{normalizer}} = Ct_{\text{normalizer-calibrator}} - Ct_{\text{normalizer-samples}}$. Triplicate measurements were taken for each of two independent trials.

Construction of pET28b-10His-Smt3-based plasmids

pET28b-10His-Smt3-SERP2475, pET28b-10His-Smt3-SERP2475-230AA were constructed via Gibson assembly (Gibson et al., 2009). Inserts were amplified from the *S. epidermidis* RP62a genome, and the backbone was amplified from a pET28b-His10Smt3 template using the primers listed in Table S8. The backbone was further subjected to DpnI (NEB) digestion. Then, inserts and backbones were purified using the EZNA Cycle Pure Kit (Omega Bio-Tek), combined in a 10:1 ratio, and subjected to Gibson Assembly. Amino acid substitutions were introduced into pET28b-10His-Smt3-SERP2475, or pET28b-10His-Smt3-SERP2475-230AA, via inverse PCR using with the primers listed in Table S8. PCR products were subjected to DpnI (NEB) digestions and purified with EZNA Cycle Pure Kit. Purified products were subjected to 5' phosphorylation with T4 polynucleotide kinase (NEB) and circularization with T4 DNA ligase (NEB) according to the manufacturer's instructions. Ligated and Gibson assembled constructs were introduced into chemically competent *E. coli* DH5α cells by heat shock (see section below titled “Transformation of *E. coli*” for details). At least three transformants were confirmed to have the desired sequence via PCR and Sanger sequencing (performed by Eurofins MWG Operon) using primers shown in Table S8. Two of the confirmed plasmids were purified using the EZNA Plasmid Mini Kit and introduced into *E. coli* Rosetta2 (DE3) cells for protein purification.

Transformation of *E. coli*

For the preparation of chemically-competent *E. coli*, overnight culture was diluted 1:100 in LB and incubated at 37°C with agitation until the OD600 reached ~0.5. The culture was placed on ice for 10 minutes, and cells were pelleted by centrifugation at 4000 x g for 5 min. The resulting pellet was resuspended in 1/10 the culture volume in transformation and storage (TSS) buffer (85% LB medium, 10% (w/v) PEG MW 8000, 5% (v/v) DMSO, 50 mM magnesium chloride). Cells were dispensed in 50 µl aliquots and stored at -80°C. For transformation, aliquots were thawed on ice for 10 min and combined with 5 µl Gibson assembled product or 1 µl purified plasmid. The mixture was kept on ice for 30 minutes and then subjected to heat-shock at 42°C for 30 seconds and immediately placed on ice for 2 min. One ml of fresh LB was added directly into the tube and incubated at 37°C for 1 hr for recovery. Finally, 200 µl was plated on an LB-agar plate containing appropriate antibiotics and incubated overnight at 37°C.

Purification of recombinant SERP2475

Recombinant SERP2475 and mutant variants encoded in pET28b-10His-Smt3-based plasmids were overexpressed and purified from *E. coli* Rosetta2 (DE3) as described previously with some modifications (Chou-Zheng and Hatoum-Aslan, 2019). Briefly, overnight cultures were diluted 1:100 in 1 L (for the truncated version) or 2 L (for the full-length version) of TB supplemented with appropriate antibiotics. Once the OD₆₀₀ reached 0.5–0.6, cell-growth was arrested on ice for 20 minutes, and protein expression was induced with 0.3 mM isopropyl-1-thio-β-d-galactopyranoside (IPTG) and 2% ethanol. Induction proceeded 16–18 hr at 17°C with constant shaking. Cells were harvested and washed with cold PBS Buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄). All subsequent steps were performed at 4°C. Each one-liter pellet was suspended in 30 ml of Buffer A (50 mM Tris-HCl, pH 9.5, 1.25 M NaCl, 200 mM Li₂SO₄, 10% sucrose, 15 mM Imidazole) containing one complete EDTA-free protease inhibitor tablet (Roche), 0.1 mg/ml lysozyme, and 0.1% Triton X-100. After 1 hr rotation, lysed cells were sonicated, and the insoluble materials were removed via centrifugation and filtration. Then, 3 ml (for 1-L pellet of truncated version), or 4 ml (for 2-L pellet

of full-length version), of slurry Ni^{2+} -NTA-agarose resin (ThermoFisher) was pre-equilibrated with Buffer A and mixed with the cleared lysates. After 1 hr incubation, the resin was collected via centrifugation and washed with 40 ml of Buffer A per one-liter pellet. The 3 ml (for 1-L pellet of truncated version), or 4 ml (for 2-L pellet of full-length version) resins were then transferred to a 5-ml gravity column (G-Biosciences) and further washed with 25 ml of Buffer A. Proteins (1 ml) were eluted into a tube containing 1 ml IMAC buffer (50 mM Tris-HCl pH 9.5, 250 mM NaCl, 10% glycerol). Proteins were eluted stepwise with 3 ml of IMAC buffer containing 50-, 100-, 200-, and 500-mM imidazole, respectively. The 200 mM imidazole aliquots were pooled and mixed with SUMO Protease (Mclab, 1000 U) and supplied buffer (salt-free) to remove the 10His-Smt3-tag. Samples were dialyzed against IMAC buffer containing 25 mM Imidazole for 3 hr. Then, 2 ml (for the truncated version), or 1.5 ml (for the full-length version), of slurry Ni^{2+} -NTA-agarose resin was equilibrated with IMAC Buffer containing 25 mM Imidazole, and incubated with the dialysate for 1 hr. The resin was collected in a 5-ml gravity column, and tag-free proteins were collected and concentrated using a 10K MWCO centrifugal filter (PALL). Concentrated proteins were further purified by size exclusion chromatography using Superdex 75 Increase 10/300 GL (Cytiva) (for the truncated version) and Superdex 200 Increase 10/300 GL (Cytiva) (for the full-length version). Collected protein fractions were subjected to nuclease assays. Fractions with the highest peak concentration were combined and concentrated for nuclease time course and helicase assays. Proteins were resolved on 15% SDS-PAGE run at 120 V for 1.5 hours and visualized with Coomassie G-250, and the concentrations were determined using Bradford reagent (Bio-Rad). At least 4 independent protein purifications were conducted for the wild-type (full-length and truncated) enzymes, and 2 independent protein purifications were conducted for each mutant variant.

Nuclease assays

For exonuclease assays, single stranded DNA substrates (Table S7) were labeled on their 5'-ends by incubating with T4 polynucleotide kinase and γ -[^{32}P]-ATP and subsequently purified over a G25 column (IBI Scientific). Radiolabeled substrates were combined with 7.5 μl of each protein fraction (for assays with individual fractions), or 130 pmols (for assays with the peak fraction) in 10 μl reactions containing nuclease buffer (25 mM Tris-HCl pH 7.5, 2 mM DTT) and 10 mM of MgCl_2 . Exonuclease reactions were incubated at 37°C for 20 minutes (for experiments with different fractions), or 5, 10, and 20 minutes. Reactions were stopped by adding an equal volume of 95% formamide loading buffer and resolved on a 15% Urea PAGE gel at 55 W for 1.5 hours. Gels were exposed to a storage phosphor screen and visualized using an Amersham Typhoon biomolecular imager. For nickase assays, 250 ng of plasmids pBR322 or pUC19 (NEB) were combined with 15 μl of protein fractions (for assays with individual fractions), or 260 pmols (for assays with the peak fraction), in 20 μl reactions containing nuclease buffer and 10 mM of MgCl_2 . Nickase reactions were incubated at 37°C for 60 minutes (for experiments with individual fractions), or 15, 30, and 60 minutes. Reactions were stopped by placing on ice for 5 min, followed by incubation with 10 μg of proteinase K for 20 minutes at room temperature. Samples were then resolved on a 1% agarose gel run at 120 V for 50 minutes and visualized with ethidium bromide under UV transillumination with an Azure 400 imager. ImageQuant software was used for densitometric analysis with the following settings: for exonuclease assays, minimum slope = 50, edge parameter = fixed at 16, and remaining settings were set to zero; for nickase assays, minimum slope = 100, edge parameter = fixed at 25, and remaining settings were set to zero. The fraction of substrate cleaved was determined using the following equation: density of cut substrates signal divided by the sum of densities of cut and uncut signals. Three independent trials were performed for each protein preparation.

Helicase Assays

Double-stranded DNA duplexes were prepared by combining 5'-radiolabeled ssDNA oligonucleotides and unlabeled complementary ssDNA oligonucleotides (Table S7) in a 1:2.5 molar ratio. The mixtures were heated to 95°C for 5 min and then slowly cooled down to room temperature over a period of 3 hours. The helicase assay was performed by first mixing radiolabeled DNA duplex with a 10-fold molar excess of unlabeled top-strand DNA to trap the complementary strand once unwound. This mixture was combined with 200 pmol SERP2475 in helicase buffer (25 mM Tris-HCl pH 9.0, 2 mM DTT, and 0.1 mg/ml BSA) supplemented with 2 mM MgCl_2 and 5 mM ATP in a 50 μl reaction. Reaction mixtures were incubated at 37°C for 1 h. As a separate positive control for unwinding, DNA substrates were heated to 95°C for 10 min in the absence of the enzyme. Reaction was stopped by adding 5 μl of the stop solution (0.1% [wt/vol] bromophenol blue, 0.1% [wt/vol] xylene cyanol, 8% [vol/vol] glycerol, 0.4% [wt/vol] SDS, 50 mM EDTA). Samples were resolved on an 8% (v/v) non-denaturing polyacrylamide gel at 130 V for 4 hours at 4°C. The gel was dried under vacuum at 80°C, exposed to a storage phosphor screen and visualized using an Amersham Typhoon biomolecular imager. Four independent trials were performed. ImageQuant software was used for densitometric analysis with default settings. The fraction of unwound (ssDNA) was determined using the following equation: density of ssDNA signal divided by the sum of densities of dsDNA and ssDNA signals.

Mass spectrometry analysis

Protein bands corresponding to the 72 kDa (full-length) and 23 kDa (truncated) variants of SERP2475 in the FT fraction from the second step of purification were excised from a 12% SDS-PAGE gel stained with Coomassie G-250. Mass spectrometry was performed by the Cancer Center Mass Spectrometry and Proteomics Shared Facility at the University of Alabama, Birmingham. The bands were digested overnight with Trypsin Gold, Mass Spectrometry Grade (Promega, cat. #V5280) following the manufacturer's instructions. Peptide extracts were reconstituted in 0.1% formic acid/ddH₂O at 0.1 $\mu\text{g}/\mu\text{l}$. Electrospray ionization tandem mass spectrometry was carried out, and the data were processed, filtered, grouped, and quantified, as previously reported in detail (Ludwig et al., 2016). The data were searched against a tailored database comprising of the *E. coli* proteome plus the protein sequence of interest (SERP2475).

Phage hybrid generation and sequencing

JBug18-Andhra Hybrids 1-8 were isolated as immune resistant mutants following challenge of LM1680/pSERP-2475 with a high titer lysate of JBug18 ($\sim 1 \times 10^{10}$ pfu/ml). To generate JBug18-Andhra Hybrids 9-18, overnight cultures of *S. epidermidis* LM1680 harboring pT181-*gp03* or pT181-*gp0304* were diluted 1:100 in fresh TSB supplemented with antibiotics and 5 mM CaCl_2 . The mixture was incubated at 37°C for an hour with agitation, then JBug18 was added to the cells in a 1:1 ratio, and the incubation continued with agitation overnight. The next day, cells were pelleted by centrifugation at 8000 x *g* for 5 min and supernatant was filtered through 0.45 mm filter. Filtered lysates were mixed with LM1680-pSERP2475 overnight culture (1:1) and the mixture was plated on TSA containing 5 mM CaCl_2 using the double-agar overlay method (Cater et al., 2017). For all phage hybrids, individual plaques were isolated and re-plated three times on LM1680/pSERP-2475 to purify. Phages were propagated and their DNA was extracted as previously described (Bari et al., 2017). Phage genomes were PCR amplified across the entire coding region for Hybrids 1-8 or *gp03-gp04* for Hybrids 9-18, and the PCR products were sequenced by the Sanger method (at Eurofins MWG Operon) using the primers listed in Table S8.

Hybrid phage genome sequence analysis

For JBug18-Andhra Hybrids 1-8, Sanger sequencing reads covering their coding regions were manually assembled using SnapGene software. For JBug18-Andhra Hybrids 9-18, a single read covered the region of interest, therefore no assembly was required. Sequences for each set of hybrids (1-8 and 9-18) were aligned with corresponding genomic regions in Andhra and JBug18 using the Clustal Omega Multiple Sequence Alignment tool (McWilliam et al., 2013). The sequence alignments (Data S2 and S3) were analyzed by a Python script developed in-house which first scans the alignment of JBug18 and Andhra, identifies each position of non-similarity, and then determines at those positions the fraction of hybrids that possess Andhra identity. The output data was exported into an Excel file, and the graphs showing the fraction of hybrids with Andhra identity at each position were generated using Microsoft Excel.

Isolation and amplification of ISP escaper phages

ISP escaper phages were isolated by plating dilutions of a concentrated wild-type phage lysate (1×10^{10} pfu/ml) atop a lawn of *S. epidermidis* LM1680 cells bearing pSERP-2475 and incubating plates overnight at 37°C. The following day, two isolated plaques were observed—these were picked using a sterile pipette tip, resuspended in 500 μl fresh TSB, and 10-fold dilutions were plated again as above. The procedure was repeated twice more to purify the phage escapers. Concentrated phage stocks were prepared from purified plaques as described in the section above (“phage propagation and enumeration”).

Genomic DNA extraction of ISP escaper phages

DNA was extracted from high titer phage lysates ($\geq 1 \times 10^9$ pfu/ml) as previously described (Cater et al., 2017). Briefly, 20 ml of phage lysate was digested with DNase I and RNase A (10 $\mu\text{g/ml}$ of each) for 30 min at 37°C. Digested lysate was combined with ten milliliters of precipitant solution (30% [wt/vol] polyethylene glycol [PEG] 8000 and 3 M NaCl) and incubated at 4°C overnight. Phages were pelleted by centrifugation for 10 min at 10,000 x *g* and 4°C. The phage pellet was resuspended in 250 μl of resuspension buffer (5 mM MgSO_4 , 10 mM EDTA (pH 8.0)), and the suspension was incubated with proteinase K (100 $\mu\text{g/ml}$) at 50°C for 30 min. The phage suspension was then combined with the resin contained in the Promega Wizard DNA cleanup kit (catalog no. A7280), and the mixture was inverted several times and applied to a minicolumn contained within the kit. The resin was washed with 2 ml of 80% isopropanol and dried by centrifugation at 13,000 x *g* for 2 min, and DNA was eluted from the resin with 100 μl of distilled water preheated to 80°C.

Escaper phage DNA sequencing and analysis

Library preparation and sequencing was performed at the Microbial Genome Sequencing Center (Pittsburgh, PA) using the Illumina Library Prep Tagmentation kit. Sequencing was performed on an Illumina NextSeq 2000. Adapters and indexes were removed by bcl2fastq v. 1.8.4 (Illumina) and FastQC v. 0.11.9 was used to confirm data quality (e.g. number of bases with quality above Q30). Wild-Type ISP sequencing reads were assembled using SPAdes v. 3.15.3 (Bankevich et al., 2012) in isolate mode with the kmer values of 21, 33, 55, 77 and 99. The resulting assembly graph was inspected using Bandage v. 0.8.1 (Wick et al., 2015) and the high coverage contig with the proper length representing the phage genome was extracted as a fasta file. The fasta file was then indexed using bowtie2 v. 2.4.4 (Langmead and Salzberg, 2012), and the escaper reads were recruited to the indexed wild-type genome sequence using bowtie2 with default parameters for alignment and scoring. Resulting bowtie2 output file (Sequence Alignment/Map - sam format) was converted to binary bam format, sorted using samtools v. 1.13 (Li et al., 2009) and indexed for coverage depth analysis. Coverage depth analysis was performed using igvtools v. 2.11.1 (Robinson et al., 2011) and the resulting wig file was analyzed to calculate and plot fraction of reads matching the wild-type ISP nucleotide at every genome position as well as the corresponding depth of coverage in reads per million.

QUANTIFICATION AND STATISTICAL ANALYSIS

Student's *t*-tests were performed to determine if observed drops in plaque counts (Figure 2) or enzymatic activities (Figure 5) were statistically significant. Tests were performed using Microsoft Excel, and a difference was deemed significant if the *p*-value was below 0.05. Details for specific experiments can be found in the corresponding figure legends.

Supplemental information

**A unique mode of nucleic acid immunity
performed by a multifunctional bacterial enzyme**

S.M. Nayeemul Bari, Lucy Chou-Zheng, Olivia Howell, Motaher Hossain, Courtney M. Hill, Tori A. Boyle, Katie Cater, Vidya Sree Dandu, Alexander Thomas, Barbaros Aslan, and Asma Hatoum-Aslan

Supplementary Figures

A unique mode of nucleic acid immunity performed by a multifunctional bacterial enzyme

S. M. Nayeemul Bari^{1,3}, Lucy Chou-Zheng^{1,3}, Olivia Howell¹, Motaher Hossain¹, Courtney M. Hill¹, Tori A. Boyle¹, Katie Cater², Vidya Sree Dandu², Alexander Thomas², Barbaros Aslan¹, and Asma Hatoum-Aslan^{1,4}

¹ Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, IL 61821, USA

² Department of Biological Sciences, University of Alabama, Tuscaloosa, AL 35487, USA

³ These authors contributed equally

⁴ Corresponding Author and Lead Contact: ahatoum@illinois.edu

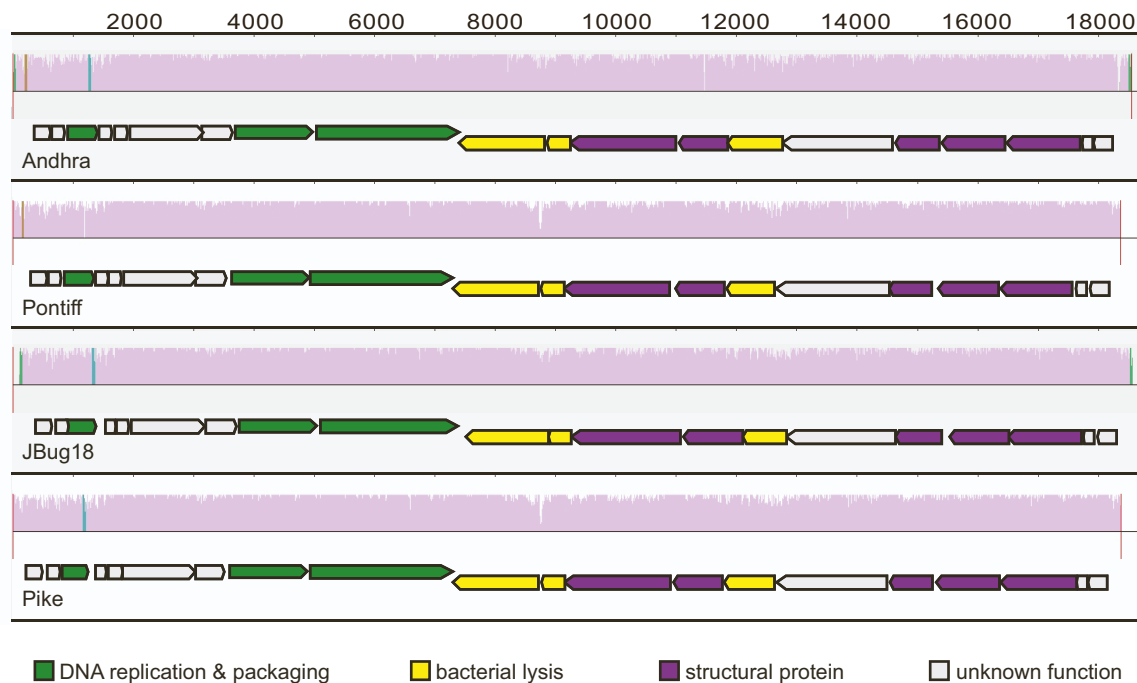


Figure S1. Four related *Podoviridae* phages with different host ranges, Related to Figure 1. Shown is a multiple genome alignment of *S. epidermidis* podophages Andhra, Pontiff, JBug18, and Pike. Genome coordinates are shown on top, and colored histograms indicate the nucleotide similarity at each position derived from a multiple sequence alignment. The open reading frames for each phage are shown underneath the corresponding histogram. The histograms were generated using the MAUVE open source software (<http://darlinglab.org/mauve/mauve.html>) and the outlines of open reading frames from the MAUVE output were overlaid using Adobe Illustrator.

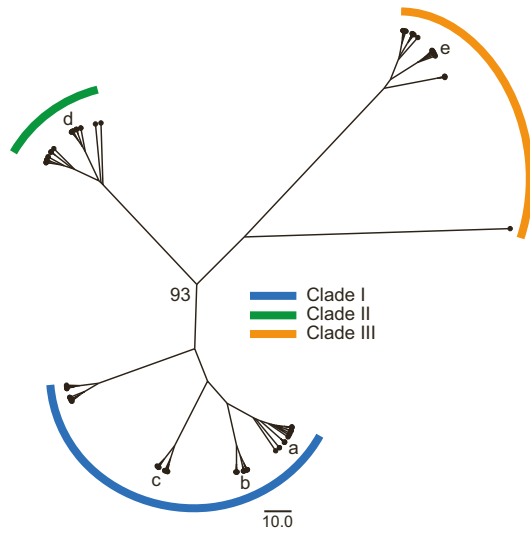


Figure S2. A dendrogram generated for SERP2475 and 99 homologs, Related to Figure 2.

Positions of homologs selected for functional characterization are indicated with lower-case letters: a. WP_045177897 from *S. aureus* MJ163; b. WP_002489608 (i.e. SERP2475) from *S. epidermidis* RP62a; c. WP_000632676 from *S. aureus* CA-347; d. WP_013870910 from *Lacinutrix* 5H-3-7-4; e. WP_101958732 from *V. vulnificus* FORC54.

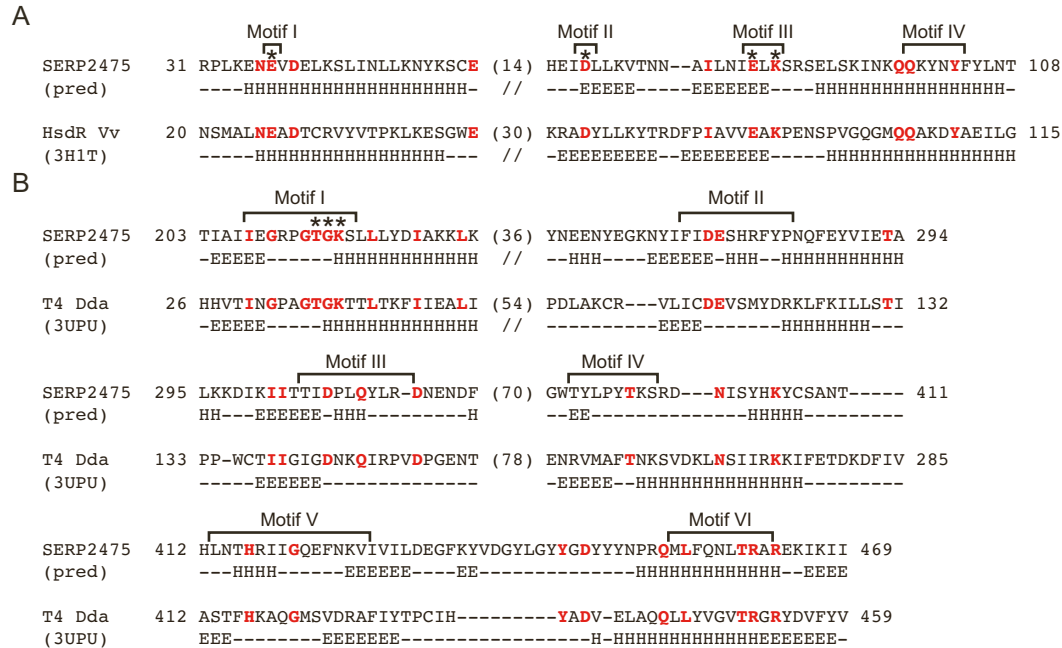


Figure S3. Predicted structural homologs of SERP2475, Related to Figure 4. Pairwise sequence alignments between SERP2475 and relevant regions of the putative HsdR endonuclease from *Vibrio vulnificus* (A) and the Dda helicase from phage T4 (Swiss PDB IDs 3H1T and 3UPU) (B) are shown as determined by HHPred. Residues colored red indicate positions of sequence identity and asterisks mark residues that were subjected to mutational analysis in this study. Predicted (pred) and actual alpha helices (H) and beta sheets (E) are indicated below each position in the alignments.

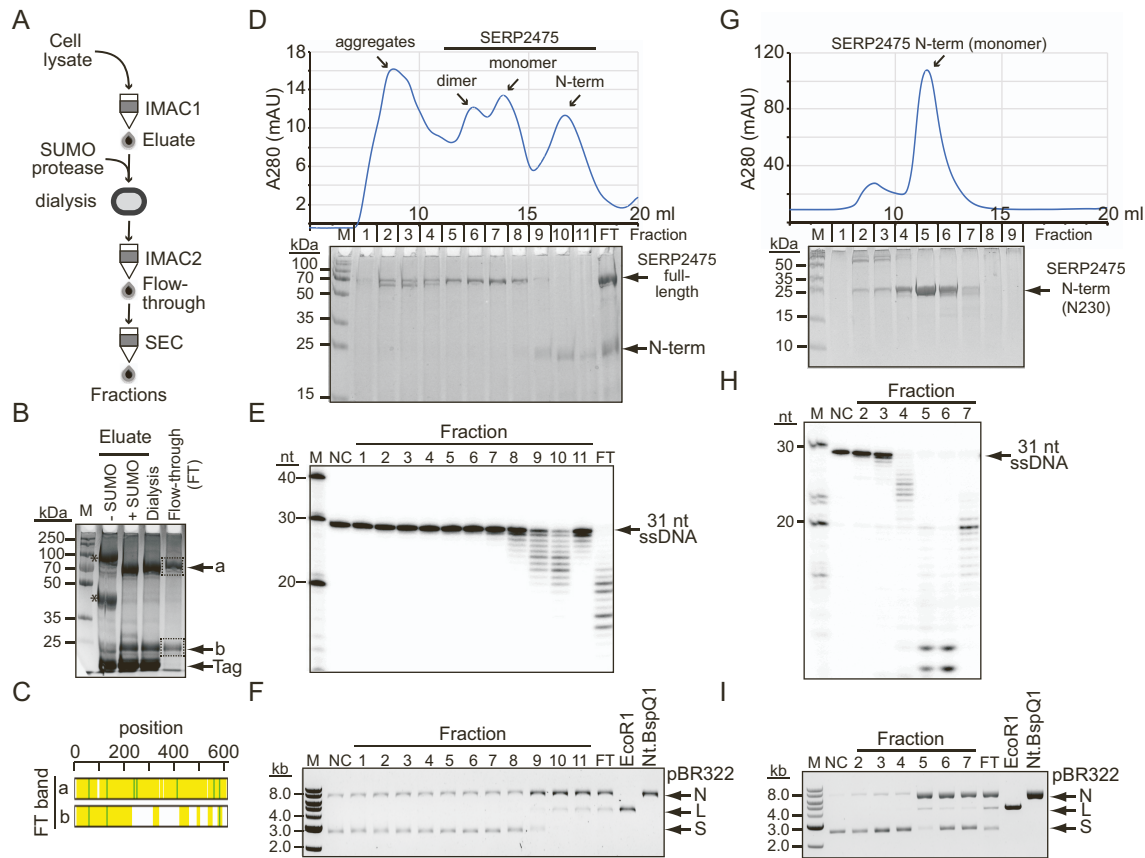


Figure S4. Purification of SERP2475 and initial characterization of its nuclease activities, Related to Figure 5. (A) Three-step protein purification process used in this study. IMAC, immobilized metal affinity chromatography; SEC, size exclusion chromatography. (B) Image of representative SDS-PAGE gel showing protein species present throughout the first two chromatography steps. Asterisks mark tagged versions of the protein before SUMO protease digestion of the 10His-Smt3 tag. Dashed boxes encompass protein bands that were excised and subjected to mass spectrometry analysis. (C) Peptide coverage of SERP2475 in excised bands. Yellow, identified peptides; green, cysteine carbamidomethylation or methionine oxidation. (D and G) Size exclusion chromatograms (top) and SDS-PAGE gels (bottom) resolving fractionated proteins in full-length (D) or the N-terminal 230 AAs (G) of SERP2475. mAU, milli absorbance units at 280 nm. (E and F) Exonuclease (E) and nickase (F) assays in

77 which indicated substrates were combined with fractions from the full-length protein prep in
78 panel D. (H and I) Exonuclease (H) and nickase (I) assays in which indicated substrates were
79 combined with fractions from the N-terminal 230 AA prep in panel G. Products of exonuclease
80 and nickase reactions were resolved on denaturing urea-PAGE and native agarose gels,
81 respectively. For nickase assays, EcoRI and Nt.BspQI were used as controls to generate linear
82 (L) and nicked (N) products, respectively, from the supercoiled (S) plasmid. Shown are
83 representatives of at least three independent trials. See also Figure S5 and Tables S5-S7.

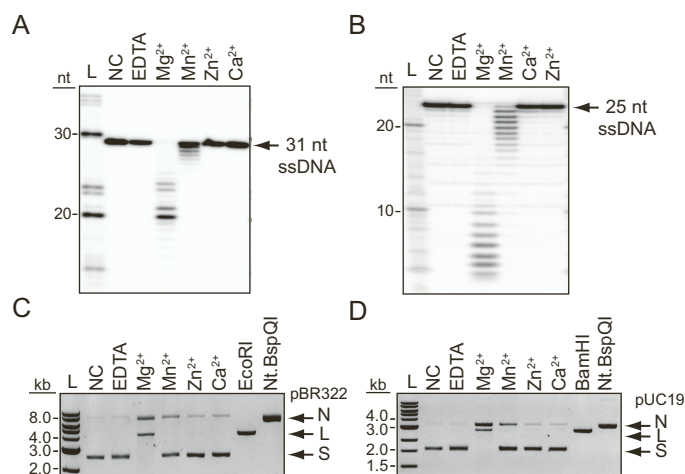


Figure S5. Activity of SERP2475 N-terminal 230 amino acids in the presence of different metals and substrates, Related to Figure 5. (A and B) Exonuclease assays are shown in which linear substrates were combined with a purified preparation of the N-terminal 230 AAs of SERP2475 (13 nM) in a reaction supplemented with indicated metals or EDTA (2 mM) and incubated at 37°C for 20 min. (C and D) Nickase assays are shown in which supercoiled plasmids were combined with a purified preparation of the N-terminal 230 AAs of SERP2475 (13 nM) in a reaction supplemented with indicated metals or EDTA (10 mM) and incubated at 37°C for 1 h. NC, negative control (no protein); N, nicked; L, linear; S, supercoiled. Shown are representative images for three independent trials. See Table S7 for linear substrates used.

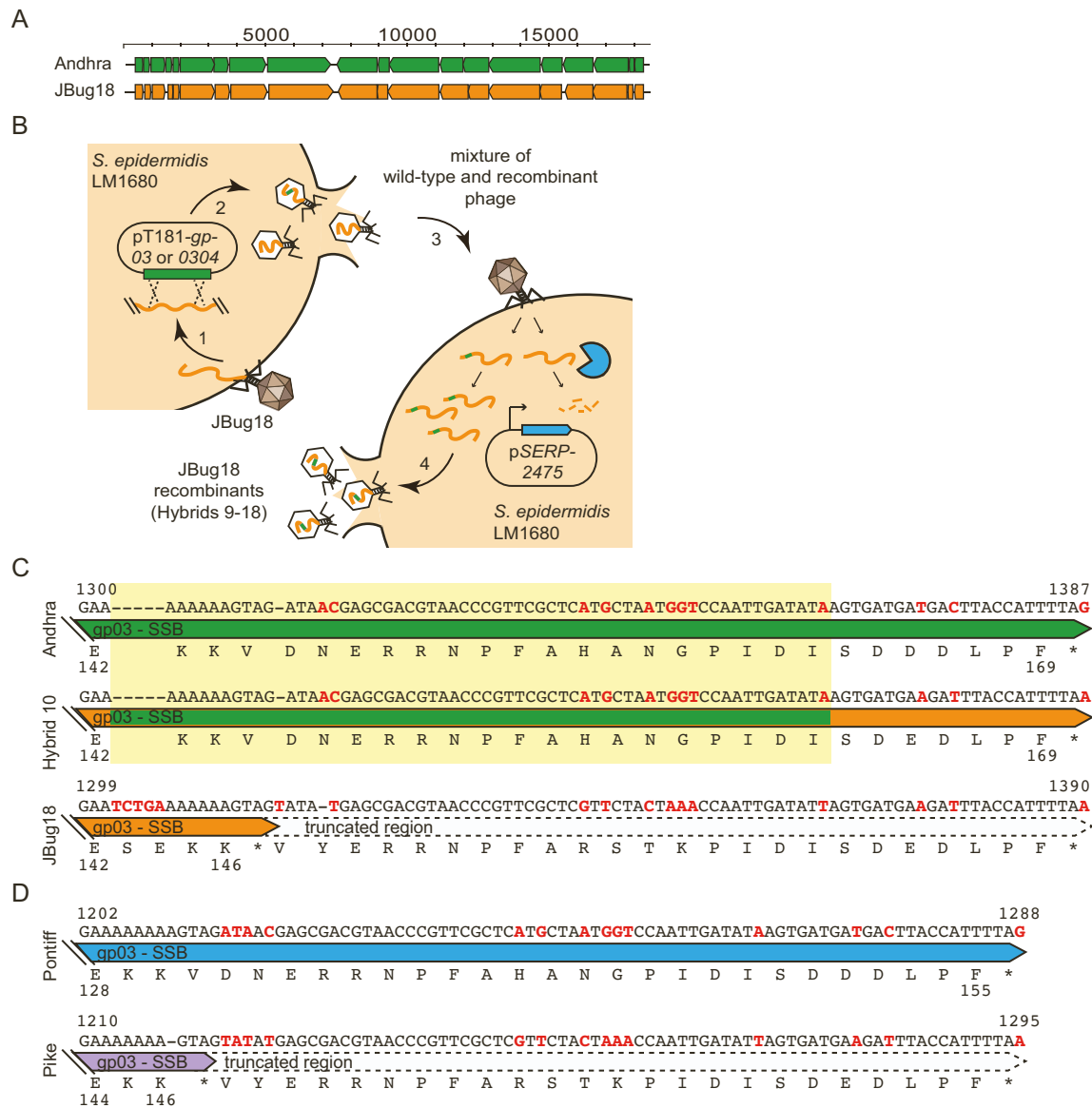


Figure S6. Generation of Hybrids 9-18 and comparison of SSB sequences, Related to Figure 6. (A) A pairwise comparison of the open reading frames of Andhra and JBug18. (B) A diagram of the method used to generate JBug18-Andhra Hybrids 9-18. (C) Sequence comparison between the SSBs of Andhra, JBug18, and Hybrid 10, which gained resistance to immunity through the acquisition of only 60 nucleotides of Andhra-derived sequence (highlighted in yellow). (D) a similar comparison between the SSBs of phages Pontiff and Pike.