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

S. M. Nayeemul Bari[1§], Lucy Chou-Zheng[1§], Olivia Howell[1], Katie Cater[2], Vidya Sree Dandu[2], Alexander Thomas[2], Barbaros Aslan[1], and Asma Hatoum-Aslan[1*]

[1] Department of Microbiology, University of Illinois, Urbana, IL 61821, USA

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

[§] These authors contributed equally

*Correspondence: ahatoum@illinois.edu

Abstract

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 anti-phage defense. However, the prevalence and effectiveness of much simpler, single-component systems that cleave DNA remain unknown. Here, we describe a novel mode of nucleic acid immunity performed by a single enzyme with nuclease and helicase activities, herein referred to as Nhi. 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 acts as both the sensor and effector to degrade 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 anti-phage 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 novel bacterial immune systems (Bernheim et al., 2021; 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 different systems that sense and destroy phage genetic material. Such systems exhibit a range of complexities, from the simpler restriction-modification (RM) systems 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 in the phage, and a DNA methyltransferase that modifies and protects the host genome from cleavage. Further, more complex RM-like systems have recently been described, including Pgl (phage growth limitation)

53 (Hoskisson et al., 2015; Sumby and Smith, 2002), DISARM (defense island associated with
54 restriction-modification) (Ofir et al., 2017), and BREX (bacteriophage exclusion) (Goldfarb et al.,
55 2015; Gordeeva et al., 2019). All of these encode DNA methyltransferases in conjunction with a
56 variety of genes such as proteases, phosphatases, and phospholipases which provide other
57 necessary functionalities for defense. CRISPR-Cas systems are even more elaborate—they
58 retain a molecular memory of phage-derived nucleic acids by capturing small stretches of their
59 genetic material and storing them in the CRISPR locus as “spacers” in between DNA repeats
60 (Hille et al., 2018). The spacer sequences are transcribed to generate small CRISPR RNAs
61 (crRNAs), which are subsequently used in combination with one or more Cas nucleases to
62 identify and eliminate phages harboring sequences complementary to the crRNAs. These
63 systems adapt to new and evolving threats by continually capturing new spacers from new
64 phage invaders. There are two classes, six types and 33 subtypes of CRISPR-Cas systems
65 currently described (Makarova et al., 2020), and different types have been shown to work
66 together (Deng et al., 2013; Hoikkala et al., 2021; Pinilla-Redondo et al., 2020; Silas et al.,
67 2017) and even synergize with RM systems (Dupuis et al., 2013) to ensure a more effective
68 defense. Such added layers have obvious advantages in protecting against diverse and
69 evolving phage predators; however, this strategy comes with the energetic cost of producing
70 and maintaining multiple components, as well as the risk that damage or loss of a single part
71 may render the system inactive. Thus, it is reasonable to speculate that bacteria also employ
72 much simpler, single-component systems to target and degrade phage nucleic acids. However,
73 the prevalence and effectiveness of such minimal systems remain poorly understood.
74 Here, we describe the discovery and characterization of a unique mode of nucleic acid
75 immunity performed by a single enzyme, herein referred to as Nhi. This enzyme provides full
76 protection against diverse staphylococcal phages, and it is sufficient to prevent phage DNA
77 accumulation in a staphylococcal strain devoid of all other known defenses. In a purified system,
78 Nhi dimerizes and possesses both nuclease and helicase activities. Further, we found that a

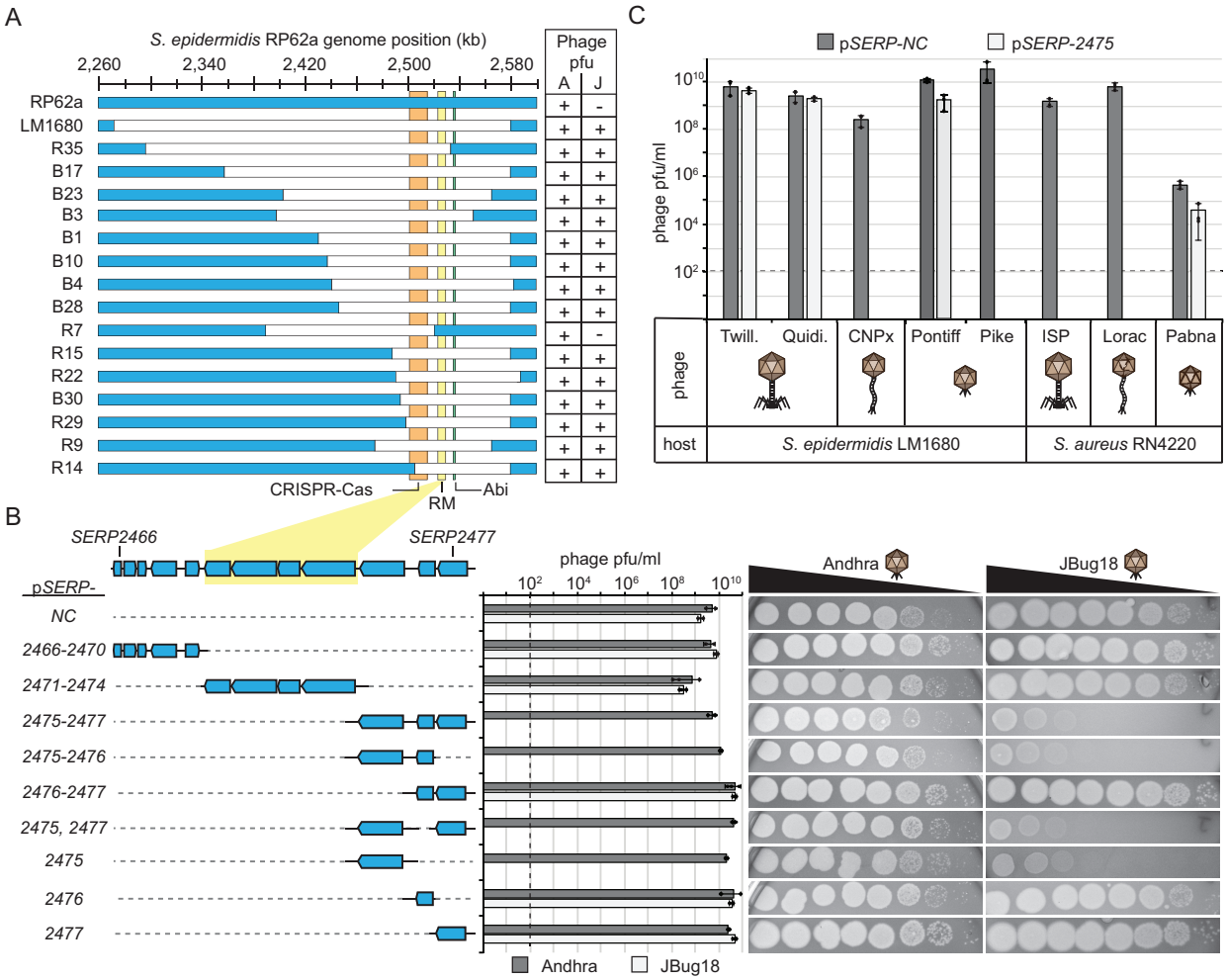
79 single-stranded DNA binding protein encoded in one group of phages can protect against Nhi's
80 effects. Based on these observations, we propose a model in which Nhi senses and degrades
81 phage-specific replication intermediates. Importantly, Nhi homologs can be found in diverse
82 bacterial phyla, and we provide evidence that they are also dedicated to immunity. Our findings
83 highlight the versatility of such compact systems as powerful weapons in anti-phage defense.
84
85 **Results**
86
87 ***SERP2475* protects against diverse staphylococcal phages**
88 We undertook this study with an aim to uncover new mechanisms of immunity in the
89 commensal opportunistic pathogen *Staphylococcus epidermidis* RP62a (Christensen et al.,
90 1987). This organism harbors a Type III-A CRISPR-Cas system (Marraffini and Sontheimer,
91 2008), an Abi mechanism (Depardieu et al., 2016), and a putative Type I RM system, all of
92 which are encoded within ~30,000 nucleotides of each other (Fig. 1A). Their close proximity is
93 consistent with recent reports that show prokaryotic immune systems typically cluster together
94 within discrete genomic loci known as defense islands (Doron et al., 2018; Gao et al., 2020;
95 Makarova et al., 2011). Importantly, key insights into CRISPR-Cas and Abi in this organism
96 were revealed by studying their molecular interactions with temperate phages ΦNM1 and CNPx,
97 respectively, closely related phages that belong to the family *Siphoviridae* (Depardieu et al.,
98 2016; Goldberg et al., 2014). Indeed, siphophages are the most common members in
99 staphylococcal phage collections (Oliveira et al., 2019), and we reasoned that the identification
100 of new immunity mechanisms might necessarily require the examination of more diverse
101 members. Toward that end, we captured and characterized four new lytic *S. epidermidis* phages
102 belonging to the family *Podoviridae*: Andhra, JBug18, Pontiff, and Pike (Cater et al., 2017;
103 Culbertson et al., 2019). These phages share over 95% sequence identity and the same 20
104 genes (Fig S1); however, we noticed that they have distinct host ranges—while Andhra and

105 Pontiff can infect the wild-type RP62a strain, JBug18 and Pike can only infect a mutant variant,
106 LM1680 (Jiang et al., 2013), which has a large (~300k nucleotide) deletion encompassing the
107 defense island and neighboring regions (Fig. 1A). These observations led to the hypothesis that
108 JBug18 and Pike are sensitive to genetic element(s) within the defense island.

109 To test this, we used a set of *S. epidermidis* RP62a mutants that bear deletions of varying
110 extents across the defense island (Jiang et al., 2013) (Fig. 1A). These strains were challenged
111 with Andhra and JBug18, representatives from the set of four with resistant and sensitive
112 phenotypes, respectively. The resulting zones of bacterial growth inhibition (plaques) were
113 enumerated, revealing that only one of the deletion mutants (R7) encodes the gene(s) required
114 for complete protection against JBug18. These observations narrowed the protective genetic
115 element(s) to a stretch of ~12,000 nucleotides containing 12 genes (designated as *SERP2466-*
116 *SERP2477*) which incidentally encompasses the RM system (Fig. 1B). To determine which of the
117 12 gene(s) are responsible for immunity, they were inserted (individually and in groups) into a
118 derivative of plasmid pC194 (Ehrlich, 1977) (herein referred to as p*SERP-*), introduced into *S.*
119 *epidermidis* LM1680, and the resulting strains were challenged with Andhra and JBug18. Through
120 this analysis, we found that a single gene of unknown function, *SERP2475* (new locus tag
121 *SERP_RS12125*) is sufficient to protect against JBug18 (Fig. 1B). A repeat of this assay with
122 phages Pontiff and Pike showed results similar to those observed with Andhra and JBug18,
123 respectively—while *SERP2475* has no effect on Pontiff, it completely protects against Pike (Fig.
124 1C).

125 To further understand the breadth of protection this gene affords, we first challenged the
126 *S. epidermidis* LM1680/p*SERP-2475* strain with additional phages from our collection from
127 different morphological families—*Herelleviridae* (Barylski et al., 2020) (formerly *Myoviridae*) and
128 *Siphoviridae*. We observed that although *SERP2475* has no noticeable impact on the lytic
129 myophages Twillingate and Quidividi (Freeman et al., 2019), it fully protects against siphophage
130 CNPx (Fig. 1C). We also tested the effectiveness of SERP2475 against *S. aureus* phages

131 representing all three families of tailed phages and found that while SERP2475 causes only
132 modest reduction in plaques of podophage Pabna (Culbertson et al., 2019), it affords full
133 protection against myophage ISP (Vandersteegen et al., 2011) and siphophage Lorac (Antoine
134 Marc et al., 2019). Taken together, these data demonstrate that *SERP2475* is sufficient to provide
135 robust protection against diverse staphylococcal phages.
136
137



138
139
140
141

142 **Figure 1.** *SERP2475* provides robust immunity against diverse staphylococcal phages. (**A**) A
143 segment of the S*. epidermidis* RP62a genome and deletion mutants. Regions encoding CRISPR-Cas, RM, and Abi systems are highlighted. Strains were challenged with Andhra (A)
144 and JBug18 (J), and resulting plaque-forming units per milliliter (pfu/ml) are indicated: +, ~1 x 10^9 pfu/ml; -, 0 pfu/ml. (**B**) Magnified view of the genomic region responsible for immunity and
145 corresponding plasmids that were created to encompass them. *S. epidermidis* LM1680 strains harboring indicated plasmids were challenged with phage, and the resulting pfu/ml are shown
146 as an average of triplicate measurements (± S.D.). Images on the right show representative plates following the application of ten-fold dilutions of Andhra and JBug18 ($1x10^0$ – $1x10^{-7}$) atop
147 *S. epidermidis* LM1680 strains bearing indicated plasmids. (**C**) *S. epidermidis* and *S. aureus* strains were challenged with phages representing all three morphological families of tailed
148 phages, and resulting pfu/ml are shown. The dotted line indicates the limit of detection for phage challenge assays. See also Figure S1.

149

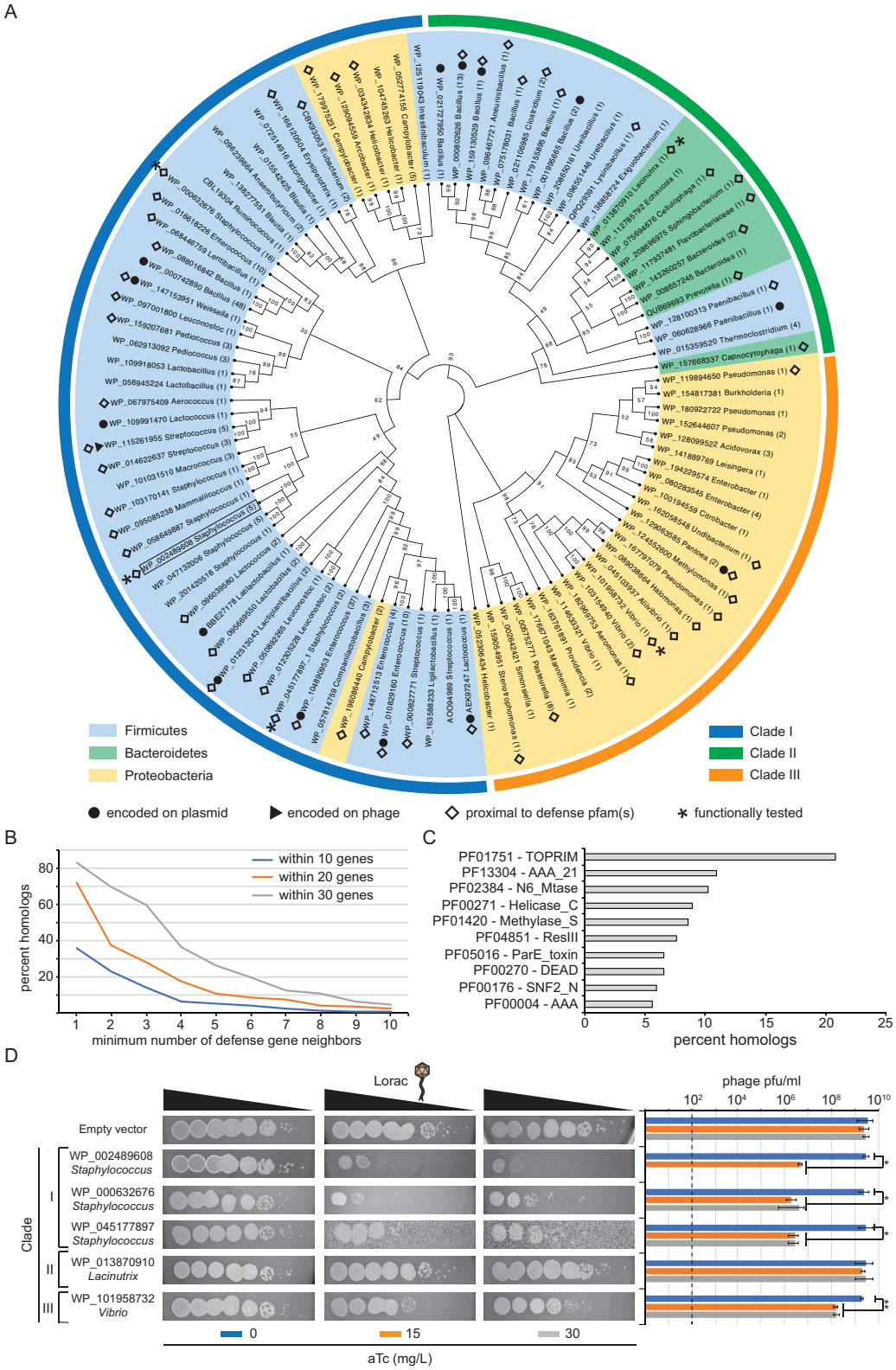
150 **SERP2475 homologs are distributed in diverse bacteria and exhibit anti-phage activity**

151 We next assessed the distribution of SERP2475 homologs and the extent to which they
152 are functionally conserved. We used tBLASTn to query three NCBI databases
153 (prok_complete_genomes, refseq_genomes, and nt) and identified 302 homologs in distinct
154 genetic backgrounds (Table S1). Although homologs were present in <1% of the fully assembled
155 microbial genomes surveyed, representatives could be found in three bacterial phyla: *Firmicutes*,
156 *Bacteroidetes*, and *Proteobacteria*. Additionally, two of the homologs were found in two
157 *Streptococcus* phages. To better understand their relationships to one another, we generated a
158 phylogenetic tree from 100 selected representatives that encompass the phylogenetic diversity of
159 the group (Fig. 2A and S2, and Supplementary File 1). This analysis revealed that the homologs
160 cluster into three distinct clades that are somewhat incongruent with host phylogeny--While clade
161 III contains homologs strictly found in *Proteobacteria*, clades I and II contain homologs originating
162 from two of the three phyla. This observation suggests that *SERP2475* and its homologs are likely
163 disseminated through horizontal gene transfer. Supporting this, many of the homologs are
164 encoded on plasmids (Fig. 2A and Table S1). To determine their level of functional conservation,
165 we first checked for proximity to genes with known defense functions. Amino acid sequences of
166 the thirty proteins encoded upstream and downstream of each homolog were extracted and
167 searched for identifiable protein domains using the hmmer software v.3.3.2 (hmmer.org). The

168 predicted protein families (pfams) of these flanking proteins were then searched against 306
169 pfams with known functions in anti-phage defense (Table S2). Figure 2B shows the fraction of
170 homologs that have between one and ten defense-related neighbors encoded within expanding
171 windows of 10, 20, and 30 genes. This data revealed that a significant fraction of SERP2475
172 homologs are indeed encoded proximal to other known defenses. For instance, 72% of homologs
173 have at least one defense neighbor encoded within 20 genes (Fig. 2B and Table S3), thus
174 supporting roles in anti-phage defense. The most frequently encountered defense neighbors
175 within 20 genes include proteins with TOPRIM (PF01751), ATPase (PF13304, PF00176, and
176 PF00004) and methyltransferase (PF02384 and PF01420) domains (Fig. 2C and Table S4).

177 To further explore the functional conservation of these homologs, we selected four
178 representatives from the three clades and tested for anti-phage activity (Figs. 2A and S2, and
179 Table S1). Importantly, these homologs share minimal (32-36%) amino acid sequence similarity
180 when compared with SERP2475. The coding sequences for SERP2475 and selected homologs
181 were inserted into a plasmid (herein referred to as pTET-) downstream of an anhydrotetracycline
182 (aTc) inducible promoter and then introduced into *S. aureus* RN4220. The resulting strains were
183 challenged with *Siphoviridae* phage Lorac in the presence and absence of the inducer. As
184 expected, SERP2475 (RefSeq ID WP_002489608) affords full protection against the phage when
185 the cells were grown in the presence of the maximum concentration of aTc (Fig. 2D). Although
186 expression of the Clade II homolog (WP_013870910 from phylum *Bacteroidetes*) resulted in no
187 detectable decrease in plaque size or number, the two *Staphylococcus* homologs from Clade I
188 (WP_000632676 and WP_045177897) and the *Vibrio* homolog from Clade III (WP_101958732
189 from phylum *Proteobacteria*) caused significant reduction in plaque numbers (1000- to 10-fold,
190 respectively). A noticeable decrease in plaque size was also observed in the latter. Interestingly,
191 one of the staphylococcal homologs (WP_045177897) also appeared to be toxic to the cells, as
192 evidenced by the 'unhealthy' appearance of the lawn when cells harboring the homolog are grown

in the presence of inducer (Fig. 2D). Altogether, these data provide evidence that distant homologs of SERP2475 are also dedicated to anti-phage defense.



196 **Figure 2.** SERP2475 homologs exhibit functional conservation. (**A**) A cladogram showing the
197 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,
198 and number of distinct species/strains within the same genus that harbor an identical or closely-
related homolog (>90% amino acid identity). Inner ring highlighting indicates the phylum, and the
199 outer ring coloration indicates the clade (I-III) in which each homolog belongs. Representatives
encoded on plasmids (closed circles), phages (closed triangle), and within 20 genes of at least one
200 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
201 with a box. (**B**) A neighborhood analysis showing the fraction of homologs (including SERP2475,
n=303) that have between 1 and 10 neighbors dedicated to anti-phage defense within 10, 20, and
202 30 flanking genes. (**C**) Top ten defense-related protein families (pfams) encoded within 20 genes
of SERP2475 homologs. (**D**) Functional analysis in which the coding sequences of selected
203 homologs were inserted into a plasmid downstream of an anhydrotetracycline (aTc) inducible
promoter and subsequently introduced into *S. aureus* RN4220. Images show ten-fold dilutions of
204 phage 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 (± S.D) are shown as a
205 representative of three independent trials. Asterisks indicate p-values <0.05 (*) and <0.0005 (**) in
a two-tailed t-test. The dotted line shows the limit of detection for this assay. See also Figure S2.

206

207 ***SERP2475* limits phage DNA accumulation**

208 We next sought to investigate the mechanism of immunity, and began by assessing which
209 stage of the phage infection cycle SERP2475 targets. As previously mentioned, common
210 strategies that bacteria employ in anti-phage defense include masking/modifying the cell surface
211 to prevent phage attachment, targeting and degrading phage-derived nucleic acids, and causing
212 programmed cell death through various Abi mechanisms (Fig. 3A). To test if *SERP2475* hinders
213 phage attachment, we conducted an adsorption assay in which *S. epidermidis* LM1680 cells
214 bearing p*SERP-2475* were combined with a defined number of phage particles for ten minutes—
215 just enough time for phages to attach to cells, but not long enough for these phages to complete
216 their replication cycle (Cater et al., 2017). The mixture was then pelleted to remove both the cells
217 and phages attached to them, and the phages remaining in suspension were enumerated to
218 determine the number that adsorbed (i.e. attached) to cells. This assay revealed that JBug18
219 attaches to LM1680/p*SERP-2475* just as efficiently as Andhra (Fig. 3B), thus ruling out an
220 adsorption-blocking mechanism.

221 A cell viability assay was next
222 performed to test for Abi. The prediction is
223 that if programmed cell death accompanies
224 immunity, then challenge with a high
225 proportion of phages to cells (≥1:1) would
226 lead to significant decline in cell viability
227 similar to that observed in the absence of
228 immune protection (Goldfarb et al., 2015;
229 Ofir et al., 2017). To test this, phages were
230 combined with LM1680/p*SERP-2475* cells in
231 liquid media at ratios of 1:1, 5:1, or 10:1, and
232 cell viability was measured following 5 hours
233 of incubation. We found that while Andhra
234 causes significant death of LM1680/p*SERP-*
235 *2475*, JBug18 elicits only a minor decrease
236 in the viability of this strain, even when
237 phages outnumber bacteria 10:1 (Fig. 3C).
238 These observations suggest that cell suicide
239 is unlikely to be the mechanism by which
240 *SERP2475* affords protection.

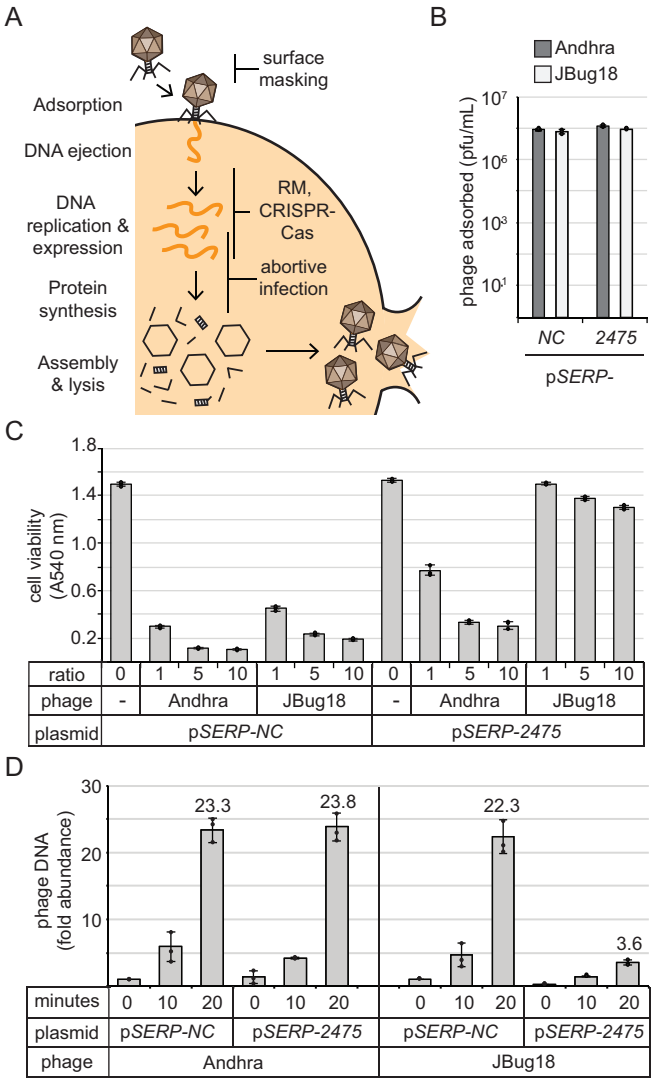


Figure 3. SERP2475 impairs phage DNA accumulation. (**A**) Illustration of the lytic phage replication cycle and common defense systems that interfere with each step. (**B** and **C**) Results of an adsorption assay (**B**) and a cell viability assay (**C**) following challenge of *S. epidermidis* LM1680 cells bearing indicated plasmids with Andhra and JBug18. (**D**) Relative abundance of phage DNA at various time points following phage infection as measured by qPCR. For all experiments, the mean ± S.D. of triplicate measurements are shown as a representative of at least two independent trials.

241 Finally, we tested whether *SERP2475*
242 impacts phage DNA levels in the cell.
243 Quantitative PCR (qPCR) was used to track
244 the accumulation of phage DNA at various
245 time points following infection with both phages. The results showed that while Andhra's DNA
246 accumulates to ~20-fold by 20 minutes post-adsorption in LM1680/p*SERP-2475*, JBug18's DNA

247 accumulates to less than 4-fold in the same time period (Fig. 3D). These observations support a
248 hypothesis whereby SERP2475 protects against phages by directly interfering with phage DNA
249 replication and/or expression.
250
251 **SERP2475 relies upon nuclease and helicase activities to perform immunity**
252 To begin to understand the catalytic function of this protein, we first conducted *in silico*
253 analyses. *SERP2475* encodes a 606 amino acid protein, and according to Interproscan (Blum et
254 al., 2021), a web-based tool that predicts protein domains and other important functional sites,
255 the most prominent features of SERP2475 are a central domain of unknown function (DUF) 2075
256 (protein family PF09848) and a P-loop NTPase domain that overlaps with it (Fig. 4A). Additionally,
257 the structural homology search tools HHPred (Zimmermann et al., 2018) and Phyre2 (Kelley et
258 al., 2015) identified a handful of superfamily 1 helicases as close homologs, including T4 phage
259 Dda helicase (He et al., 2012) (E-value 1.2 x 10^{-21}). The latter tools also predicted structural
260 similarities between the N-terminus of SERP2475 and a putative HsdR restriction endonuclease
261 from *Vibrio vulnificus* (Uyen et al., 2009) (E-value 1.1 x 10^{-6}). Although SERP2475 shares minimal
262 amino acid sequence similarity in pairwise comparisons with full length HsdR_Vv and Dda (19%
263 and 22%, respectively), we located conserved residues in key motifs corresponding to a putative
264 PD-(D/E)XK nuclease domain on its N-terminus (Fig. 4B) and putative helicase domain spanning
265 the central portion of the protein (Fig. 4C). Further, we observed a striking conservation of these
266 critical residues in the sequence alignment with SERP2475 and the 99 representative homologs
267 (Fig. 4D), suggesting that these domains are likely important for protein function. To confirm, we
268 constructed two mutant versions of p*SERP-2475* that encode alanine substitutions in the putative
269 nuclease and helicase domains (asterisks in Fig. 4B and C). The strains were then challenged
270 with phages Andhra and JBug18, and the results showed a complete restoration of JBug18
271 replication in both mutant strains (Fig. 3E). These collective observations support the hypothesis
272 that SERP2475 and its homologs use nuclease and helicase activities to achieve immunity.

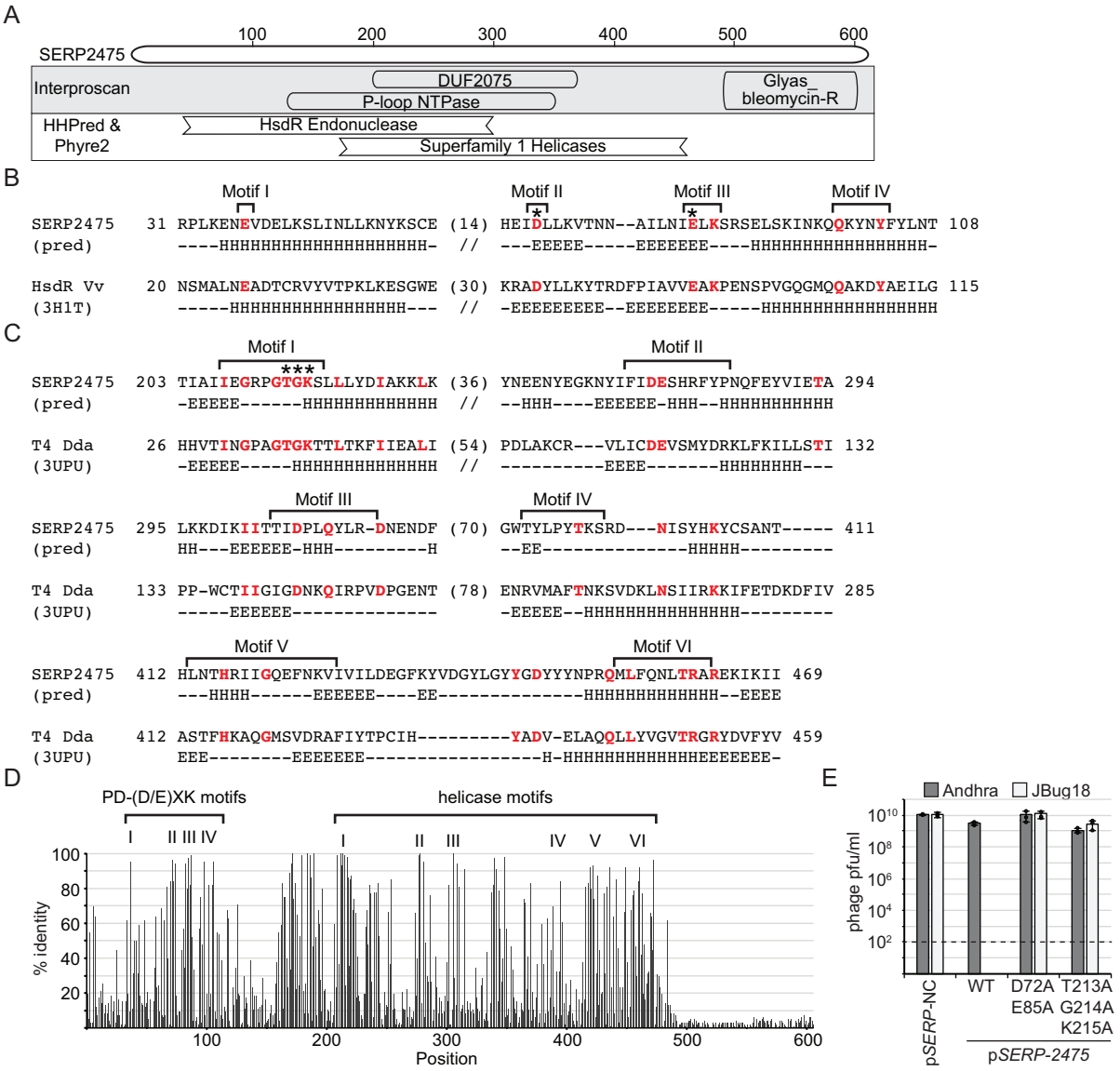


Figure 4. Conserved nuclease and helicase domains are required for immunity *in vivo*. (**A**) Predicted domains and structural homologs of SERP2475. (**B** and **C**) Pairwise sequence alignments between SERP2475 and relevant regions of the putative HsdR endonuclease from *Vibrio vulnificus* and the Dda helicase from phage T4 (Swiss PDB IDs 3H1T and 3UPU, respectively) 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. (**D**) The fraction of 99 homologs that possess amino acids identical to those of SERP2475 at each position in their multiple sequence alignment. Positions of critical amino acids comprising the PD-(D/E)XK nuclease and helicase motifs are indicated at the top. (**E**) *S. epidermidis* LM1680 strains bearing the indicated plasmids were challenged with phages Andhra and JBug18, and resulting pfu/ml were enumerated. Shown are an average of triplicate measurements (±S.D.). The dotted line indicates the limit of detection for this assay.

281 To test for these enzymatic functions, we next took a biochemical approach. *SERP2475*
282 was introduced in the pET28b expression vector downstream of an N-terminal 10His-Smt3 tag,
283 overexpressed in *E. coli*, and subjected to a three-step purification process involving two steps of
284 affinity purification followed by size exclusion chromatography (Fig. 5A). During the first two steps,
285 we noticed that SERP2475 co-purifies with a smaller species, and following digestion of the
286 10His-Smt3 tag by the SUMO protease, both exhibit a reduction in size of ~14 kDa (Fig. 5B). This
287 suggested that the smaller species is also tagged, and thus comprises an N-terminal fragment of
288 the full-length protein. To confirm, we excised the bands in the SDS-PAGE gel corresponding to
289 the full length (~72 kDa) and putative truncated variant (~23 kDa) in the flow-through fraction from
290 the second purification step and subjected the proteins to mass spectrometry analysis. The results
291 indicated that SERP2475 is indeed the most abundant protein in both bands (Tables S5 and S6),
292 with the truncated version showing dense peptide coverage over only the first ~230 amino acids
293 (AAs) (Fig. 5C). The subsequent size exclusion chromatography step successfully separated the
294 full-length and truncated versions, and also revealed that the full-length species forms a dimer in
295 solution, as evidenced by the presence of two adjacent peaks in the chromatogram that contain
296 the pure full-length protein (Fig. 5D).
297 We next conducted *in vitro* functional analyses. Nuclease assays were first performed with
298 the fractionated protein and revealed both 3'-5' exonuclease and plasmid nicking activities (Fig.
299 5E and F, respectively) stemming from fractions containing the N-terminal fragment. Interestingly,
300 these activities were nearly absent in fractions containing the full-length protein, suggesting a
301 mechanism of autoinhibition. To confirm that the N-terminal domain is sufficient to produce both
302 activities, we introduced the region of *SERP2475* encoding the first 230 AAs of the protein into
303 the pET28b expression vector downstream of the 10His-Smt3 tag and purified this truncated
304 version to homogeneity using the same purification protocol (Fig. 5A). We observed that the N-
305 terminal fragment elutes as a monomer (Fig. 5G), indicating that the dimerization domain is likely
306 located elsewhere on the protein. Importantly, nuclease assays confirmed that the N-terminal

307 domain is sufficient to produce robust exonuclease and nickase activities (Fig. 5H and I,
308 respectively). The enzyme is unable to further degrade the nicked or linearized double-stranded
309 DNA products when given up to 4 hours of incubation (Fig. 5I), suggesting that supercoiling is
310 essential for cleaving double-stranded DNA. Both activities require either Mg^{2+} or Mn^{2+} and can
311 be observed using different DNA substrates (Fig. S3), providing evidence that these activities are
312 sequence non-specific.

313 We also tested for helicase activity. As mentioned earlier, the central portion of SERP2475
314 has conserved helicase motifs and many of its predicted structural homologs are superfamily 1
315 helicases (Fig. 4). These enzymes are known to function as a monomer or dimer to unwind
316 double-stranded substrates using energy supplied by ATP (Fairman-Williams et al., 2010). To test
317 for this activity, the full-length fractions containing the monomer and dimer peaks (Fig. 5D) were
318 pooled, concentrated, and incubated with ATP, Mg^{2+}, and various DNA substrates (Fig. 5J). The
319 results showed that SERP2475 can indeed unwind double-stranded DNA when offered a 5'- or
320 3'- single-stranded overhang (Fig. 5K and L). Altogether, these results demonstrate that
321 SERP2475 possesses exonuclease and nickase activities stemming from its N-terminus, as well
322 as bidirectional helicase activity.

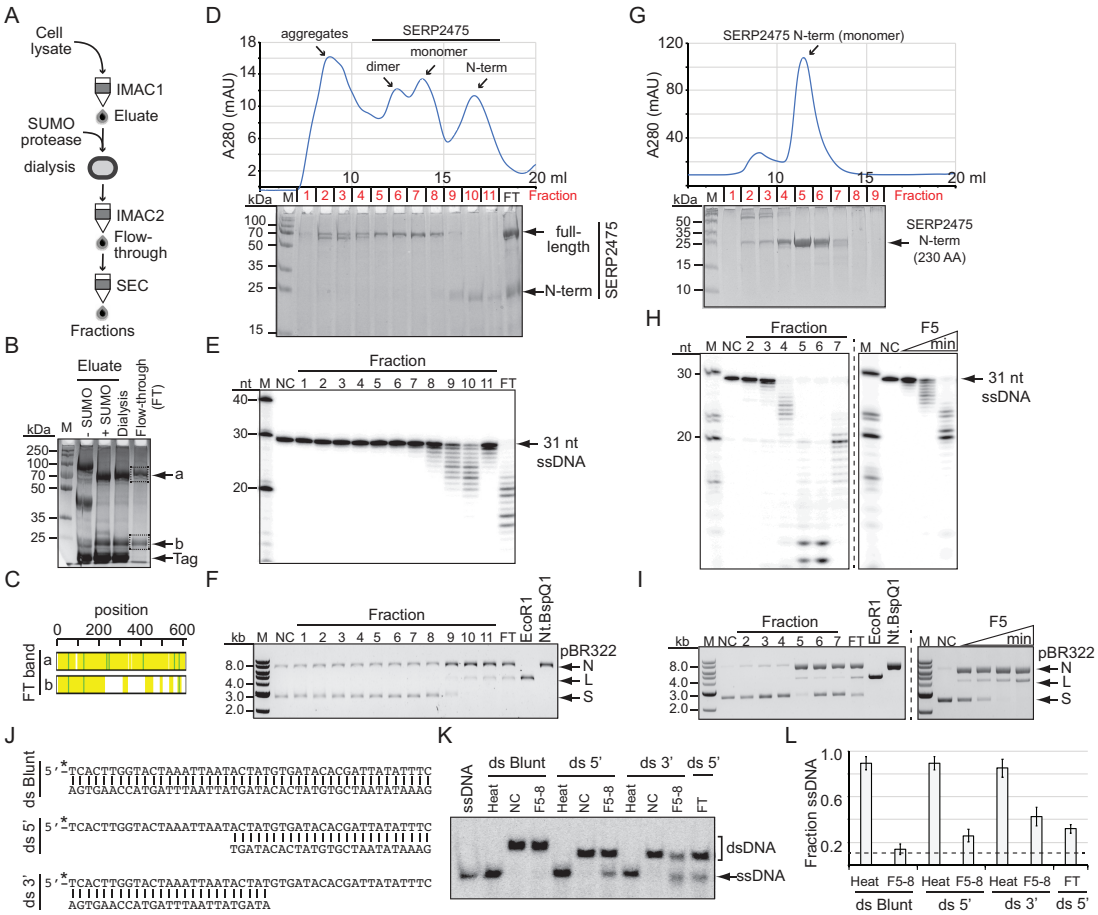


Figure 5. SERP2475 has nuclease and helicase activities. (**A**) Three-step protein purification process used in this study. IMAC, immobilized metal affinity chromatography; SEC, size exclusion chromatography. (**B**) Image of representative SDS-PAGE gel showing protein species present throughout the first two chromatography steps. 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 which indicated substrates were combined with fractions from the full-length protein prep in panel D. (**H** and **I**) Left, Exonuclease (H) and nickase (I) assays in which indicated substrates were combined with fractions from the N-terminal 230 AA prep in panel G. Right, time course assays in which substrates were incubated with fraction 5 (F5) in a 13 nM reaction for various amounts of time (5, 10, and 20 min (H) or 30, 60, 120, and 240 min (I)). Products of exonuclease and nickase reactions were resolved on denaturing urea-PAGE and native agarose gels, respectively. For nickase assays, EcoRI and Nt.BspQI were used as controls to generate linear (L) and nicked (N) products, respectively, from the supercoiled (S) plasmid. Shown are representatives of at least three independent trials. (**J**) Substrates used in helicase assays. (**K**) Helicase assays in which indicated substrates were incubated at 37°C for 1 h with or without full-length SERP2475 pooled fractions 5-8 (F5-8) in a 4 nM reaction and resolved using native PAGE. As a positive control for complete unwinding, substrates were heated 95°C for 2 min. (**L**) Quantification of unwound (ssDNA) substrate following heating or incubation with full-length SERP2475 F5-8. Shown is an average of four independent trials (± S.D.). M, molecular weight ladder; kb, kilobase; nt, nucleotide; kDa, kilodalton; NC, negative control, ss, single-stranded; ds, double-stranded. See also Figure S3.

332 **A single-stranded DNA binding protein protects against SERP2475 immunity**

333 To further refine our understanding of SERP2475's targeting specificity, we sought to
334 determine how Andhra resists immunity. Andhra and JBug18 encode the same 20 proteins (Fig.
335 S4A), and a pairwise alignment of their coding regions show that they differ at only 705 positions
336 by either a single-nucleotide polymorphism (SNP) or a gap (Supplementary file 2). To narrow
337 down which SNPs and/or gaps in Andhra are important for resistance to immunity, we first
338 attempted to isolate naturally-evolved JBug18 mutants that can escape immunity by plating
339 concentrated phage preparations with LM1680/p*SERP-2475*. After several failed attempts at
340 recovering plaques, one attempt yielded resistant phages, which upon further inspection were
341 found to possess hybrid genomes that contain a patchwork of Andhra and JBug18 sequences.
342 These hybrids necessarily arose through the inadvertent mixing of the two phages and
343 propagation on the same host strain. Nonetheless, this fortuitous accident proved invaluable in
344 helping to pinpoint the region required for immune resistance—since all hybrids can escape
345 immunity, we reasoned that they must share Andhra-derived sequences in the region(s) required
346 for resistance. To test this, we purified and sequenced eight such hybrids, and determined the
347 fraction of hybrids that possess Andhra identity at each of the 705 differing positions across their
348 coding regions. We found that all eight hybrids share Andhra identity at positions 891-2117 in the
349 alignment (Fig. 6A and Supplementary file 2). This region overlaps gene products (*gp*) *03-06* in
350 the phage genomes and encompasses 69 SNPs and gaps, of which, 64 are concentrated within
351 *gp03* and *gp04* (Fig. 6B). Accordingly, we speculated that one or both of the latter genes are
352 responsible for resistance.

353 To narrow down the protective region even further, a second set of resistant JBug18
354 hybrids were generated which bear Andhra-derived sequences in *gp03* and *gp04*. This was
355 accomplished by introducing Andhra's *gp03* and/or *gp04* coding regions into *S. epidermidis*
356 LM1680 on plasmids, and then propagating JBug18 on these strains to allow the phage to

recombine with the Andhra-derived sequences (Fig. S4B). The resulting phages were then plated on LM1680/p*SERP-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-nucleotide stretch spanning positions 1302-1362 in Andhra's genome (Fig. 6C and Supplementary file 2). This region overlaps *gp03*, which encodes a single-stranded DNA binding protein (SSB, Fig 6D, and Supplementary file 3). Importantly, JBug18 has a 5-nucleotide insertion in this region and consequently harbors a truncated variant of the SSB (Fig. 6D and S4C). However, by acquiring the 60-nucleotide stretch from Andhra, all ten hybrids had restored the reading frame and hence encode a full-length SSB, suggesting that the C-terminus of the SSB is essential for escape from immunity. In agreement with these observations, phages Pontiff and Pike possess the expected *gp03* genotypes: While the resistant phage Pontiff encodes a full-length SSB, the sensitive phage Pike encodes a truncated version, this time due to a single nucleotide deletion (Fig. S4D).

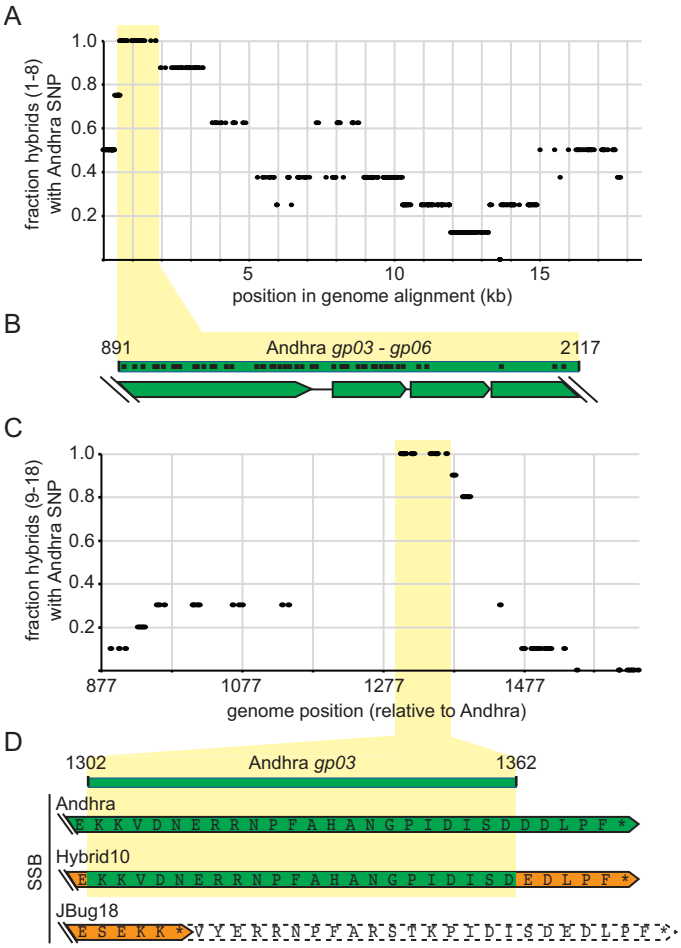


Figure 6. A phage-encoded single-stranded DNA binding protein protects against immune function. (**A** and **C**) Fractions of JBug18-Andhra Hybrids that harbor Andhra identity at positions where the parental phage genomes differ. Segments common to all hybrids are highlighted in yellow. (**B** and **D**) Expanded views of the highlighted common regions. Black dots in (B) indicate relative positions where SNPs and/or gaps occur in an alignment between the parental phages. In (D), green and orange arrows with solid borders delimit phage-encoded SSBs, and the white arrow with dotted border delimits a truncated portion. See also Figure S4.

383 **Discussion**

384

385 Here, we describe a new mode of nucleic acid immunity performed by a single enzyme
386 with nuclease and helicase activities. Taking into account these demonstrated functions, we
387 propose to name this enzyme and its homologs Nhi (Nuclease-helicase mediated immunity). One
388 attribute that sets Nhi apart from other DNA-targeting immune systems is that it does not rely
389 upon recognition of specific DNA sequences (Fig. 4), and yet abrogates phage DNA accumulation
390 without causing appreciable cell death (Fig. 3). These observations beg the question: What is the
391 basis for phage specificity? Our finding that the SSB in one group of phages can protect against
392 immunity (Fig. 6) supports a preliminary model in which Nhi senses and degrades replication
393 intermediates that are unique to these phages (Fig. 7). Staphylococcal *Podoviridae* phages use
394 a protein-priming mechanism of replication (Salas et al., 2016; Vybiral et al., 2003), in which
395 terminal proteins (TPs) covalently linked to the 5'-ends of their linear dsDNA provide a hydroxyl
396 group upon which to initiate replication (Fig. 7A). The initial stages of replication involve local
397 unwinding of DNA ends by double-stranded DNA binding proteins and subsequent release of the
398 free 3'-end. Under normal circumstances, the 3'-end is captured by the phage DNA polymerase
399 (DNAP) in complex with a second TP, whereupon replication ensues. We speculate that Nhi may
400 compete with DNAP for the free 3'-end and use its 3'-exonuclease and helicase activities to
401 processively degrade the phage genome. SSBs are known to bind and protect DNA and regulate
402 the replication machinery, particularly through interactions with their C-terminus (Shereda et al.,
403 2008). The precise mechanism by which the C-terminus of the *Podoviridae* phages' SSB protects
404 against Nhi's effects remains to be determined.

405 Remarkably, phages with very different mechanisms of replication also succumb to Nhi.
406 These include the Twort-like phage ISP, and lambda-like phages Lorac and CNPx (Fig. 1).
407 Although the replication mechanism of the former remains poorly understood (Klumpp et al.,
408 2010), the latter undergo several rounds of theta replication early in the infection cycle, followed

409 by rolling-circle replication (Casjens and Hendrix, 2015; Narajczyk et al., 2007) (Fig. 7B). We
410 speculate that Nhi may also recognize and nick supercoiled theta replication intermediates. Since
411 the full-length Nhi forms a dimer in solution (Fig. 5D), it has the potential to introduce two adjacent
412 nicks on opposing strands and cause double-stranded breaks in the phage genome. Indeed, even
413 the monomeric 230 AA fragment caused some plasmid linearization (Fig. 5F and I), presumably
414 through simultaneous nicking on opposite strands that co-localized by chance. Since full-length
415 Nhi is largely devoid of nuclease activity *in vitro* (Fig. 4), we speculate such phage-specific
416 replication intermediates, perhaps in combination with one or more phage-encoded proteins
417 bound to them could constitute recognition elements that trigger Nhi activation.

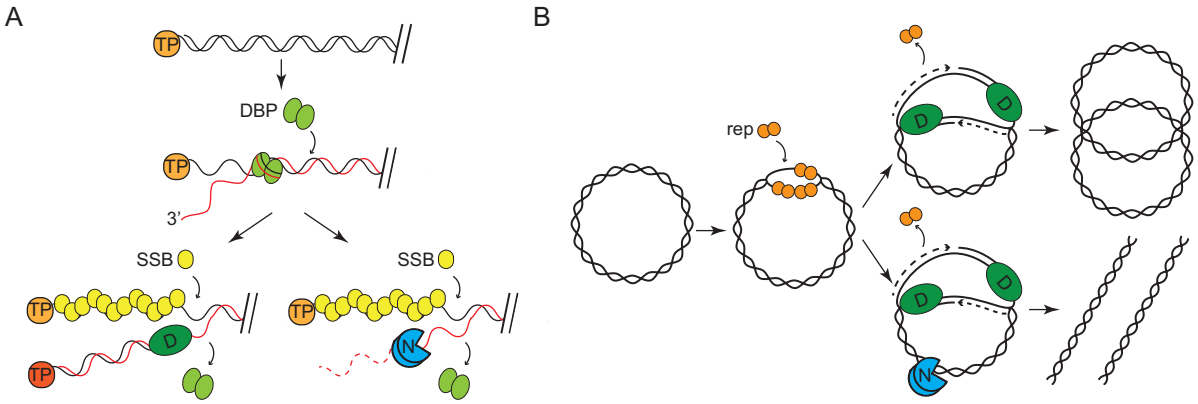
418 It is unclear how the full length version of Nhi is truncated within the cells prior to
419 purification (Fig. 5B and C), and we cannot exclude the compelling possibility that generation of
420 the truncated form might constitute a natural part of its mechanism of action. Nonetheless, its
421 unexpected appearance allowed us to reconstitute the nuclease activities stemming from Nhi's
422 N-terminal PD-(D/E)XK domain. These domains are found in a broad superfamily of metal-
423 dependent nucleases involved in diverse functions, including DNA restriction, recombination, and
424 repair (Steczkiewicz et al., 2012). Members of this group include restriction enzymes, holliday
425 junction resolvases, herpesvirus exonucleases, and many others from in all kingdoms of life. Nhi's
426 helicase motifs comprise a separate domain that overlap with the DUF2075 (Fig. 4). In addition
427 to the phage T4 Dda helicase, close predicted structural homologs include the human helicase
428 Upf1 (E-value 2.5 x 10^{-23}), and mouse nuclease-helicase DNA2 (E-value 1.2 x 10^{-21}), which are
429 involved in nonsense-mediated mRNA decay and DNA replication/repair, respectively
430 (Chakrabarti et al., 2011; Zhou et al., 2015). All three are superfamily 1B helicases, which unwind
431 double-stranded substrates in the 5'-3' direction. Interestingly, Nhi's helicase activity is bipolar
432 (Fig. 5K), which undoubtedly allows it to act on more diverse substrates. We also noted the
433 presence of a C-terminal domain (Glyas_Bleomycin-R) that is poorly conserved across Nhi
434 homologs (Fig. 4A and D). Such domains are found in a group of metalloenzymes that perform a

435 variety of activities, including isomerizations, epimerizations, and oxidative cleavage of C-C bonds
436 (Armstrong, 2000). Whether and how this domain coordinates with the others to achieve immunity
437 are subjects of ongoing work.

438 In contrast to the *S. epidermidis* Nhi, other homologs tested showed more modest anti-
439 phage activity (Fig. 2). This could possibly be explained by incompatibility with the foreign phage
440 target or heterologous host background, which together might lessen their apparent effectiveness.
441 In support of this, the homologs that were most effective against *S. aureus* RN4220 phage Lorac
442 originated from two different *S. aureus* strains. Nonetheless, the observation that the *V. vulnificus*
443 homolog could still afford protection is remarkable and underscores the versatility of this mode of
444 prokaryotic immunity. Interestingly, one of the *S. aureus* homologs abrogated cell growth while
445 also blocking phage replication (Fig. 2), suggesting that it may cause collateral cleavage of host
446 DNA, and as a consequence, slow cell division. Indeed, the latter represents a valid strategy for
447 nucleic acid immunity, as some CRISPR-Cas systems are known to simultaneously damage
448 phage and host nucleic acids and cause cell dormancy and/or death during persistent phage
449 infection (Meeske et al., 2019; Ofir et al., 2017; Rostøl and Marraffini, 2019; Watson et al., 2019).

450 Finally, it bears mentioning that the SERP2475 homologs identified in this study represent
451 but a small subset of DUF2075 domain-containing proteins. According to Interproscan, these
452 constitute a large family of conserved proteins with over 7,000 members that can be found in
453 organisms spanning all domains of life. Although the majority of members are encoded in bacteria,
454 many are found in eukaryotes, a handful of which are in humans. Of these, the Schlafen (slfn)
455 family proteins Slfn5, Slfn11, and Slfn13 have been shown to restrict the replication of diverse
456 viruses, including human immunodeficiency virus (HIV) and herpes simplex virus (HSV) (Kim et
457 al., 2021; Li et al., 2012; Valdez et al., 2019; Yang et al., 2018). Notably, Slfn13 also relies upon
458 endonuclease activity for anti-viral defense (Yang et al., 2018). Such functional conservation
459 across phylogenetic boundaries has become a recurring theme in recently-described defense
460 systems (Bernheim et al., 2021; Cohen et al., 2019; Kazlauskiene et al., 2017; Niewoehner et al.,

461 2017), and we anticipate continued investigation of prokaryotic DUF2075 proteins has the
462 potential to seed new insights into human immunity.
463
464



465
466 **Figure 7.** Proposed mechanism for SERP2475 (Nhi) immunity. (**A**) Protein-priming mechanism
467 of replication in staphylococcal *Podoviridae* phages. (**B**) Theta replication in staphylococcal
468 *Siphoviridae* phages. TP, terminal protein; DBP, double-stranded DNA binding protein; rep, replication initiator protein; D, DNA polymerase; N, Nhi. Refer to the discussion for more details.

Acknowledgements

AH-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.

Author Contributions

S.M.N.B., L.C-Z., O.H., K.C., V.D., and A.T., designed, created, and tested plasmid constructs; S.M.N.B. performed adsorption, cell viability, and phage hybridization assays; LCZ performed qPCR assays and analyses; S.M.N.B. and L.C-Z. performed protein purification; S.M.N.B., L.C-Z., and A.H-A. performed biochemical assays and analyses; B.A. performed all bioinformatic analyses; and A.H-A. conceived the study and wrote the manuscript. All authors have read and approved the manuscript.

Declaration of Interests

The authors declare no competing interests.

484 **Methods**

485 ***Bacterial strains and growth conditions***. S*. epidermidis* RP62a and mutant variants were a
486 generous gift from Luciano Marraffini. *S. epidermidis* strains were grown in Brain Heart Infusion
487 (BHI, BD Diagnostics), *S. aureus* strains were grown in Tryptic Soy Broth (TSB, BD
488 Diagnostics), *E. coli* DH5α was grown in Luria Bertani (LB) broth (VWR), and *E. coli* Rossetta2
489 (DE3) was grown in Terrific Broth (TB, VWR) for protein purification. Growth media was
490 supplemented with the following: 10 µg/ml chloramphenicol (to select for pC194-based
491 plasmids), 10 µg/ml tetracycline (to select pT181-based plasmids), 10 µg/ml erythromycin (to
492 select for pTET-based plasmids), 15 µg/ml neomycin (to select for *S. epidermidis* cells), 30
493 µg/ml chloramphenicol (to select for *E. coli* Rossetta2 plasmids) and 50 µg/ml kanamycin (to
494 select for pET28b-10HisSmt3-based plasmids).

495 ***Phage propagation and plate infection assays.*** *S. epidermidis* phages (Andhra, JBug18,
496 Pontiff, Pike, Quidividi, and Twillingate) and *S. aureus* phages (ISP, Lorac, and Pabna) were
497 propagated on their respective hosts to create stocks. Stock concentrations in plaque-forming
498 units per ml (pfu/ml) were determined as previously described (Antoine Marc et al., 2019; Cater
499 et al., 2017; Culbertson et al., 2019; Freeman et al., 2019). For plate infection assays, 10-fold
500 dilutions of phage stocks were spotted atop a lawn of indicated strains using the double-agar
501 overlay method (Cater et al., 2017). For assays using the anhydrotetracycline- (aTc-) inducible
502 system (Fig. 2), plates and top agar were supplemented with 15 or 30 mg/L aTc, as indicated.
503 Following overnight incubation, phage pfu/ml were enumerated.

504 ***Construction of pC194, pT181, and pTET-based plasmids and introduction into***
505 ***staphylococcal strains.*** All pC194-, pT181-, and pTET-based plasmids were constructed using
506 either inverse PCR or Gibson Assembly(Gibson et al., 2009) with the primers listed in Table S7
507 as described previously (Walker et al., 2017). pAH011 (Hatoum-Aslan et al., 2011), a derivative
508 of pC194 (Ehrlich, 1977), was used as the backbone for plasmids designated as p*SERP*- in this
509 study. Plasmid pT181 (Khan et al., 1981) was used as backbone for pT181-*gp03* and pT181-

510 *gp0304*. pTarget (Samai et al., 2015) was used as the backbone for all plasmids designated as
511 pTET- in this study. All assembled plasmids were first introduced into *S. aureus* RN4220
512 (pC194- and pTET-based plasmids) or OS2 (pT181-based plasmids) via electroporation, and
513 inserted sequences were confirmed by PCR amplification and Sanger sequencing (performed
514 by Eurofins MWG Operon) using primers shown in Table S7. Confirmed plasmids were purified
515 using EZNA Plasmid Mini Kit (Omega Bio-tek), and where indicated, introduced into *S.*
516 *epidermidis* LM1680 via electroporation as previously described (Walker et al., 2017).
517 ***Homolog identification.*** To identify homologs, the amino acid sequence of SERP2475 was
518 independently queried against three databases (prok_complete_genomes, refseq_genomes,
519 and nt) using NCBI's tBLASTn webserver and hits were downloaded in xml format. Using a
520 Python script utilizing Biopython v. 1.7.8 functions (Cock et al., 2009), xml files were parsed and
521 combined into a unique set of hits. Fully annotated complete genomes containing the homologs
522 from the combined list of hits were downloaded from NCBI in genbank (gbk) format (~3GB).
523 Each genome was parsed and coding sequence (CDS) features of corresponding BLAST hits
524 were extracted. Proteins with unique accession numbers and unique sequences (excluding
525 pseudogenes) were retained and combined into a fasta file. Hits shorter than 200 amino acids
526 were eliminated from the list to obtain the final set of homologs in distinct genetic backgrounds
527 (Table 1).
528 ***Phylogenetic tree generation.*** An iterative process was used to select the final set of 100
529 homologs and build the tree. The fasta file with all identified homologs was first submitted to the
530 MAFFT webserver (used May 8th, 2021, https://mafft.cbrc.jp/alignment/server/) to obtain a
531 multiple sequence alignment (MSA) with the E-INS-I option selected along with the remainder
532 parameters set to their default values. Upon inspection of the MSA, low-scoring homologs that
533 were also observed to be from adequately represented genera/species but introducing large
534 gaps in the alignment were removed from consideration. The MSA computation was then
535 repeated as above. Using this second MSA, a preliminary phylogenetic tree was generated with

536 IQ-TREE21 multi-core v. 1.6.12 (Nguyen et al., 2014) with the optimal substitution model
537 LG+R6 that provided the lowest Bayesian information criterion. One thousand ultra-fast
538 bootstraps were performed to evaluate node support (options –bb 1000 –wbt). Upon inspection
539 of the resulting tree, hard polytomies (Sayyari and Mirarab, 2018) that resulted in many
540 extremely short branches were identified and, to preserve the phylogenetic diversity, only one
541 representative homolog for each were retained to perform the final MAFFT alignment. A final
542 MSA was computed with the remaining homologs (n=100 including SERP2475) and the final
543 phylogenetic tree was computed using IQ-TREE with the same parameters noted above.
544 RAxML-NG v. 1.0.2 (Kozlov et al., 2019) was used to confirm the tree with similar corresponding
545 parameters (100 bootstraps). Fig Tree v. 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) was
546 used for tree visualization, and Adobe Illustrator was used to overlay highlights and markings
547 relevant to the study.
548 ***Homolog neighborhood analysis.*** For the neighborhood analysis, the genomes containing the
549 original 303 homologs (including SERP2475) were parsed and the amino acid sequences of
550 neighboring proteins encoded on either side of the Blast hit (window sizes of 10, 20, and 30)
551 were extracted into individual fasta files using a Python script. Pfam 34.0 database was
552 downloaded (Pfam-A.hmm) and hmmpress (hmmer 3.3.2, hmmer.org) was used to index it. For
553 each of the neighborhood fasta files, a Python script utilizing hmmscan was used to obtain and
554 generate a new set of files that included protein family (pfam) predictions of its contents.
555 Predicted pfams were then searched against a set of 306 pfams with known defense-related
556 functions (Table S3)—this list was compiled from the old and newly-identified defense pfams
557 cited in (Gao et al., 2020). A Python script was then developed and utilized to determine the
558 defense related neighbors for each homolog and analyzed to generate the plots for each
559 neighbor window size.
560 ***Phage adsorption and cell viability assays.*** Overnight cultures of *S. epidermidis*
561 LM1680/p*SERP-NC* and LM1680/p*SERP-2475* were diluted 1:100 in fresh BHI supplemented

562 with antibiotics and 5 mM $CaCl_2$, and incubated at 37°C with agitation for one hour. For the
563 adsorption assay, Andhra or JBug18 were added to cultures (0.01:1 phage:cell ratio) and
564 incubated at 37°C for 10 min. Cells along with adsorbed phages were pelleted at 8000 *x g* for 5
565 min at 4°C and resulting supernatants were passed through 0.45 mm syringe filter. The number
566 of free phages in the supernatants were enumerated by the double-agar overlay method (Cater
567 et al., 2017). The number of adsorbed phages were determined by subtracting the number of
568 phages in suspension from the number that was initially added. Triplicate samples were
569 prepared for each treatment, and two independent trials were conducted. For cell viability
570 assays, 200 µl of the bacterial cultures were distributed into a 96-well microtiter plate (into
571 triplicate wells for each treatment), and phages were added to cells at ratios of 1:1, 5:1, or 10:1.
572 These bacteria-phage mixtures were incubated at 37°C with agitation for an additional five
573 hours. 25 µl of 0.1% (w/v) 2,3,5 triphenyltetrazolium chloride (TTC) was added into each well
574 and the microtiter plate was incubated at 37°C for an additional 30 mins to allow the colorless
575 TTC to become enzymatically reduced to the red 1,3,5-triphenylformazan product by actively
576 growing bacterial cells. The relative numbers of viable cells were then determined by measuring
577 the absorbance at 540 nm. Triplicate measurements were averaged, and two independent trials
578 were conducted.
579 ***Phage infection time course followed by quantitative PCR.*** Phage infection time course
580 assays were conducted in liquid media as previously described (Chou-Zheng and Hatoum-
581 Aslan, 2019). Briefly, *S. epidermidis* LM1680 mid-log cells bearing p*SERP-NC* or *pSERP-2475*
582 were infected with Andhra or JBug18 (phage:cell ratio of 0.5:1), cells were harvested at 0, 10, or
583 20 minutes post-infection, and their total DNA was extracted. Each qPCR reaction (25 µl)
584 contained 500 ng of total DNA as template, 0.4 nM of phage-specific primers (N233 and N234)
585 or host-specific primers (S001 and S002) (Table S7), and 1X PerfeCTa SYBR Green SuperMix
586 (Quanta Biosciences). Separate standard reactions containing 10^2–10^9 DNA molecules were
587 also prepared using purified Andhra phage DNA extract, JBug phage DNA extract, or bacterial

588 genomic DNA extract. A CFX Connect Real-Time PCR Detection System (Bio-Rad) was used to
589 amplify the DNA templates. Phage DNA copy number was normalized against host values, and
590 the normalized value for the 0 min time point was set to one to obtain the relative DNA
591 abundance for the rest of the time points as described previously (Chou-Zheng and Hatoum-
592 Aslan, 2019). Triplicate measurements were taken for each of two independent trials.
593 ***Construction of pET28b-10His-Smt3-based plasmids.*** pET28b-10His-Smt3-*SERP2475* and
594 pET28b-10His-Smt3-*SERP2475*-230AA were constructed via Gibson assembly. Briefly, inserts
595 were obtained by amplifying the complete and partial portion of *SERP2475* from *S. epidermidis*
596 RP62a genome, and the backbone was amplified from a pET28b-His10Smt3 template.
597 Corresponding primers are listed in Table S7. The backbone PCR products were subjected to
598 DpnI (NEB) digestion. Then, all PCR products were purified using the EZNA Cycle Pure Kit and
599 Gibson assembled. Assembled constructs were introduced into *E. coli* DH5α via chemical
600 transformation, and transformants were confirmed to have the desired plasmids using PCR and
601 DNA sequencing with primers T7P and T7T (Table S7). Confirmed plasmids were purified using
602 the EZNA Plasmid Mini Kit and introduced into *E. coli* Rosetta2 (DE3) cells for protein
603 purification.
604 ***Purification of recombinant SERP2475 and SERP2475-230AA.*** Recombinant proteins
605 encoded in pET28b-10His-Smt3-based plasmids were overexpressed and purified as described
606 previously with some modifications (Chou-Zheng and Hatoum-Aslan, 2019). Briefly, overnight
607 cultures were diluted 1:100 in 1 L (for the truncated version) or 2 L (for the full-length version) of
608 TB supplemented with appropriate antibiotics. Once the OD_{600} reached 0.5–0.6, cell-growth was
609 arrested on ice for 20 minutes, and protein expression was induced with 0.3 mM isopropyl-1-
610 thio-β-d-galactopyranoside (IPTG) and 2% ethanol. Induction proceeded 16-18 hr at 17°C with
611 constant shaking. Cells were harvested and washed with cold PBS Buffer (137 mM NaCl, 2.7
612 mM KCl, 10 mM Na_2HPO4, 1.8 mM KH_2PO4). All subsequent steps were performed at 4°C.
613 Each one-liter pellet was suspended in 30 ml of Buffer A (50 mM Tris–HCl, pH 9.0, 1.25 M

614 NaCl, 200 mM Li_2SO4, 10% sucrose, 15 mM Imidazole) containing one complete EDTA-free
615 protease inhibitor tablet (Roche), 0.1 mg/ml lysozyme, and 0.1% Triton X-100. After 1 hr
616 rotation, lysed cells were sonicated, and the insoluble materials were removed via centrifugation
617 and filtration. Then, 3 ml (for 1-L pellet of truncated version), or 4 ml (for 2-L pellet of full-length
618 version), of slurry Ni^{2+}-NTA-agarose resin (ThermoFisher) was equilibrated with Buffer A and
619 mixed with the cleared lysates. After 1 hr incubation, the resin was collected via centrifugation
620 and washed with 40 ml of Buffer A. The resins containing the full-length protein from individual
621 1-L pellets were pooled. Resins were then transferred to a 5-ml gravity column (G-Biosciences)
622 and further washed with 25 ml of Buffer A. Proteins were collected into 1-ml aliquots and eluted
623 in stepwise elution with 3 ml of IMAC buffer (50 mM Tris–HCl pH 9.0, 250 mM NaCl, 10%
624 glycerol) containing 50-, 100-, 200-, and 500-mM imidazole, respectively. The 10His-Smt3-tag
625 was removed from most concentrated eluates using SUMO Protease (McLab, 1000 U) and
626 supplied buffer (salt-free). Samples were dialyzed against IMAC buffer containing 25 mM
627 Imidazole for 3 hr. Then, 2 ml (for the truncated version), or 1.5 ml (for the full-length version), of
628 slurry Ni^{2+}-NTA-agarose resin was equilibrated with IMAC Buffer containing 25mM Imidazole,
629 and incubated with the dialysate for 1 hr. The resin was collected in a 5-ml gravity column, and
630 tag-free proteins were collected and concentrated using a 10K MWCO centrifugal filter (PALL).
631 Concentrated proteins were further purified by size exclusion chromatography using Superdex
632 75 Increase 10/300 GL (Cytiva) (for the truncated version) and Superdex 200 Increase 10/300
633 GL (Cytiva) (for the full-length version). Collected protein fractions were subjected to nuclease
634 assays. Fractions with the highest peak concentration were combined and concentrated for
635 nuclease time course and helicase assays. Proteins were resolved on 12% SDS-PAGE run at
636 120 V for 1.5 hours and visualized with Coomassie G-250, and the concentrations were
637 determined using Bradford reagent (Bio-Rad).
638 ***Nuclease assays.*** For exonuclease assays, single stranded DNA substrates were labeled on
639 their 5'-ends by incubating with T4 polynucleotide kinase and γ-[^{32}P]-ATP and purified over a

640 G25 column (IBI Scientific). Radiolabeled substrates were combined with 7.5 µl of each protein
641 fraction, or 130 pmols for time course assays, in 10 µl reactions containing nuclease buffer (25
642 mM Tris-HCl pH 7.5, 2 mM DTT) and 10 mM of $MgCl_2$. Nuclease reactions were incubated at
643 37°C for 20 minutes (for experiments with different fractions), or 5, 10, and 20 minutes (for time
644 course assays). Reactions were stopped by adding an equal volume of 95% formamide loading
645 buffer and resolved on a 15% Urea PAGE gel at 55 W for 1.5 hours. Gels were exposed to a
646 storage phosphor screen and visualized using an Amersham Typhoon biomolecular imager. For
647 nickase assays, 250 ng of plasmids pBR322 or pUC19 (NEB) were combined with 15 µl of
648 protein fractions, or 260 pmols for time course assays, in 20 µl reactions containing nuclease
649 buffer (25 mM Tris-HCl pH 7.5, 2 mM DTT) and 10 mM of $MgCl_2$. Nickase reactions were
650 incubated at 37°C for 60 minutes (for experiments with different fractions), or 30, 60, 120, and
651 240 minutes (for time course assays). Reactions were stopped by placing on ice for 5 min,
652 followed by incubation with 10 µg of proteinase K for 20 minutes at room temperature. Samples
653 were then resolved on a 1% agarose gel run at 120 V for 50 minutes and visualized with
654 ethidium bromide under UV transillumination with an Azure 400 imager.
655 ***Helicase Assay***. Double-stranded DNA duplexes were prepared by combining 5′-radiolabeled
656 ssDNA oligonucleotides and unlabeled complementary ssDNA oligonucleotides (as shown in
657 Fig. 5J) in a 1:2.5 molar ratio. The mixtures were heated to 95°C for 5 min and then slowly
658 cooled down to room temperature over a period of 3 hours. The helicase assay was performed
659 by first mixing radiolabeled DNA duplex with a 10-fold molar excess of unlabeled top-strand
660 DNA to trap the complementary strand once unwound. This mixture was combined with 200
661 pmol SERP2475 in helicase buffer (25 mM Tris-HCl pH 9.0, 2 mM DTT, and 0.1 mg/ml BSA)
662 supplemented with 2 mM $MgCl_2$ and 5 mM ATP in a 50 µl reaction. Reaction mixtures were
663 incubated at 37°C for 1h. As a separate positive control for unwinding, DNA substrates were
664 heated to 95°C for 10 min in the absence of the enzyme. Reaction was stopped by adding 5 µl
665 of the stop solution (0.1% [wt/vol] bromophenol blue, 0.1% [wt/vol] xylene cyanol, 8% [vol/vol]

666 glycerol, 0.4% [wt/vol] SDS, 50 mM EDTA). Samples were resolved on an 8% (v/v) non-
667 denaturing polyacrylamide gel at 130 V for 4 hours at 4°C. The gel was dried under vacuum at
668 80°C, exposed to a storage phosphor screen and visualized using an Amersham Typhoon
669 biomolecular imager. Four independent trials were performed. ImageQuant software was used
670 for densitometric analysis. The fraction of unwound (ssDNA) was determined using the following
671 equation: density of ssDNA signal divided by the sum of densities of dsDNA and ssDNA signals.
672 ***Mass spectrometry analysis.*** Protein bands corresponding to the 72 kDa (full-length) and 23
673 kDa (truncated) variants of SERP2475 in the FT fraction from the second step of purification
674 were excised from a 12% SDS-PAGE gel stained with Coomassie G-250. Mass spectrometry
675 was performed by the Cancer Center Mass Spectrometry and Proteomics Shared Facility at the
676 University of Alabama, Birmingham. The bands were digested overnight with Trypsin Gold,
677 Mass Spectrometry Grade (Promega, cat. #V5280) following the manufacturer's instructions.
678 Peptide extracts were reconstituted in 0.1% formic acid/ddH_2O at 0.1 µg/µl. Electrospray
679 ionization tandem mass spectrometry was carried out, and the data were processed, filtered,
680 grouped, and quantified, as previously reported in detail (Ludwig et al., 2016). The data were
681 searched against a tailored database comprising of the *E. coli* proteome plus the protein
682 sequence of interest (SERP2475).
683 ***Phage hybrid generation and sequencing.*** JBug18-Andhra Hybrids 1-8 were isolated as
684 immune resistant mutants following challenge of LM1680/p*SERP-2475* with a high titer lysate of
685 JBug18 (~1 x 10^{10} pfu/ml). To generate JBug18-Andhra Hybrids 9-18, overnight cultures of *S.*
686 *epidermidis* LM1680 harboring pT181-*gp03* or pT181-*gp0304* were diluted 1:100 in fresh TSB
687 supplemented with antibiotics and 5 mM $CaCl_2$. The mixture was incubated at 37°C for an hour
688 with agitation, then JBug18 was added to the cells in a 1:1 ratio, and the incubation continued
689 with agitation overnight. The next day, cells were pelleted by centrifugation at 8000 *x g* for 5 min
690 and supernatant was filtered through 0.45 mm filter. Filtered lysates were mixed with LM1680-
691 p*SERP2475* overnight culture (1:1) and the mixture was plated on TSA containing 5 mM $CaCl_2$

692 using the double-agar overlay method (Cater et al., 2017). For all phage hybrids, individual
693 plaques were isolated and re-plated three times on LM1680/p*SERP-2475* to purify. Phages
694 were propagated and their DNA was extracted as previously described (Bari et al., 2017).
695 Phage genomes were PCR amplified across the entire coding region for Hybrids 1-8 or *gp03-*
696 *gp04* for Hybrids 9-18, and the PCR products were sequenced by the Sanger method (at
697 Eurofins MWG Operon) using the primers listed in Table S7.
698 ***Hybrid phage genome sequence analysis.*** For JBug18-Andhra Hybrids 1-8, Sanger
699 sequencing reads covering their coding regions were manually assembled using SnapGene
700 software. For JBug18-Andhra Hybrids 9-18, a single read covered the region of interest,
701 therefore no assembly was required. Sequences for each set of hybrids (1-8 and 9-18) were
702 aligned with corresponding genomic regions in Andhra and JBug18 using the Clustal Omega
703 Multiple Sequence Alignment tool (https://www.ebi.ac.uk/Tools/msa/clustalo/). The sequence
704 alignments (Supplementary files 2 and 3) were analyzed by a Python script developed in-house
705 which first scans the alignment of JBug18 and Andhra, identifies each position of non-similarity,
706 and then determines at those positions the fraction of hybrids that possess Andhra identity. The
707 output data was exported into an Excel file, and the graphs showing the fraction of hybrids with
708 Andhra identity at each position were generated using Microsoft Excel.
709 ***Data availability.*** The raw phage sequence reads that support the findings of Figure 6 are
710 publicly available through Figshare (10.6084/m9.figshare.9598040). The Accession codes for
711 Andhra, JBug18, Pontiff, and Pike genomes are KY442063, MH972263, MH972262, and
712 MH972261, respectively. Phages, mutant derivatives, and constructs can be made available
713 upon written request.
714 ***Code availability.*** The Python code library used for homolog and neighborhood analysis in
715 Figure 2 is freely available at GitHub (https://github.com/ahatoum/Nhi). The Python code written
716 to analyze the data for Figure 6 is freely available at GitHub (https://github.com/ahatoum/Hybrid-
717 phage-genome-sequence-analysis).

Supplementary Information

The following supplementary files are also available:

Supplementary Figures 1-4

Table S1. Accompanies figure 2. SERP2475 and 302 homologs in unique genetic backgrounds

Table S2. Accompanies figure 2. Protein families (Pfams) associated with anti-phage defense

Table S3. Accompanies figure 2. Defense-related Pfam neighbors within 20 genes of SERP2475 homologs

Table S4. Accompanies figure 2. Rank ordering of defense-related Pfams within 20 genes of SERP2475 and its homologs

Table S5. Accompanies figure 5C. Mass spectrometry analysis of proteins contained within "band a".

Table S6. Accompanies figure 5C. Mass spectrometry analysis of proteins contained within "band b".

Table S7. Oligonucleotides used for cloning and PCR in this study.

Supplementary file 1. SERP2475 and 99 selected homologs

Supplementary file 2. Multiple sequence alignment for JBug18-Andhra Hybrids 1-8

Supplementary file 3. Multiple sequence alignment for JBug18-Andhra Hybrids 9-18

References

Antoine 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*, e00586-19.

Armstrong, R.N. (2000). Mechanistic Diversity in a Metalloenzyme Superfamily. Biochemistry *39*, 13625–13632.

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., Adriaenssens, E.M., and Consortium, I.R. (2020). ICTV Virus Taxonomy Profile : Herelleviridae. J. Gen. Virol. *101*, 362–363.

Bergh, O., Borsheim, K.Y., Bratbak, G., and Heldal, M. (1989). High abundance of viruses found in aquatic environments. Nature *340*.

Bernheim, A., Millman, A., Ofir, G., Meitav, G., Avraham, C., Shomar, H., Rosenberg, M.M., Tal, N., Melamed, S., Amitai, G., et al. (2021). Prokaryotic viperins produce diverse antiviral molecules. Nature *589*, 120–124.

Blum, M., Chang, H.-Y., Chuguransky, S., Grego, T., Kandasaamy, 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*, 344–354.

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. *2*, e00040-17.

Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Hir, L., 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., et al. (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, 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*, e00054-19.

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., 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.-È., Villion, M., Magadán, A.H., and Moineau, S. (2013). CRISPR-Cas and restriction–modification systems are compatible and increase phage resistance. Nat. Commun. 1–7.

Ehrlich, S.D. (1977). Replication and expression of plasmids from Staphylococcus aureus in Bacillus subtilis. Proc. Natl. Acad. Sci. U. S. A. *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 Quidividi, Terranova, and Twillingate. Microbiol. Resour. Announc. *8*, e00598-19.

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 Meth *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., Truncait, L., Morgan, R.D., Ivanov, 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. *108*, 21218–21222.

He, X., Byrd, A.K., Yun, M., Iv, C.W.P., 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., Bratovič, M., Ressel, S., and Charpentier, E. (2018). The Biology of CRISPR-Cas: Backward and Forward. Cell *172*, 1239–1259.

Hoikkala, V., Ravantti, J., Díez-villaseñor, C., Tiirola, 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*, e03338-20.

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.

Kazlauskiene, 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. *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.

Klumpp, 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., 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.

Li, M., Kao, E., Gao, X., Sandig, H., Limmer, K., Pavon-Eternod, M., Jones, T.E., Landry, S., Pan, T., Weitzman, M.D., et al. (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., 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.

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.

Marraffini, L.A., and Sontheimer, E.J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science *322*, 1843–1845.

Meeske, A.J., Nakandakari-Higa, S., and Marraffini, L.A. (2019). Cas13-induced cellular dormancy prevents the rise of CRISPR-resistant bacteriophage. Nature *570*, 241–245.

Millman, A., Bernheim, A., Stokar-Avihail, A., Fedorenko, T., Voichek, M., Leavitt, A., Oppenheimer-Shaanan, Y., and Sorek, R. (2020). Article Bacterial Retrons Function In Anti-Phage Defense Bacterial Retrons Function In Anti-Phage Defense. Cell *183*, 1–11.

Narajczyk, M., Barańska, S., Węgrzyn, A., and Węgrzyn, G. (2007). Switch from θ to σ replication of bacteriophage λ DNA: factors involved in the process and a model for its regulatio.

Mol. Genet. Genomics *278*, 65–74.

Nguyen, L., Schmidt, H.A., Haeseler, A. Von, and Minh, B.Q. (2014). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Mol. Biol. Evol. *32*, 268–274.

Niewoehner, O., Garcia-doval, C., Rostøl, 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. (2017). 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*, 1–14.

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.

Rostøl, J.T., and Marraffini, L.A. (2019). Non-specific degradation of transcripts promotes plasmid clearance during type III-A CRISPR–Cas immunity. Nat. Micro *4*, 656–662.

Salas, M., Holguera, I., Redrejo-Rodríguez, M., and de Vega, M. (2016). DNA-Binding Proteins Essential for Protein-Primed Bacteriophage Φ29 DNA Replication. Front. Mol. Biosci. *3*, 1–21.

Samai, P., Pyenson, N., Jiang, W., Goldberg, G.W., Hatoum-Aslan, A., and Marraffini, L.A. (2015). Co-transcriptional DNA and RNA Cleavage during 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.

Shereda, R.D., Kozlov, A.G., Lohman, T.M., Cox, M.M., Keck, J.L., Shereda, R.D., Kozlov, A.G.,

Lohman, T.M., Shereda, R.D., Kozlov, A.G., et al. (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.

Steczkiewicz, 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.

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 *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*, e00104-19.

Vandersteegen, K., Mattheus, W., Ceyssens, P., Bilocq, F., Vos, D. De, Pirnay, J., Noben, J., Merabishvili, M., Lipinska, U., Hermans, K., et al. (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.

Walker, F.C., Chou-Zheng, L., Dunkle, J.A., and Hatoum-Aslan, A. (2017). Molecular determinants for CRISPR RNA maturation in the Cas10 – Csm complex and roles for non-Cas nucleases. Nucleic Acids Res. *45*, 2112–2123.

Watson, B.N.J., Vercoe, R.B., Salmond, G.P.C., Westra, E.R., Staals, R.H.J., and Fineran, P.C.

(2019). Type I-F CRISPR-Cas resistance against virulent phages results in abortive infection and provides population-level immunity. Nat. Commun. *10*, 5526.

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*, 1–13.

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., 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.

Supplementary Figures

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

S. M. Nayeemul Bari[1§], Lucy Chou-Zheng[1§], Olivia Howell[1], Katie Cater[2], Vidya Sree Dandu[2], Alexander Thomas[2], Barbaros Aslan[1], and Asma Hatoum-Aslan[1*]

[1] Department of Microbiology, University of Illinois, Urbana, IL 61821, USA

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

[§] These authors contributed equally

*To whom correspondence should be sent: ahatoum@illinois.edu

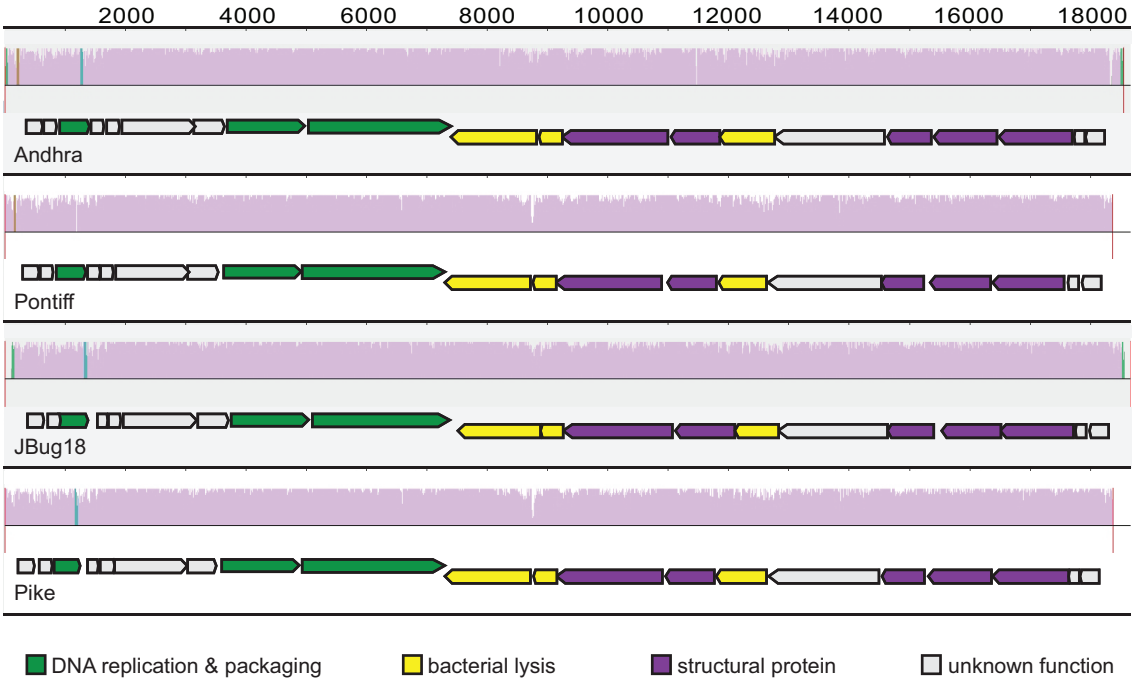


Figure S1. To accompany Figure 1. Four related *Podoviridae* phages with different host ranges. Shown is a multiple genome alignment of *S. epidermidis* podophages Andhra, Pontiff, JBug18, and Pike. Genome coordinates are shown on top, and colored histograms indicate the nucleotide similarity at each 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.

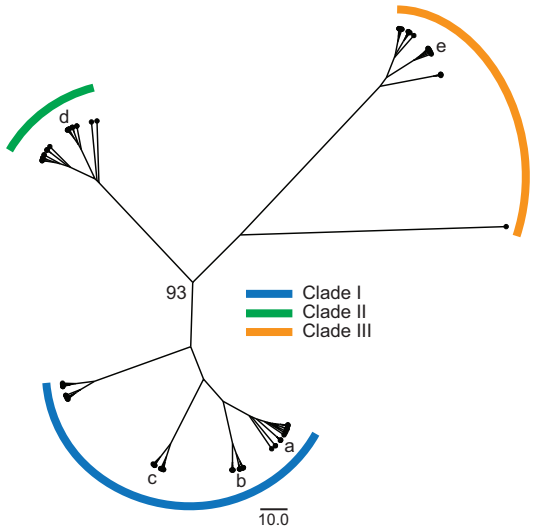


Figure S2. To accompany Figure 2. A dendrogram generated for SERP2475 and 99 homologs. Positions of homologs selected for functional characterization are indicated with lower-case letters: a. WP_045177897 from *S. aureus* MJ163; b. WP_002489608 (i.e. SERP2475) from *S. epidermidis* RP62a; c. WP_000632676 from *S. aureus* CA-347; d. WP_013870910 from *Lacinutrix* 5H-3-7-4; e. WP_101958732 from *V. vulnificus* FORC54.

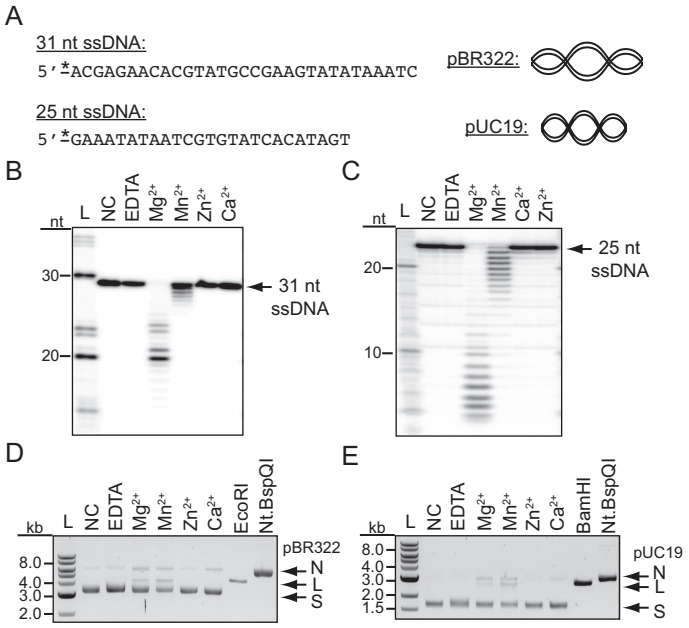


Figure S3. To accompany Figure 5. Activity of SERP2475 N-terminal 230 amino acids in the presence of different metals and substrates. (**A**) Illustration of different substrates used for nuclease assays. Asterisks indicate the position of the radiolabel on the 5'-end of linear substrates. Exonuclease assays (**B** and **C**) and nickase assays (**D** and **E**) are shown in which various 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 (for exonuclease assays) or 1 h (for nickase assays). Shown are representative images for three independent trials.

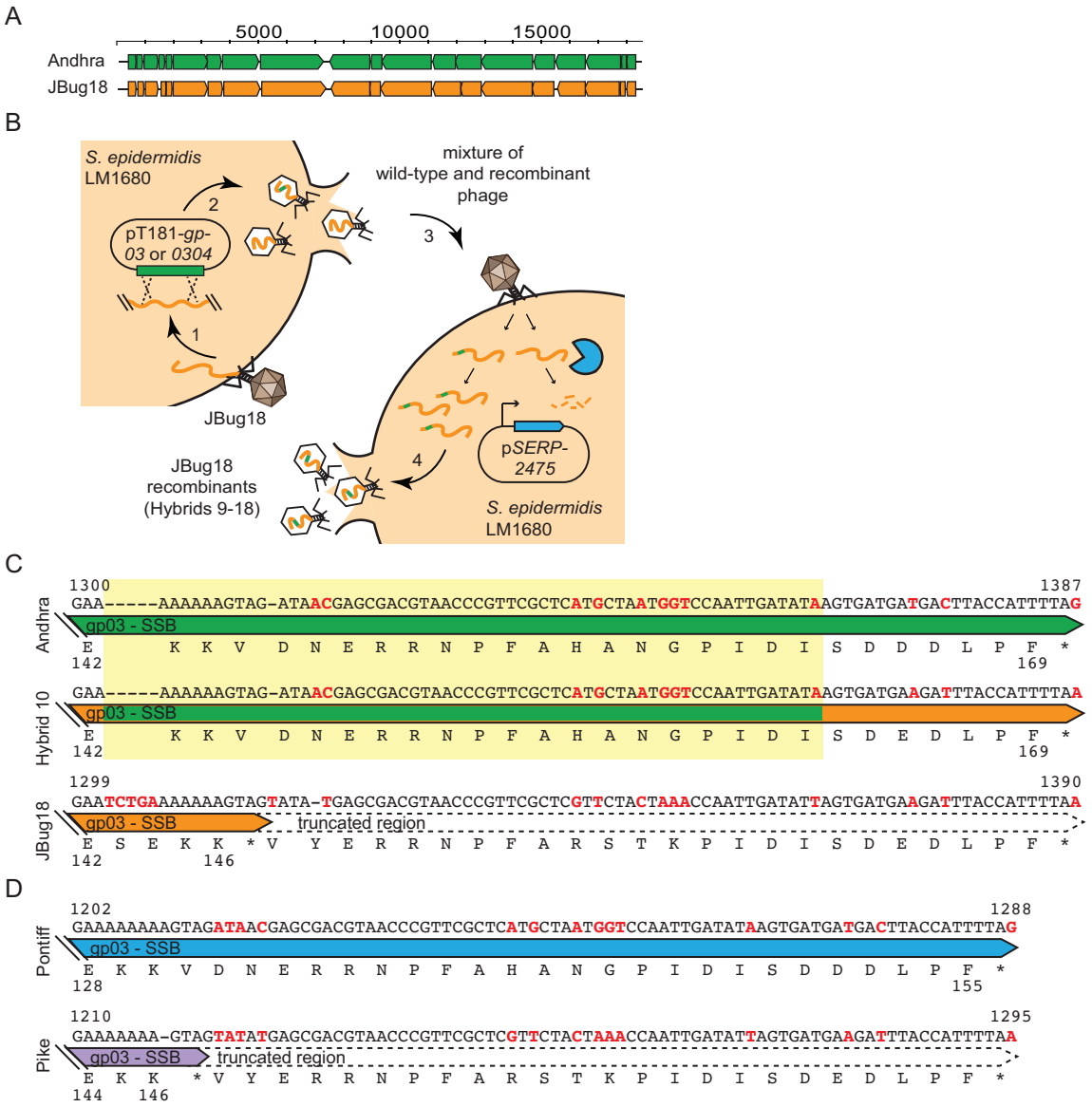


Figure S4. Accompanies Fig. 6. Generation of Hybrids 9-18 and comparison of SSB sequences. (**A**) A pairwise comparison of the open reading frames of Andhra and JBug18. (**B**) A diagram of the method used to generate JBug18-Andhra Hybrids 9-18. (**C**) Sequence comparison between the SSBs of Andhra, 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.