Communication between viruses guides lysis-lysogeny decisions

Zohar Erez[#1], **Ida Steinberger-Levy**[#1,2], **Maya Shamir**[1], **Shany Doron**[1], **Avigail Stokar-Avihail**[1], **Yoav Peleg**[3], **Sarah Melamed**[1], **Azita Leavitt**[1], **Alon Savidor**[4], **Shira Albeck**[3], **Gil Amitai**[1,#], and **Rotem Sorek**[1,#]

[1]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 7610001, Israel

[2]Department of Biochemistry and Molecular Genetics, Israel Institute for Biological Research, Ness-Ziona, Israel

[3]Israel Structural Proteomics Center (ISPC), Faculty of Biochemistry, Weizmann Institute of Science, Rehovot 7610001, Israel

[4]de Botton Institute for Protein Profiling, The Nancy and Stephen Grand Israel National Center for Personalized Medicine, Weizmann Institute of Science, Rehovot 7610001, Israel

[#] These authors contributed equally to this work.

Abstract

Temperate viruses can become dormant in their host cells, a process called lysogeny. In every infection, such viruses need to decide between the lytic and the lysogenic cycles, i.e., whether to replicate and lyse their host or to lysogenize and keep the host viable. Here we show that viruses (phages) of the spBeta group use a small-molecule communication system to coordinate lysis-lysogeny decisions. During infection of its *Bacillus* host cell, the phage produces a 6aa communication peptide that is released to the medium. In subsequent infections, progeny phages measure the concentration of this peptide and lysogenize if the concentration is sufficiently high. We found that different phages encode different versions of the communication peptide, demonstrating a phage-specific peptide communication code for lysogeny decisions. We termed this communication system the "*arbitrium*" system, and further show that it is encoded by 3 phage genes: *aimP*, producing the peptide, *aimR*, the intracellular peptide receptor, and *aimX*, a negative regulator of lysogeny. The *arbitrium* system enables an offspring phage to communicate with its predecessors, i.e., to estimate the amount of recent prior infections and hence decide whether to employ the lytic or lysogenic cycle.

Introduction

Temperate phages may choose to infect either through the lytic or the lysogenic cycles[1]. Whereas the lytic cycle leads to lysis of the bacterial cell, in the lysogenic cycle the phage genome integrates into the bacterial genome, and the lysogenized bacterium becomes immune to further infection by the same phage[1]. In accordance, growth dynamics of bacteria infected by temperate phages presents partial, but not full, lysis of the culture, followed by culture recovery due to growth of lysogenized bacteria.

The factors influencing lysis-lysogeny decisions were so far studied in depth mainly for the phage lambda, infecting *Escherichia coli*[2]. The decision was shown to be probabilistic in nature, but influenced from the nutritional state of the infected bacterium as well as by the number of co-infecting phage particles[2,3]. In the current study we report that phages infecting *Bacillus* species can rely on small molecule communication to execute lysis-lysogeny decisions.

Results

Conditioned media promote phi3T lysogeny

We initiated this study by attempting to test the hypothesis, which was not verified here eventually, that bacteria secrete communication molecules to alert other bacteria from phage infection. For this, we prepared conditioned media from cultures of *Bacillus subtilis* str. 168 infected by four different phages: phi29, phi105, rho14, and phi3T (Fig 1A). To prepare these conditioned media we infected bacteria growing at mid-log phase with the phage at a multiplicity of infection (MOI) of 1, and collected the conditioned media three hours following infection (or without infection, for control media). The collected media were finely filtered to retrieve small molecules and eliminate remaining bacteria and phages, and a fresh batch of bacteria was added to the conditioned and control media and subsequently infected by the same phage (Fig 1A).

For one of the phages we tested, phi3T, the infection dynamics in the conditioned medium, as inferred from the bacterial growth curve, was dramatically different than the dynamics in the control medium: whereas a substantial fraction of the infected bacterial culture had lysed in the control medium two hours post infection, the culture grown in the conditioned medium appeared to be much more protected from lysis (Fig 1B). This effect was not detected for any of the other three phages we tested, for which we did not observe a difference in infection dynamics between the control and conditioned media (Extended data fig 1). Moreover, the conditioned medium prepared from phi3T infection did not affect the infection dynamics of other phages, and *vice versa*, conditioned media prepared from other phages did not affect the infection dynamics of phi3T (Extended data fig 1). These results imply that a small molecule is released to the medium during infection of *B. subtilis* by phi3T and this molecule can affect infection dynamics of downstream infections by this phage.

Quorum sensing (QS) in Bacilli and other Firmicutes is typically based on short peptides that are secreted into the medium and sensed by intra-cellular or membrane bound

receptors4–6. We therefore examined whether the active substance in the medium is proteinaceous by treating the conditioned medium with proteinase K. Infection dynamics in the proteinase-treated conditioned medium were similar to the dynamics in the control medium, suggesting that the active component in the medium is indeed proteinaceous (Extended data fig 2). Since communication peptides in *Bacillus* quorum sensing systems are frequently imported into the cell by the oligopeptide permease transporter (OPP), we tested phage infection dynamics in bacteria in which an essential subunit of the OPP transporter, *oppD*, was deleted (Fig 1C). The phage-derived conditioned medium lost its effect when the bacteria lacking the functional OPP were infected by phi3T, suggesting that the active substance in the conditioned medium is a 3aa-20aa long peptide, which is the size range of peptides that can be imported by the OPP transporter of Gram positive bacteria 7.

Examination of the phage infection dynamics in the *oppD* mutant showed increased culture lysis in both control and conditioned media as compared to infection of WT bacteria (Fig 1C). Since phage phi3T is a temperate phage 1,8, the increased lysis observed in the infection dynamics curve of the *oppD* mutant was suggestive that the active peptide released to the medium may promote lysogeny of the phage. To check this hypothesis we examined phage phi3T integration into the *B. subtilis* genome during infection using a semi-quantitative PCR assay. Indeed, increased lysogeny was observed when the bacterial culture was infected in the conditioned medium (Fig 1D). Combined, these results imply that during phi3T infection a short peptide is released to the medium and, as this peptide accumulates, it acts as a communication agent affecting the lysis/lysogeny decision of later generations of the phage progeny. We named the putative communication molecule *arbitrium* (the Latin word for "decision").

Phage-produced peptide induces lysogeny

Phi3T was isolated 4 decades ago and was characterized as belonging to the spBeta group of phages, although to date its genome has not been sequenced8. To search for the possible genetic system encoding the *arbitrium* peptide, we sequenced and analyzed the genome of phi3T, which assembled into a single 128 kbps contig containing 201 predicted genes, 128 (64%) of which show significant homology to spBeta. We found three ORFs that were predicted to have an N-terminal signal peptide, suggesting that they are secreted or membrane-localized. While two of these genes seemed irrelevant (one integral membrane protein and the other was a large nuclease), the third gene exhibited features reminiscent of *Bacillus* quorum sensing peptides (Fig 2A). Peptides belonging to the Phr family of quorum sensing systems in *B. subtilis* are typically processed from a pre-pro-peptide that contains an N-terminal signal sequence, which is recognized by the Sec system and cleaved upon secretion4. Once outside the cell, the pro-peptide is further processed by *B. subtilis* extracellular proteases to produce the mature short (5-6 aa) peptide that is typically found on the C-terminal end of the pro-peptide4. Our candidate gene encoded a short ORF (43 aa), and displayed both an N-terminal signal sequence and the consensus cleavage site for peptide maturation at its C-terminus (Fig 2A, Extended data table 1). If this phi3T-encoded protein is secreted and matured extracellularly, then the predicted mature communication peptide after pro-peptide cleavage would be Ser-Ala-Ile-Arg-Gly-Ala (SAIRGA). Indeed, mass spectrometry analysis confirmed the presence of the SAIRGA peptide in the

conditioned medium but not in the control medium or in medium derived from phi3T in which the gene encoding the peptide was silenced (Extended data fig 3).

To test whether the predicted mature peptide is indeed the *arbitrium* molecule that influences the phage lysogeny decision, we infected bacteria with phi3T in LB medium supplemented with increasing amounts of synthesized SAIRGA peptide. A clear concentration-dependent effect on the phage infection dynamics was observed, such that reduced culture lysis was apparent when the medium contained higher concentrations of the synthesized peptide (Fig 2B). These effects were specific to that peptide, and were neither observed for shorter versions of the peptide (SAIRG or AIRGA), nor for PhrC, a known quorum sensing peptide of *B. subtilis* (Fig 2B, Extended data fig 4). The maximal effect on the culture growth curve was observed at SAIRGA peptide concentration of 500 nM, above which the effect seemed saturated (Fig 2B).

To verify that the observed effect of the SAIRGA peptide on the dynamics of the infected culture was the result of increased lysogeny, we directly sequenced total DNA of bacteria collected from a time course experiment during infection by phi3T with and without the peptide. By comparing the fraction of sequencing reads passing through the intact phage integration site in the bacterial genome to reads demonstrating phage integration at that site, we directly quantified the fraction of lysogenized bacteria at each time point. We found consistently elevated lysogeny in the presence of the SAIRGA peptide, such that 48% (+/-7.9%) of the bacteria were lysogenized at 60 minutes post infection, as compared to 18% (+/-3.3%) of bacteria grown without the addition of the synthesized peptide (Fig 2C). These results imply that the phage-encoded gene we identified is secreted and processed into the mature *arbitrium* communication peptide that further affects the phage lysis/lysogeny decision. We denoted this gene *aimP*.

To examine the dynamics of *arbitrium* peptide accumulation during the course of phage infection, we infected bacteria with phi3T in low MOI (1:1000 phage:bacteria), and recorded the peptide concentrations in the medium using targeted mass spectrometry during several cycles of phage replication, in parallel to recording phage and bacterial counts. Our results show that the peptide accumulates to detectable quantities after 3 cycles of phage infection (t=120), and significantly accumulates after 4 infection cycles when the phage:bacteria ratios reach ~1:1 (Figure 3). After 150-180 minutes the peptide reaches concentrations that are effective in lysogeny promotion as measured in Figure 2B.

The *aimP* gene is located immediately downstream of a gene encoding a 378 aa long protein, suggesting that these two genes may be functionally linked. This upstream gene encodes a predicted tetratricopeptide repeat (TPR) domain, typical of intracellular peptide receptors of the RRNPP family in QS systems of Gram positive bacteria[9–11] (Fig 2A). We therefore hypothesized that this upstream gene, which we denoted *aimR*, is the receptor of the AimP-derived *arbitrium* peptide. To test this hypothesis we purified a C-terminal His-tagged AimR, and used microscale thermophoresis (MST) to measure the binding between the purified receptor and the synthesized *arbitrium* peptide. This analysis showed high-affinity binding, at an effective peptide concentration of EC_{50}=138 nM (118-162 nM at confidence interval of 95%), between the phi3T AimR receptor and the cognate SAIRGA peptide (Fig 2D),

confirming that AimR most probably functions as the intracellular receptor of the *arbitrium* SAIRGA peptide.

Phage-specific communication code

To appreciate the abundance of this system in nature, we used a homology search to find homologs of the *aimR* gene in available sequenced genomes. We found 112 instances of AimR homologs, virtually all of them in *Bacillus* phages or in prophages found integrated within Bacilli genomes, suggesting that this gene primarily fulfills a phage-related function (Fig 4A; Supplementary table 1). In all cases, *aimR* homologs were found upstream of *aimP* candidate genes, i.e., short polypeptides encoding an N-terminal signal peptide, followed by a pro-peptide conforming with the processing maturation signal of the *Bacillus* extracellular proteases (Supplementary table 1; Extended data table 1). Although the sequences of the predicted mature peptides were diverse, all of them maintained strict rules for their sequence composition, with an obligatory glycine residue at the 5th position, glycine or alanine at the 6th position, and a preference for positively charged residue at the 4th position (Fig 4B, 4C). While most (72%) of the phages harboring homologs of the *arbitrium* system belonged to the spBeta group, *arbitrium* was also found in other types of phages, including phi105-like and Mu-like phages (Supplementary table 1).

To test the hypothesis that the phage-encoded communication peptides guide phage lysogeny in a sequence-specific manner, we examined the infection dynamics of the spBeta phage, which is heteroimmune relative to phi3T, and in which we identified a homolog of the AimR-AimP system. The predicted mature AimP-derived *arbitrium* peptide of spBeta was GMPRGA, a sequence that differs by the 3 N-terminal amino acids from the SAIRGA peptide of phi3T. Whereas the GMPRGA peptide promoted lysogeny of spBeta, it did not affect the lysogeny profile of phi3T; and similarly, the phi3T-derived SAIRGA peptide had no effect on the infection dynamics of spBeta (Figs 4D-4F). In accordance, the spBeta-derived GMPRGA peptide did not show specific binding to the phi3T AimR receptor (Fig 2D). These results demonstrate a sequence-specific peptide code that guides phage lysogeny in a phage-specific manner.

Mechanism of the *arbitrium* system

In communication systems of Gram positive bacteria, the binding of the communication peptide to its receptor usually leads to reprograming of the transcriptional response. This can occur either directly, when the receptor is a transcription regulator such as in the cases of the PrgX12,13 in Enterococci, the PlcR14–16 of the *Bacillus cereus* group, and in other systems17,18; or indirectly, as in the case of Rap/Phr systems of Bacilli, in which the receptor is a phosphatase that regulates downstream transcriptional regulators by dephosphorylation19,20, or steric interference21. The presence of a predicted helix-turn-helix (HTH) motif in the N-terminus of AimR suggested that the receptor of the *arbitrium* system directly binds DNA. To examine whether AimR binds the phage DNA *in vivo*, we engineered a His-tagged *aimR* gene into a *B. subtilis* 168 strain in which a dCas9 (CRISPRi) technology22 was used to silence the expression of the phage AimR gene, but not the cloned His-tagged AimR (Methods). We then performed a ChIP-seq assay 15 minutes after phage infection with and without the presence of the *arbitrium* peptide. Sequencing of the DNA

bound to AimR clearly showed that AimR binds a single site in the phage genome, directly downstream of the *aimP* gene (Fig 5A, 5B). Moreover, this binding only occurred when the *arbitrium* peptide was lacking from the medium, suggesting that binding of the *arbitrium* peptide to its AimR receptor leads to dissociation of the receptor from its binding site on the phage DNA.

During the process of AimR purification we noticed that the protein migrates as homodimer in a gel filtration column. Upon addition of the phi3T-derived SAIRGA peptide, however, the protein strictly migrated as a monomer (Fig 5C). This observation was corroborated using a cross linking assay (Extended data fig 5). These results suggest that the *arbitrium* peptide transfers the signal via alteration of the oligomeric state of its AimR receptor from a DNA binding dimer to a peptide-bound, dissociated monomer. Addition of the spBeta GMPRGA peptide did not lead to a change in the AimR oligomeric state, pointing, again, to the high specificity between the peptide and its receptor in the *arbitrium* system (Fig 5C, Extended data fig 5).

To examine whether binding of the *arbitrium* peptide to its AimR receptor leads to a transcriptional response in the phage genome we applied RNA-seq to RNA extracted from bacteria during a time course of infection with and without the peptide. The most dramatic change in the expression was observed for a single transcript, which we denoted *aimX*, that was immediately downstream to the AimR DNA binding site (Fig 5D-G, Extended data fig 6). This transcript showed substantial expression in the absence of the *arbitrium* peptide starting 10 minutes after infection, but its expression was reduced more than 20 fold when the medium was supplemented by 1 μM of the SAIRGA peptide (Fig 5D-G).

The above results suggest that AimR, when bound to the phage DNA as a dimer in the absence of the *arbitrium* peptide, is a transcriptional activator of AimX. Indeed, when AimR was silenced using dCas9, the expression of AimX was dramatically reduced (Fig 5D). Moreover, silencing of AimR resulted in increased lysogeny, suggesting that binding of AimR to the phage DNA inhibits lysogeny (or promotes lysis), possibly by activating the expression of AimX (Fig 5H).

Since the AimR knockdown did not lead to a dramatic transcriptional effect for any transcript except AimX at 20 minutes post infection (Extended data fig 6), we hypothesized that the main function of AimR is to control the expression of *aimX*, and that the expressed AimX works downstream to the AimR-AimP communication system to execute the lysis-lysogeny decision. Consistent with this hypothesis, knockdown of AimX using dCas9 resulted in increased lysogeny (Fig 5H). Moreover, complementing with an ectopic AimX on the background of AimR knockdown (in which AimX expression is naturally silenced, Fig 5D) resulted in culture lysis, demonstrating that AimX functions downstream to AimR, and works as the inhibitor of the lysogenic or the promoter of the lytic cycle (Fig 5H).

The *aimX* transcript encodes a short (51 aa) ORF, followed by an unusually long stretch of additional 62 bases of non-coding RNA region that terminates in a stem-loop structure. A homolog of this ORF follows the *aimR-aimP* operon only in 17 of the 112 phages in which the *arbitrium* system was identified (Supplementary table 1). We therefore examined the

expression of the *arbitrium* locus in the spBeta phage, in which the AimX ORF is absent. We found that although no detectable ORF was present in the region downstream *aimP*, a non-coding RNA (ncRNA) was expressed from that intergenic region in spBeta (Extended data fig 7). Moreover, this ncRNA was responsive to the arbitrium peptide such that in the presence of the spBeta *arbitrium*, its transcription was abolished, akin to AimX in phage phi3T (Extended data fig 7). These results imply that AimX possibly exerts its function as a regulatory non-coding RNA in a manner that remains to be clarified in future studies.

Combined together, our results point to the following model: upon initial infection of a bacterial culture, phi3T expresses the early operon AimR-AimP. AimR, as a dimer, activates the expression of AimX, which, in turn, blocks the pathway to lysogeny and/or promotes the lytic cycle in a manner yet unknown. At the same time, AimP is secreted into the medium and processed into the mature *arbitrium* peptide. After several cycles of infection, *arbitrium* peptides will accumulate in the medium. Now, when a phage particle infects a yet-uninfected bacterium, the concentration of the *arbitrium* peptide, which is internalized into the bacteria by the OPP transporter, will be high enough to bind the AimR receptor. Upon binding, the AimR receptor changes it oligomeric state from the active dimer to the inactive monomer, silencing the AimX lysogeny-inhibitor and leading to lysogeny (Figure 6).

Discussion

We have shown that phages belonging to the spBeta group of phages use communication peptides in order to decide whether to enter a lytic cycle or lysogenize the infected bacterium. In a sense, the communication mechanism we described allows an offspring phage to communicate with its ancestors, i.e., measure the amount of the predecessor phages that completed successful infections in prior cycles. The biological logic behind this strategy is clear: when a single phage encounters a bacterial colony, there is ample prey for the progeny phages that are produced from the first cycles of infection, and hence a lytic cycle is preferred. In later stages of the infection dynamics the number of bacterial cells is reduced to a point that progeny phages are at risk of no longer having a new host to infect. Then, it is logical for the phage to switch into lysogeny to preserve chances for viable reproduction. The *arbitrium* system provides an elegant mechanism for a phage particle to estimate the amount of recent prior infections and hence decide whether to employ the lytic or lysogenic cycle.

Although the presence of the *arbitrium* peptide in the medium has a dramatic effect on phage lysogeny (Fig 2B, 2C) this effect is not absolute: even at maximal peptide concentrations only ~50% of the cells become lysogenized after 60 minutes, indicating that probability of lysogeny significantly rises, but is still a stochastic event (Fig 2C). Similarly, in the absence of the peptide, the phage does not become obligatorily lytic, as observed from an experiment with phi3T in which *aimP* was deleted (Extended data fig 8). Therefore, there seems to be a basal, possibly stochastic, tendency of the phage to lysogenize independently, which is supplemented by the *arbitrium* system that allows reference to prior infections.

Peptide communication systems are known to exist on conjugative plasmids both in Bacilli[23,24] and in Enterococci[12,13], where they can regulate the plasmids horizontal

transfer. The presence of bacterial quorum sensing systems on phage genomes was noted previously[4,25], but their roles were either unknown or linked to interference with the host bacterial communication[25]. To our knowledge this study is the first demonstration of actual small-molecule communication between viruses. We found this communication to be manifested through the *arbitrium* system in a large group of *Bacillus* phages; we envision that such a strategy will also be discovered for other phages utilizing different communication systems. Moreover, this strategy may not be limited to phages, and possibly also guide dormancy/lysis decisions in viruses infecting eukaryotes.
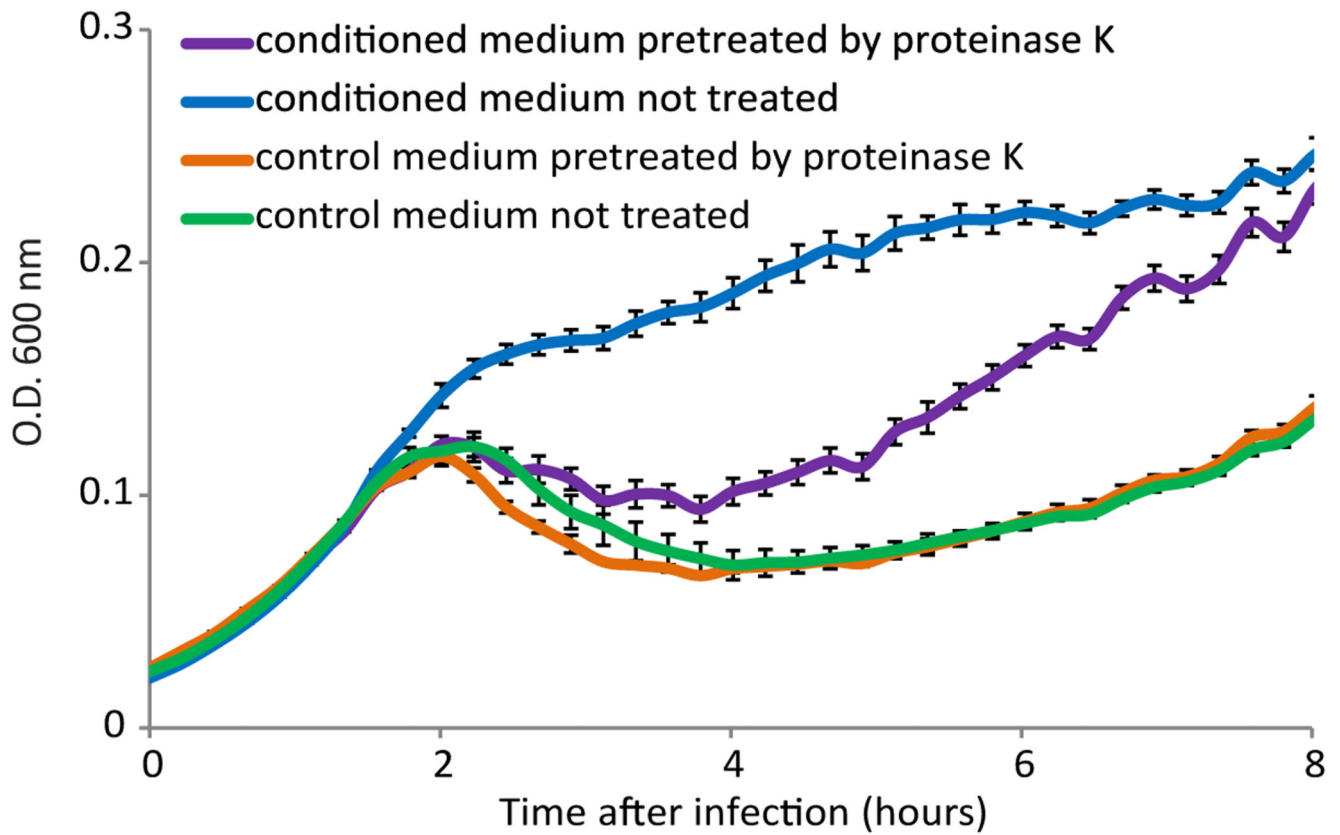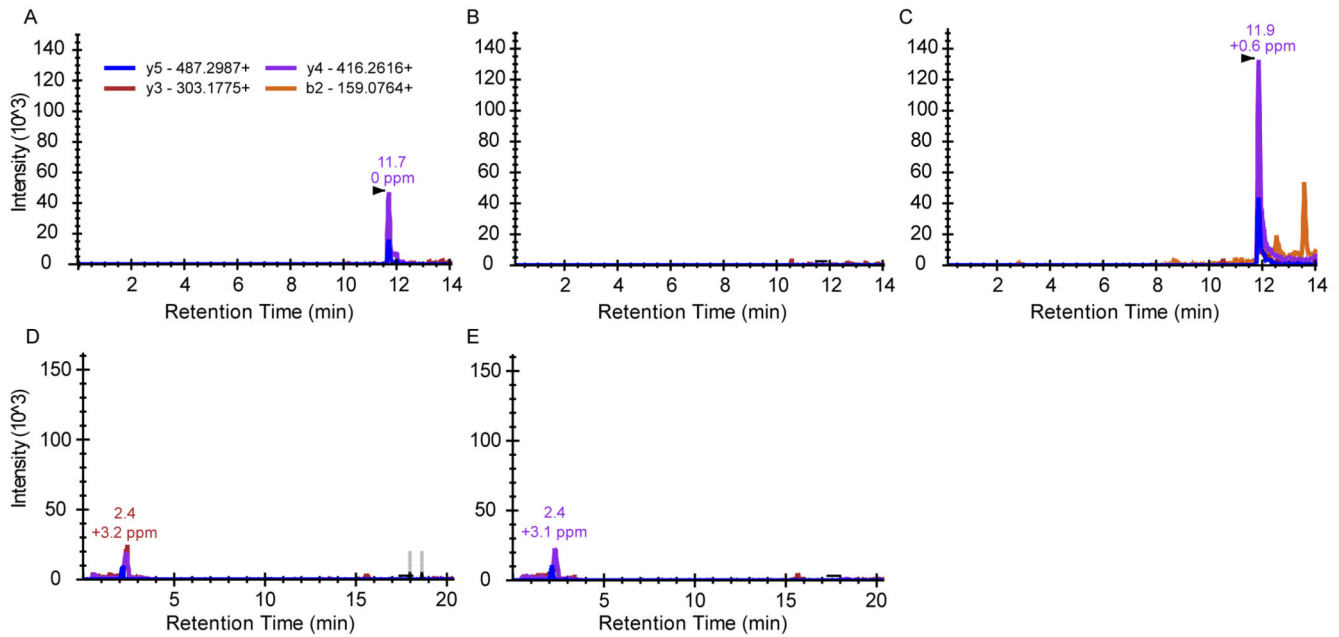
Extended Data
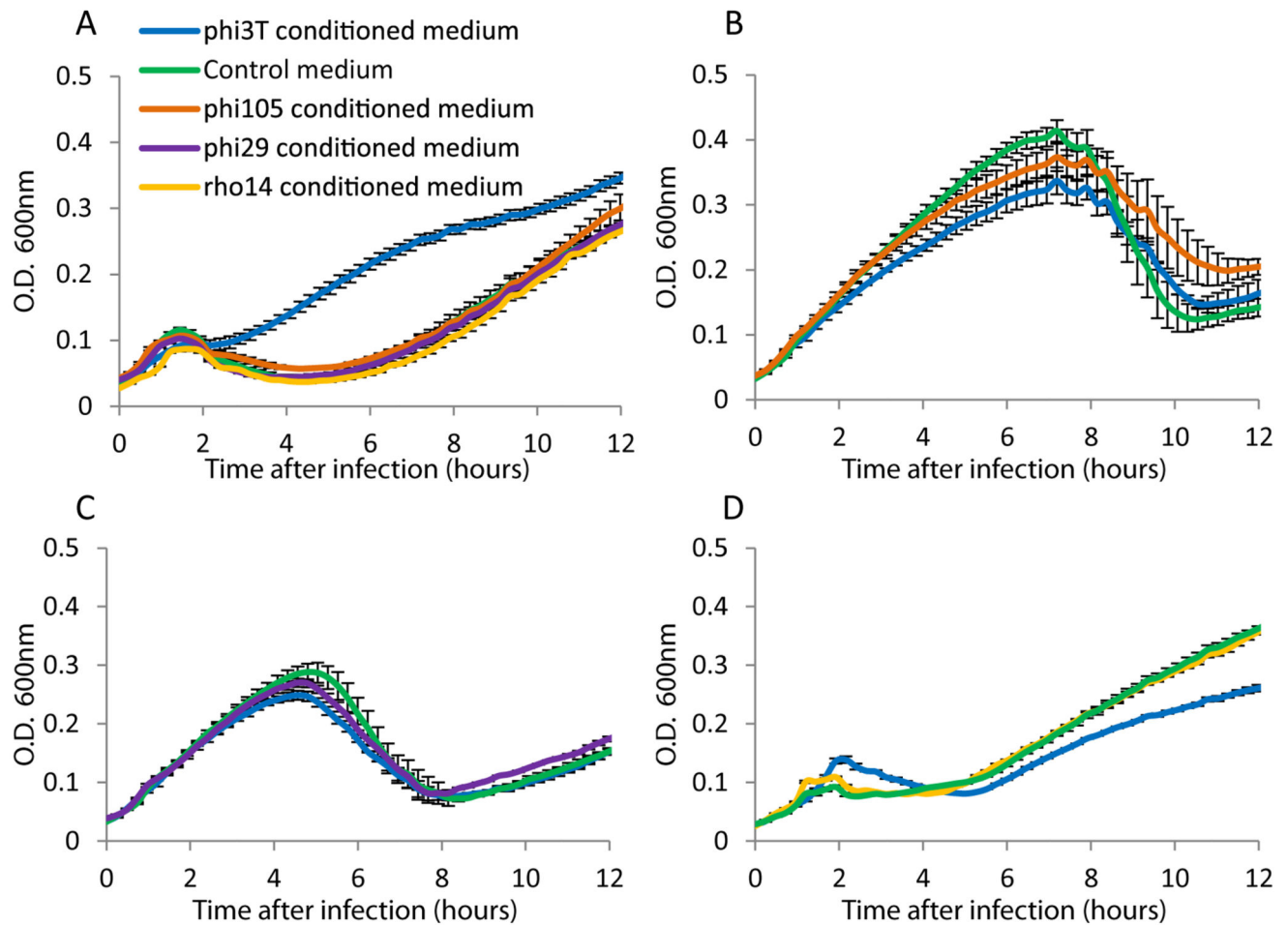
Extended data figure 1. Specificity of conditioned media.
Conditioned media were prepared using initial infection of *B. subtilis* 168 by four phages: phi3T, phi105, phi29 and rho14. Presented are growth curves of *B. subtilis* 168 in the different media, infected with each phage at MOI=0.1. (A) Infection with phi3T. (B) Infection with phi105. (C) Infection with phi29. (D) Infection with rho14. Data represent average of 3 replicates, and error bars represent standard error.

Extended data figure 2. Proteinase K treatment reduces the effect of conditioned medium.
Growth curves of *B. subtilis* 168 infected by phi3T at MOI=0.1, in control and conditioned media, with and without pre-treatment with proteinase K. Data represent average of 3 technical replicates, and error bars represent standard error.

Extended data figure 3. Mass spectrometry verifies the presence of the SAIRGA peptide in conditioned medium.

Presented are extracted ion chromatograms of targeted MS analysis experiment. B2, Y3, Y4 and Y5 refer to fragmentation products of the peptides made by the instrument in the MS/MS process, with the expected m/z ion masses indicated. Y-axis represents the ion intensity of each fragment ion (arbitrary units). (A) Reference synthesized SAIRGA peptide at 100 nM concentration in LB. (B) Control medium from *B. subtilis* 168. (C) Conditioned medium from phi3T-infected *B. subtilis* 168. Arrowhead depicts the expected retention time of the SAIRGA peptide. (D) Control medium from *B. subtilis* 168 expressing dCas9 with a spacer targeting *aimP*. (F) Conditioned medium derived from phi3T-infected *B. subtilis* 168 expressing dCas9 with a spacer targeting *aimP*.
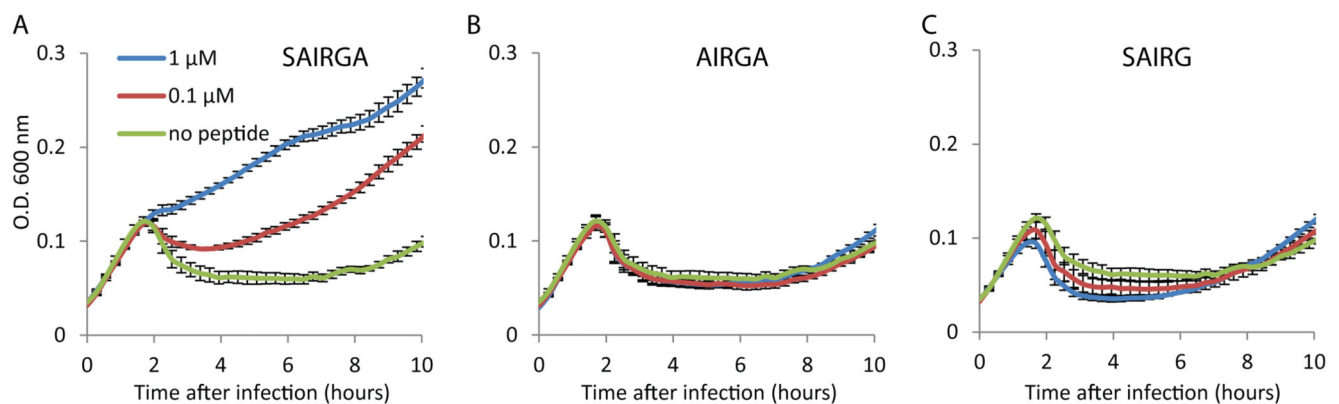

Extended data figure 4. 5aa versions of the 6aa *arbitrium* peptide do not guide lysogeny.

Growth curves of *B. subtilis* 168 infected by phi3T at MOI=0.1, in LB media supplemented with synthesized SAIRGA, AIRGA or SAIRG peptide. Shown is average of 3 biological replicates, each with 3 technical replicates. Error bars represent standard error.

Europe PMC Funders Author Manuscripts
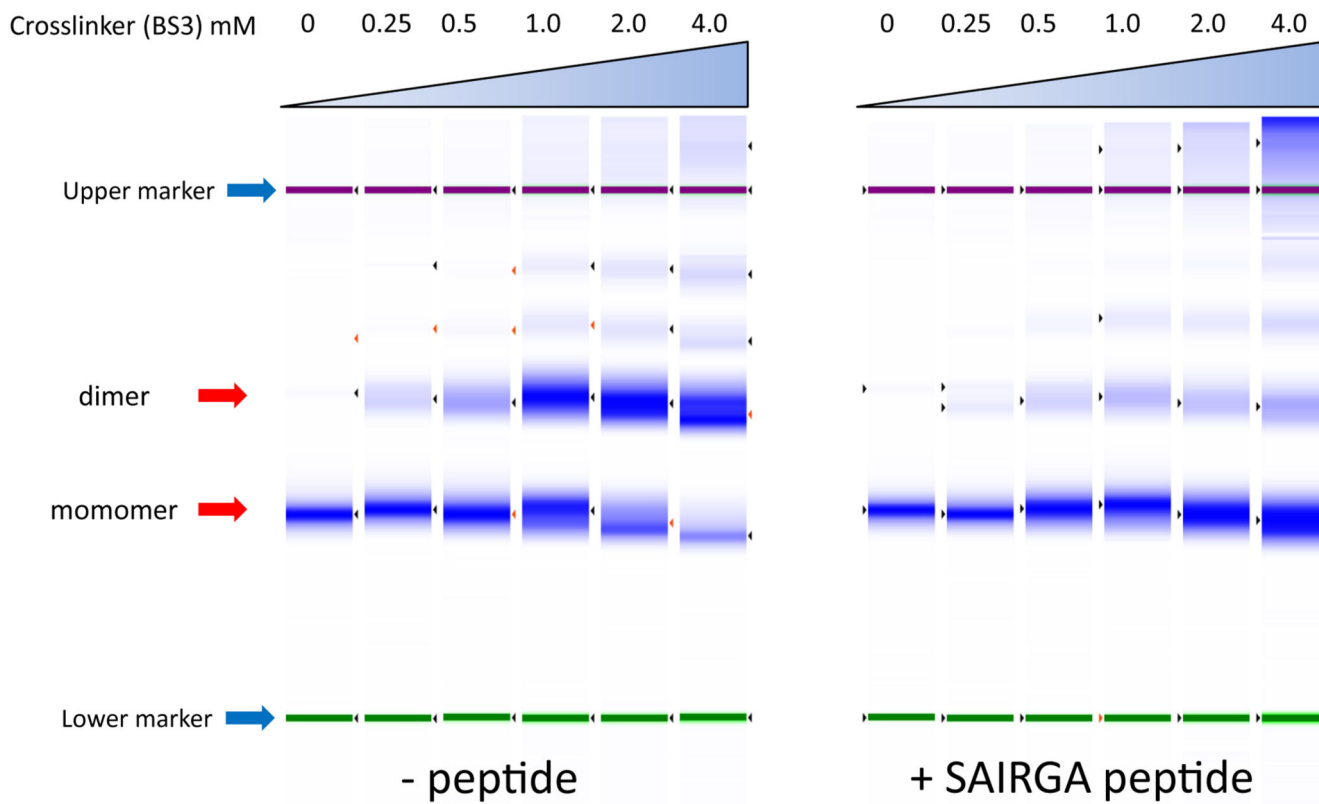
Extended data figure 5. Exposure of AimR to SAIRGA peptide reduces propensity for dimerization.

Purified AimR was eluted from a gel-filtration column, dialyzed, then mixed with SAIRGA peptide (f.c. 100 μM) dissolved in water or with equal amount of volume without peptide and incubated at room temperature for 5 minutes. The protein samples were then mixed for 30 minutes with different concentrations of the crosslinker BS3 bis(sulfosuccinimidyl)suberate. Presented are electrophoresis results analyzed using the TapeStation instrument (Agilent Technologies), showing that in the absence of peptide, purified AimR tends to preferentially be cross linked into dimers, whereas in the presence of the SAIRGA peptide, dimerization is significantly reduced.
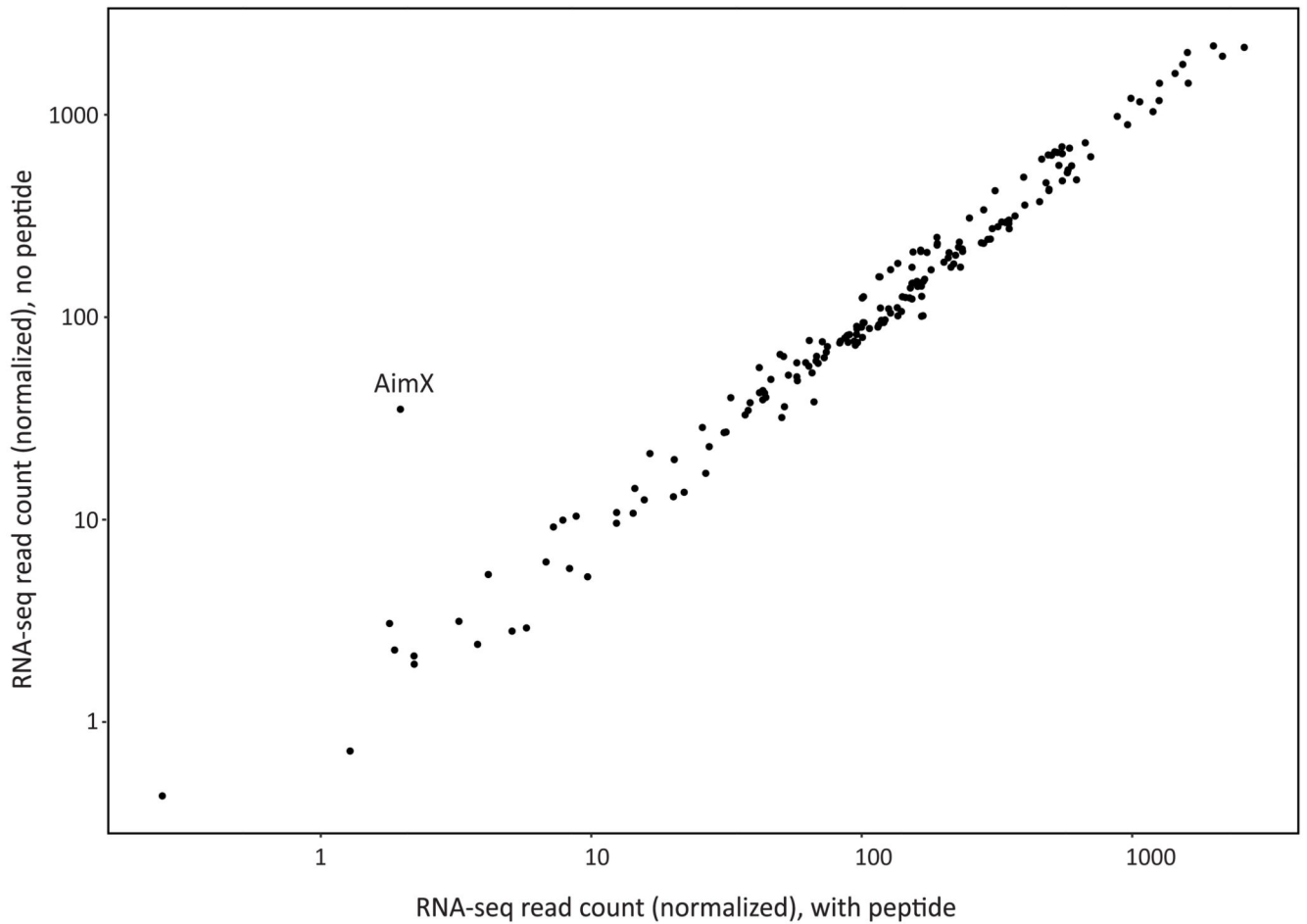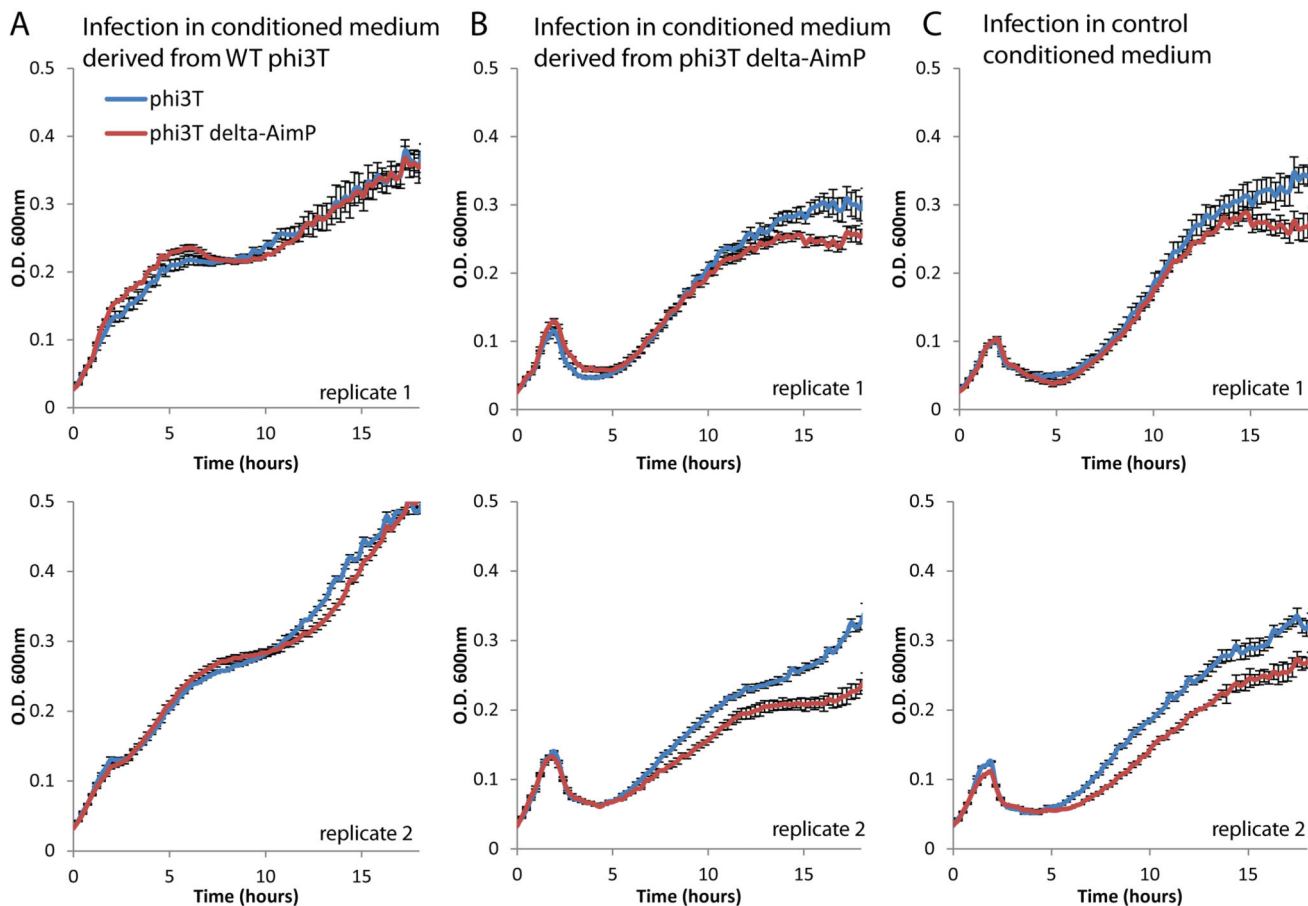
Extended data figure 6. Phage gene expression 20 minutes post infection.
Each dot represents a single phage gene. Axes represent average RNA-seq read count per gene from 3 replicates, after normalization to control for RNA-seq library size. X axis, expression when phage infection was in the presence of 1 μM of SAIRGA peptide; Y axis, no peptide.
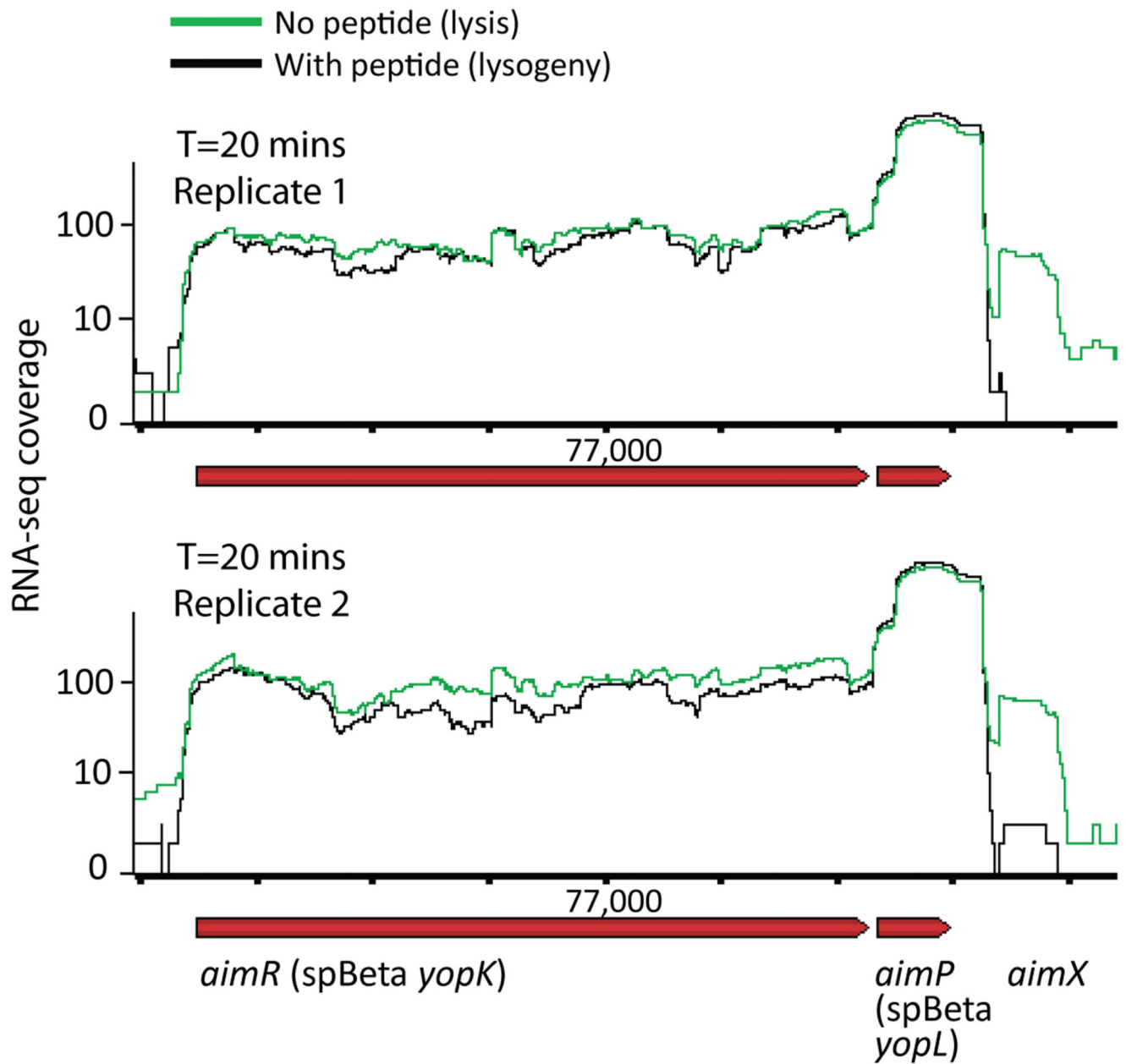
Extended data figure 7. Expression of the *arbitrium* locus in phage spBeta during infection.
RNA-seq coverage of the *arbitrium* locus at 20 minutes post infection, with (black) or without (green) 1 μM of synthesized GMPRGA peptide in the medium. RNA-seq coverage was normalized to the number of reads mapped to the phage genome in each RNA-seq library. Shown are two replicates of the experiment.

A Infection in conditioned medium derived from WT phi3T

B Infection in conditioned medium derived from phi3T delta-AimP

C Infection in control conditioned medium

Extended data figure 8. Infection experiments with phi3T delta-*aimP* and derived conditioned media.

Growth curves of *B. subtilis* 168 infected with phi3T or phi3T delta-*aimP* in (A) conditioned medium derived from phi3T; (B) conditioned medium derived from phi3T delta-*aimP*; and (C) control medium. Each medium was generated in biological replicate, and the two replicates are presented separately. Each curve represents the average of three technical replicates, and error bars represent standard error.

Extended data table 1
Signal for cleavage by extracellular proteases in the AimP pre-pro-peptide.

Shown are the amino acid sequences of *B. subtilis* quorum sensing Phr genes divided to their domains4. Recognition signal for *B. subtilis* extracellular proteases is in green. Signal peptide in the phi3T AimP protein was predicted by the SignalP 4.1 web server26.

	Signal peptide	Pro-peptide	Mature peptide
PhrA	MKSKWMSGLLLVAVGFSFTQVMVHA	GETANTEGK**TFHIA**	ARNQT
PhrC	MKLKSKLFVICLAAAIFTAAGVASANA	EAL**DFHVT**	ERGMT

	Signal peptide	Pro-peptide	Mature peptide
PhrF	MKLKSKLLLSCLALSTVGVATTIANA	PTH**QIEVA**	QRGMI
PhrG	MKRFLIGAGVAAVILSGWFIA	DHQTHSQ**EMKVA**	EKMIG
Phr-pTA1060	MKFKGLFSAVLIVSLLVGAGYSFV	HHD**EVSVA**	SRNAT
AimP	MKKVFFGLVILTALAISFVAGQQSVSTASA	SD**EVTVA**	SAIRGA

Extended Materials and Methods

Phages and strains in this study—*Bacillus subtilis* str. 168 was obtained from the Bacillus Genetic Stock Center (BGSC). *Bacillus subtilis* BEST7003 was obtained from Prof. Itaya Mitsuhiro at Keio University, Japan. The DS4979 strain (*oppD*::kan)27 was obtained from Dr. Avigdor Eldar at Tel Aviv University. The 3610 strain was obtained from Dr. Ilana Kolodkin-Gal at the Weizmann Institute. *Bacillus* phages phi3T, spBeta, phi105, rho14 and phi29 were all obtained from the BGSC.

Oligos and reagents—All oligos used in this study were purchased from either Sigma (St. Louis, Missouri) or Integrated DNA Technologies (IDT, San Jose, California). Native synthetic peptides were purchased from Peptide 2.0 Inc. (Chantilly, Virginia), at 98% purity, desalted. Isotopically labeled peptide was purchased from GenScript (Piscataway, New Jersey), purified using HPLC to 95% purity.

Phi3T genome sequencing—Phage DNA was extracted using the Qiagen DNeasy blood and tissue kit (CAT# 69504) and subjected to Illumina sequencing on a MiSeq machine with single read length of 50 bases. Genome was assembled using the Velvet software package28 with hash_length=31. Genome was assembled into 3 contigs that were further manually assembled based on end-homologies. Assembled genome was permuted to conform with the spBeta phage genome coordinates (accession NC_001884). ORFs were predicted using Glimmer29 with options -o50 -g110 -t30, and products were assigned to the ORFs using rpsblast against the Conserved Domain Database (CDD) version 3.12 with e-value of 1e-05. Intergenic regions were searched for short ORFs using blastx against the nr database (e-value <= 1e-05), and the ORFs were adjusted accordingly. ORFs were also annotated using blastp against the spBeta phage proteins (e-value <= 1e-05) (Supplementary table 2). The annotated genome was deposited in Genbank under accession KY030782.

Preparation of conditioned and control media—Overnight cultures of *B. subtilis* 168 were diluted 1:50 in LB media supplemented with 0.1 mM $MnCl_2$ and 5 mM $MgCl_2$, and incubated at 37°C with shaking until reaching O.D.$_{600}$=0.5. For the conditioned medium only, phages were added to the bacterial culture at MOI=1. Cultures were then incubated for 3 hours at 37°C with shaking. The media were centrifuged at 4000 rpm for 10 minutes at 4°C and the supernatant was filtered with 0.2 μm filter (GE Healthcare Life Sciences, Whatman, CAT# 10462200). Phages and large molecules were further filtered out from the media by using Amicon Ultra centrifugal filters at a cutoff of 3,000 NMWL (3 kDa)

(Milipore, CAT# UFC900324). A plaque assay was performed in order to verify that no phages were left in the medium. The medium was kept at 4°C for up to two months.

For the Proteinase K assay 7.5 mg (per reaction) of Proteinase K–Agarose from *Tritirachium album* (Sigma, CAT# P9290) were washed twice with 750 μl of sterile water and then resuspended with 750 μl of LB supplemented with 0.1 mM $MnCl_2$ and 5 mM $MgCl_2$. Next the tubes were centrifuged again, and the supernatant was discarded. 1.5 ml of phi3T-derived conditioned medium or control medium was added to a tube containing the washed proteinase K. The media were incubated for 2 hours at 37°C with the proteinase K. The media were centrifuged and the supernatants were collected for the infection assay.

Growth dynamics of phage-infected cultures—Overnight cultures of bacteria were diluted 1:100 in LB media and incubated at 37°C with shaking until reaching $O.D_{600}=0.1$. The bacterial culture was centrifuged at 4000 rpm for 10 minutes at room temperature. The supernatant was discarded and the pellet was resuspended in LB medium supplemented with 0.1 mM $MnCl_2$ and 5 mM $MgCl_2$ at 10% of the initial volume. The concentrated bacterial culture was added to control medium, conditioned medium or medium supplemented with synthesized *arbitrium* peptide in a ratio of 1:9 (bacteria to medium). The culture was incubated for 1 hour at room temperature. The culture was then infected with phages at MOI=0.1. Optical density measurements at a wavelength of 600 nm were taken using a TECAN Infinite 200 plate reader in a 96- well plate.

For infection experiments that did not include conditioned medium or addition of a synthesized peptide, the diluted overnight culture was grown to early-logarithmic phase and then infected as above.

Infection dynamics—for the infection dynamics experiments described in Figure 3, an overnight culture of *B. subtilis* 168 was diluted 1/100 and incubated at 37°C with shaking until reaching $O.D_{600}=0.5$. Next, 0.1 mM $MnCl_2$ and 5 mM $MgCl_2$ were supplemented to the medium, and the bacterial culture was split into two flasks. One flask was supplemented with phage phi3T at MOI=0.001, the second was not infected and used as control. Samples for measuring PFU/ml, CFU/ml and peptide concentration were taken every 15 minutes during the time course of 180 minutes. $O.D_{600}$ was also measured every 15 minutes. For CFU measurements, samples were taken from the uninfected culture, diluted in 1xPBS, and plated on LB agar plates. For the PFU measurements, samples were taken from the infected culture, centrifuged for 5 minutes at 4000 RPM and in 4°C. Next, the supernatant was filtered using 0.2 μm filter (GE Healthcare Life Sciences, Whatman, CAT# 10462200). Phage concentrations from these media were measured using a drop assay in serial dilution. For peptide concentration, the media that were filtered previously with the 0.2 μm filter were filtered twice using Amicon Ultra centrifugal filters at a cutoff of 3,000 NMWL (3 kDa) (Milipore, CAT# UFC900324). The media were sent to targeted mass spectrometry analysis. Infection dynamics experiments were done in biological triplicates.

Semi quantitative PCR assay for lysogeny—An overnight culture of bacteria was diluted 1/100 until reaching $O.D_{600}=0.1$. Medium was replaced (with conditioned medium or control medium) as described above, and incubated for 1 hour at room temperature.

Bacteria were infected by phi3T at MOI=5. Cell pellets were collected at times 0, 15, 30, 40 and 60 minutes post infection in the presence of conditioned or control medium. DNA was extracted using Qiagen DNeasy blood and tissue kit (CAT# 69504). Multiplex PCR assays to detect phage phi3T DNA (both as free phage DNA and as a lysogen), *B. subtilis* DNA, and the junction between integrated phage and bacterial genome were performed as previously described by Goldfarb *et al* 30. As controls, genomic DNA from uninfected *B. subtilis* 168 bacteria, as well as a phi3T lysogen (1L1 clone obtained from BGSC), were subject to the same multiplex PCR procedure.

Mass spectrometry—Conditioned media was filtered using 3 kDa MW cutoff filters (Millipore). The low molecular weight fraction was spiked with isotopically labeled internal standard (IS) peptide (SA{Ile(13C6,15N)}RGA) at a final concentration of 50 pg/μl, and then subjected to desalting using the Oasis HLB uElution plates (Waters Corp.). Samples were dried and stored at -80°C until analysis. ULC/MS grade solvents were used for all chromatographic steps. Each sample was loaded using split-less nano-Ultra Performance Liquid Chromatography (10 kpsi nanoAcquity; Waters, Milford, MA, USA). The mobile phase was: A) H_2O + 0.1% formic acid and B) acetonitrile + 0.1% formic acid. Desalting of the samples was performed online using a reversed-phase C18 trapping column (180 μm internal diameter, 20 mm length, 5 μm particle size; Waters). The peptides were then separated using a T3 HSS nano-column (75 μm internal diameter, 250 mm length, 1.8 μm particle size; Waters) at 0.35 μl/minute. For the experiments that are presented in Figure 3 and Extended data figure 3D-E, peptides were separated on the T3 HSS nano-column at 45°C and eluted from the column into the mass spectrometer using the following gradient: 4% to 15%B in 15 minutes, 15% to 90%B in 5 minutes, maintained at 90% for 5 minutes and then back to initial conditions. For the experiments that are presented in Extended data figure 3A-C, peptides were separated on the T3 HSS nano-column at 55°C and eluted from the column into the mass spectrometer using the following gradient: 4% to 35%B in 50 minutes, 35% to 90%B in 5 minutes, maintained at 90% for 5 minutes and then back to initial conditions.

The nanoUPLC was coupled online through a nanoESI emitter (10 μm tip; New Objective; Woburn, MA, USA) to a quadrupole orbitrap mass spectrometer (Q Exactive Plus, Thermo Scientific) using a FlexIon nanospray apparatus (Proxeon). Data was acquired in parallel reaction monitoring (PRM) mode, targeting precursor masses 287.6690 and 291.1776, the doubly charged form of native and isotopically labeled (IS) SAIRGA peptide, respectively. MS2 resolution was set to 35,000 and the maximum injection time set to 200 msec, automatic gain control was set to 2e5. Raw data was imported to Skyline software31 version 3.5. Product ion intensities were extracted and the total area under the curve (AUC) was calculated. Peptide concentration was determined based on the ratio of the AUC from the native versus IS, multiplied by the concentration of the standard. A calibration curve was run to verify that the measurements are within the linear dynamic range. Since the IS was added to the filtered conditioned media, we performed a recovery experiment to calculate the recovery factor in order to obtain the concentration in the unfiltered conditioned medium. We calculated recovery of 40%, thus all concentration measurements were multiplied by 2.5 to yield actual concentration in the pre-filtered medium.

AimR expression and purification—AimR was cloned into the expression vector pET28a (Novagen) using Transfer-PCR (TPCR)32. Cloning was performed using the following primers TP28_aimR_F (5' TTTGTTTAACTTTAAGAAGGAGATATACCATGATTAAGAATGA ATGCGAAAAGG) and TP28_aimR_cHis_R (5'-CTTTGTTAGCAGCCGGATCTTAGTGGTG GTGGTGGTGGTGAATAGAGATAAGGTTTAATAATTCAAG). The integrity of the newly constructed clone, designated pET28-AimR-cHis, was verified by Sanger sequencing. The AimR clone was expressed in *E. coli* BL21(DE3) cells. Freshly transformed BL21(DE3) cells harboring pET28-AimR-cHis were inoculated as a starter culture into 250 ml non-baffled flask containing 100 ml of LB medium supplemented with 30 μg/ml Kanamycin. Growth was performed at 37°C and 250 rpm agitation for ~16 hours. Following overnight growth, 12.5 ml from the starter culture were inoculated into 5 liter non-baffled flasks containing 1.25 liter of LB medium (1:100 dilution of the starter culture, starting OD_{600} ~0.05) supplemented with 30 μg/ml Kanamycin. Large-scale growth and protein expression was performed using a constant shaking of 200 rpm. Cell growth was performed at 37°C, until OD_{600} reached ~0.7. Protein expression was performed at 15°C for about 18 hours using 200 μM IPTG (Isopropyl β-D-1-thiogalactopyranoside) as an inducer.

For protein purification, the cell pellet was resuspended in lysis buffer (50 mM Tris pH 8, 0.3 M NaCl, 20 mM Imidazole, 2 mM DTT, 0.2 mg/ml Lysozyme, 1 μg/ml DNAse, protease inhibitor cocktail (Calbiochem)), disrupted by a cell disrupter at 4°C and clarified at 15,000 g for 30 minutes. The clarified lysate was loaded onto a HisTrap_FF_5 ml column (GE Healthcare) and washed with buffer containing 50 mM Tris pH 8, 0.3 M NaCl, 20 mM imidazole and 2 mM DTT. AimR was eluted from the column in one step with the same buffer containing 0.5 M imidazole. Fractions containing AimR were pooled and injected to a size exclusion column (HiLoad_16/60_Superdex_200_prepgrade, GE_Healthcare) equilibrated with 20 mM Tris pH 8, 0.3 M NaCl, 2 mM TCEP. Fractions containing pure AimR were pooled and flash frozen in aliquots using liquid nitrogen. Purity was verified by the protein running as a single band on SDS PAGE gel stained with coomassie blue (10 μl at 2 mg/ml loaded per lane), and migration as a single peak on an analytical size exclusion column.

Pure AimR was injected to an analytical gel filtration column (Superdex_200_Increase_10/30 GL, GE Healthcare) equilibrated with buffer containing 20 mM Tris pH 8, 0.3 M NaCl, 2 mM TCEP. The migration position of pure AimR was compared to that of AimR-peptide mixtures at the following molar ratios: AimR and SAIRGA peptide (1:2), AimR and GMPRGA peptide (1:1). The column was calibrated (inset of Figure 5C) by monitoring the migration positions of the following known proteins/ polymers: blue dextran (2000 kDa), Thyroglobulin (669 kDa), Apoferritin (443 kDa), beta-amylase (200 kDa), alcohol dehydrogenase (150 kDa), Albumin (66 kDA), Carbonic anhydrase (29 kDa).

Cross linking assays—Cross-linking was generated by reacting the amine-reactive reagent BS3 with the AimR protein in the presence or absence of the 6 aa peptide SAIRGA. Purified AimR (1.37 mg/ml) was eluted from a gel-filtration column in 20 mM Tris pH 8, 300 mM NaCl, 2 mM TCEP. The protein was then dialyzed against 20 mM Hepes pH 7.5,

150 mM NaCl, 2 mM DTT in a 10 KDa cutoff/3 ml G2 Dialysis Cassette (TermoFisher scientific) according to the manufacturer instructions. The dialyzed protein was centrifuged at 21,000 g 4°C for 5 minutes yielding 0.5 mg/ml (11.1 μM). Protein was then mixed with SAIRGA peptide (f.c. 100 μM) dissolved in water or with equal amount of volume without peptide and incubated at room temperature for 5 minutes. The protein samples were then mixed for 30 minutes with 5 different bis(sulfosuccinimidyl)suberate (BS3) concentrations (0.25,0.5, 1.0, 2.0, 4.0 mM). BS3 was prepared by dissolving 1 mg in 70 μl water yielding a 25 mM stock concentration. After the incubation period, samples were quenched for 15 minutes (25°C) with Tris-HCl pH 7.5 in a final concentration of 50 mM. Results were analyzed by electrophoresis in protein Screen Tape using a TapeStation (Agilent Technologies). Sizes of dimer and monomer were verified using a Novex Bolt 4-12% Bis-Tris gel using an SDS-PAGE apparatus (Novex Bolt).

Microscale Thermophoresis (MST)—Two-step purified 6xHis-tagged AimR stored in Tris/NaCl buffer (50 mM Tris pH 8.0, 150 mM NaCl, 2 mM TCEP) at -80°C was thawed on ice and centrifuged at 21,000 g for 10 minutes at 4°C prior to the analysis. Peptides (SAIRGA and GMPRGA) were solubilized in 50 mM Tris-HCl pH 8.0, 150 mM NaCl to a final concentration of 100 μM. AimR was diluted to 200 nM and was incubated with 16 different peptide concentrations varying between 9-4000 nM, which were prepared in Tris/NaCl buffer containing 0.1% [v/v] Pluronic acid (NanoTemper). Roughly 3 μl were loaded into NT.LabelFree Zero-Background Premium Coated Capillaries (NanoTemper) and inserted into a Monolith NT.LabelFree device (NanoTemper). MST experiments were performed at 60% MST power (infra-red laser) and 20% LED power at 23 °C using the Monolith NT.LabelFree instrument (Nanotemper). Ratios between normalized initial fluorescence and post-temperature-jump and thermophoresis were calculated and averaged from 3 independent runs (runs were incubated for 20 minutes at room temperature before the measurement). A plot of fluorescence ratios versus peptide concentration was used to assess the binding capacity of the phage protein and its cognate peptide ligand.

ChIP-Seq—For the ChIP-seq experiments, cell cultures of *B. subtilis* 168 AimR-His tag complementation strain (described in CRISPRi section of the methods) were grown in 100 ml to O.D of 0.1 in LB at 37°C. Then, 50 ml of culture was centrifuged at 4000 rpm for 10 minutes at room temperature. The pellets were then suspended with 25 ml LB either containing or lacking the peptide (SAIRGA) at a final concentration of 1 μM. The cultures were then placed for an additional hour in incubation at room temperature with shaking. After 1 hour, the cultures were infected with phage (MOI=0.5). After 15 minutes of phage infection the cultures were centrifuged for 5 minutes at 4°C 3000 g. The supernatant was discarded and the pellets were resuspended with 1 ml of ice-cold 1xPBS (10 mM Phosphate, 137 mM NaCl, 2.7 mM KCl, pH of 7.4).

For the formaldehyde fixation of protein to DNA, the 1 ml PBS resuspended pellets were then mixed with 62.5 μl formaldehyde (Thermoscientific 16% formaldehyde solution (w/v) methanol-free ampule) yielding a final formaldehyde concentration of 1% (w/v) within the solution. The formaldehyde-containing cell suspension was incubated at room temperature for 10 minutes with mild agitation. After 10 minutes, 75 μl 2 M Glycine (f.c. 150 mM) was

added to quench residual formaldehyde. Glycine-containing samples were kept on ice for an additional 10 minutes followed by centrifugation at 5500 g, 4°C for 1 minute. Supernatant was discarded and pellets were washed with a 1 ml of ice-cold 1xPBS. Centrifugation and wash were repeated 3 times for each sample.

The cell pellets were then suspended with 600 μl lysis buffer containing 50 mM Tris-HCl pH 7.5, 150 mM NaCl and a protease inhibitor mix (cOmplete ULTRA Tablets Roche). The lysis-buffer-containing cells were applied on a lysing matrix B (MP Biomedicals) 0.1 mm silica beads. The mixture and beads were placed in a FastPrep-24 (MP Biomedicals) apparatus and shaken aggressively for 20 sec 6m/sec at 4°C. The beads were then separated from lysed cells by centrifugation of 10,000 g for 1 minute at 4°C according to the manufacturer instructions. Then, 300 μl of the supernatant, containing the lysed cell mixture, was transferred into a 1.5 ml Bioruptor® Plus TPX microtubes (diagenode) and kept on ice for 10 minutes (according to manufacturer's instructions). The sample was then sonicated at 4°C with full power for 15 minutes (30 sec off/on cycles) using the BioRuptor plus (Diagenode) apparatus. Sonication sheared the DNA to an average size of ~500 bp. Samples were then centrifuge at 20,000 g for 10 minutes at 4°C.

For the IP experiments, supernatant containing the lysis buffer and cellular content was then mixed with Triton X-100 and deoxycholate yielding a final IP-buffer composition containing 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% [vol/vol] Triton X-100, 0.1% [wt/vol] sodium deoxycholate supplemented with proteases inhibitors (cOmplete ULTRA Tablets Roche). Anti-6X His tag® ChIP-grade antibody (abcam (ab9108)) was then added to sonicated samples and gently mixed over night at 4°C. In parallel, Protein G Dynabeads (100.04D; Invitrogen) were washed three times with IP buffer.

DNA-protein-antibody complexes (300 μl) were captured with a 100 μl Dynabeads protein G by mixing them for 1 hour at room temperature with rotation. 0.72 μl of 0.5 M EDTA (f.c. 1 mM) was added to that mixture to prevent DNase activity at room temperature. Beads were applied to a magnetic stand (Qiagen) and washed three times with IP buffer (200 μl) at room temperature. Two elution steps were applied with 100 μl and 50 μl of elution buffer (50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 1% [wt/vol] SDS) for 15 minutes at 65°C on a rocking platform. Eluate (100 μl) was incubated with 5 μl of proteinase K (20 mg ml^{-1}) for 1 hour at 50°C and then for 6 hours at 65°C. Immunoprecipitated DNA was recovered using a QIAquick PCR Purification kit (Qiagen). Immunoprecipitated DNA was then converted into NGS libraries using an existing protocol[33] and was sequenced on a NextSeq500 Illumina machine generating 75 nt-long reads.

Sequencing-based assay for lysogeny—An overnight culture of bacteria was diluted 1/100 until reaching O.D$_{600}$=0.1. Medium was replaced as described above (LB supplemented with 0.1 mM MnCl$_2$ and 5 mM MgCl$_2$ with or without SAIRGA), and incubated for 1 hour at room temperature. Bacteria were infected by phi3T at MOI=2. Cell pellets were collected at times 5, 10, 20 and 60 minutes post infection in the presence or absence of SAIRGA peptide at 1 μM final concentration in the medium. DNA was extracted using Qiagen DNeasy blood and tissue kit (CAT# 69504) and subjected to Illumina-based whole genome sequencing on NextSeq500. The relative abundance of lysogens in each

sample was estimated using the number of reads mapped to the uninterrupted integration site versus reads mapped to the integration junction spanning the prophage DNA on one end and the bacterial DNA on the other end.

CRISPRi experiments—Construction of strains silencing phage genes was done by inserting a dCas9 construct controlled by a xylose promoter22 (a gift from the Carol Gross lab, UCSF), into the *lacA* region in *B. subtilis* 168 genome, and sgRNA with spacers targeting the gene of choice under constitutive promoter to *thrC* region (Spacer targeting *aimR*: ACCATTTACTTTTTCATAAC, spacer targeting *aimX*: TTTCCGCTTCATTCTCAAGA, spacer targeting *aimP*: TGTGCTTACTGATTGTTGGC). Infection assays in CRISPRi strains were performed in LB supplemented with 0.1 mM MnCl$_2$, 5 mM MgCl$_2$ and 0.2% xylose.

For complementation assays of *aimR* on the background of *aimR*-silenced CRISPRi strain, *aimR* was amplified from the phi3T genome using the fwd primer AAGAATTCCTCATTGTGTTTAGGTAAAATAAGAAATTC and the rev primer AACTGCAGTTAGTGGTGGTGGTGGTGGTGAATAGAGATAAGGTTTAATAATTCAAG (that includes a 6xHis tag). These primers amplify *aimR* together with 158 bases from its upstream region, with a 6xHis at its C-terminus. The amplified fragment was cloned into the pBS1C plasmid (received from BGSC). Then, the native protospacer adjacent motif in the complemented *aimR* gene was changed by a synonymous point mutation (C->A at codon #20 of the *aimR* gene) using a primer set containing the point mutation and Gibson assembly. The modified gene was then integrated into the *amyE* locus in the *B. subtilis* genome. The *aimX* complementation was constructed on the background of *aimR*-silenced CRISPRi strain. For this, *aimX* was amplified from the phi3T genome using the fwd primer AAACTAGTTTTAAGGGAAAGTTCCAGAAATTC and rev primer AACTGCAGTCCGTTGCCAATAGATTATGC. These primers amplify *aimX* together with 60 bases from its upstream region and 107 bases from its downstream region (containing the gene terminator). The amplified *aimX* was cloned into a pBS1C plasmid modified to contain a xylose promoter, and was then integrated into the *amyE* locus in the *B. subtilis* genome.

RNA-seq—For determining the difference in gene expression with and without the peptide, bacteria (*B. subtilis 168* or *B. subtilis* BEST7003) were incubated for 1 hour in LB medium supplemented with 0.1 mM MnCl$_2$ and 5 mM MgCl$_2$ and in the presence or absence of 1 μM synthesized SAIRGA peptide. Next the bacteria were infected with phi3T (MOI=0.1). Cell pellets were then collected at times 0, 5, 10 and 20 minutes post infection.

RNA extraction and RNA-seq was done as described in Dar *et al.* 34. Briefly, pellets were next lysed using the Fastprep homogenizer (MP Biomedicals, Santa Ana, California), and RNA was extracted with the FastRNA PRO blue kit (MP Biomedicals, 116025050) according to the manufacturer's instructions. RNA samples were treated with TURBO deoxyribonuclease (DNase) (Life technologies, AM2238) and fragmented with fragmentation buffer (Ambion) in 72°C for 1:45 minutes. The reactions were cleaned by adding x2.5 SPRI beads. The beads were washed twice with 80% EtOH, and air dried for 5 minutes. The RNA was eluted using H$_2$O. rRNA was depleted by using the Ribo-Zero rRNA Removal Kit (epicenter, MRZB12424). Strand-specific RNA-seq was performed using the

NEBNext Ultra Directional RNA Library Prep Kit (NEB, E7420) with the following adjustments: all cleanup stages were performed using x1.8 SPRI beads, and only one cleanup step was performed after the end repair step.

For determining the effect of CRISPRi silencing of the *aimR* gene, bacteria at early logarithmic stage were infected with phi3T (MOI=0.1), and cell pellets were collected 20 minutes post infection. RNA-seq libraries were prepped as described above.

RNA-seq libraries were sequenced using the Illumina NextSeq500 platform. Sequenced reads were demultiplexed and adapters were trimmed using fastx_clipper with default parameters. Reads were mapped to the reference genomes (gene annotation and sequences were downloaded from Genbank: NC_000964 for *Bacillus subtilis* str. 168, AP012496 for *Bacillus subtilis* BEST7003, NC_001884 for spBeta and KY030782 for phi3T) using NovoAlign (Novocraft) V3.02.02 with default parameters. All downstream analyses and normalized genome-wide RNA-seq coverage maps were generated as described in Dar *et al*34.

Differential expression analysis—Reads per gene were calculated for each biological replicate of time-point 20 minutes post infection, with and without the synthetic peptide, and normalized relative to the total mapped reads hitting the phage genome in each replicate. Log10 transformation of the average of 3 replicates per gene in each condition was used to plot Extended data figure 6 and calculate the fold change of gene AimX.

Identification of AimR homologs and the arbitrium peptide code—Homologs for the phi3T AimR receptor were searched for using the BLAST option in the Integrated Microbial Genomes (IMG) web server (https://img.jgi.doe.gov/cgi-bin/mer/main.cgi). The phi3T AimR was provided as a query sequence and was searched against all isolate genomes with an e-value threshold of 1e-35. The gene neighborhood for each AimR homolog was visually inspected via the IMG "gene neighborhood" representation, and genes found located next to proteins annotated as phage proteins were considered as found in a prophage. The immediate downstream gene for each AimR homolog was considered the respective AimP gene if it contained a signal peptide as predicted by the IMG web server. If no immediate downstream gene was annotated, the intergenic region immediately downstream to the AimR homolog was translated using the Expasy Translate Tool (http://web.expasy.org/translate/), and short translated ORFs were inspected for the AimP signature. Phages harboring the arbitrium system were typed according to the following parameters: prophage genome length, prophage gene order, and Phaster35 and Virfam36 classifications. Results of this analysis are presented in Supplementary table 1.

Construction of aimP phage deletion strain—In order to exchange *aimP* with an antibiotic cassette, a plasmid composed of four fragments was constructed as follows: (i) the upstream region flanking *aimP* (1000 bp upstream to *aimP* start codon) was amplified using the fwd primer TGAAGAGCAACTTTTAAGTG and the rev primer TATTCTCACCTCCTTTCAAAT. (ii) the Spectinomycin resistance cassette was amplified from plasmid p15a-pDR110 using the fwd primer GTTTGACAAATTTGAAAGGAGGTGAGAATACGAATGGCGATTTTCGTTC and the

rev primer GTATTATGTATTACCTATTCAATTATTTAACCCCCTATGCAAGGGTTTATT. (iii) the downstream region flanking *aimP* (1001 bp downstream to *aimP* stop codon) was amplified using the fwd primer TTAAATAATTGAATAGGTAATACATAATAC and the rev primer CAGATTTCTATTTGCTTCATTAAG. (iv) the origin of replication and Ampicillin resistance was amplified from the plasmid pBS1C using the fwd primer TTATGACTTAATGAAGCAAATAGAAATCTGCGGGACTTACCGAAAGAAAC and the rev primer AAGAAGTATTCACTTAAAAGTTGCTCTTCATTCTTGACACTCCTTATTTGA. The four fragments were assembled together using NEBuilder® HiFi DNA Assembly Master Mix (NEB, E2621S). The plasmid was transformed into NEB® 10-beta Competent *E. coli* (NEB, C3019I) for amplification of the plasmid, and then transformed into *B. subtilis* BEST7003 harboring a phi3T prophage. The deletion of *aimP* was verified using Illumina sequencing. The phages were then induced using Mitomycin C (Sigma, M0503). Conditioned media preparation and infection dynamics assays with this phage were performed as described above.

Data availability

The phi3T genome was deposited in GenBank under accession number KY030782. Phage gene annotations are in Supplementary table 2. AimR homologs list and the corresponding peptides for each one are in Supplementary table 1. RNA-seq and ChIP-seq data were deposited in the European Nucleotide Database (ENA), accession number PRJEB18541.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Rutberg, L. *The Molecular Biology of Bacilli* Vol. 1 *Bacillus subtilis*. Ch. Temperate bacteriophages of *Bacillus subtilis*. Dubnau, DA., editor. Academic press; 1982. p. 247-268.

2. Oppenheim AB, Kobiler O, Stavans J, Court DL, Adhya S. Switches in bacteriophage lambda development. Annu Rev Genet. 2005; 39:409–29. [PubMed: 16285866]

3. Zeng L, et al. Decision making at a subcellular level determines the outcome of bacteriophage infection. Cell. 2010; 141:682–691. [PubMed: 20478257]

4. Pottathil M, Lazazzera BA. The extracellular Phr peptide-Rap phosphatase signaling circuit of *Bacillus subtilis*. Front Biosci. 2003; 8:d32–45. [PubMed: 12456319]

5. Perego M. Forty years in the making: understanding the molecular mechanism of peptide regulation in bacterial development. PLoS Biol. 2013; 11:e1001516. [PubMed: 23526885]

6. Waters CM, Bassler BL. Quorum sensing: cell-to-cell communication in bacteria. Annu Rev Cell Dev Biol. 2005; 21:319–46. [PubMed: 16212498]

7. Lanfermeijer FC, Detmers FJ, Konings WN, Poolman B. On the binding mechanism of the peptide receptor of the oligopeptide transport system of *Lactococcus lactis*. EMBO J. 2000; 19:3649–56. [PubMed: 10899119]

8. Tucker RG. Acquisition of thymidylate synthetase activity by a thymine-requiring mutant of *Bacillus subtilis* following infection by the temperate phage phi 3. J Gen Virol. 1969; 4:489–504. [PubMed: 4979518]

9. Rocha-Estrada J, Aceves-Diez AE, Guarneros G, de la Torre M. The RNPP family of quorum-sensing proteins in Gram-positive bacteria. Appl Microbiol Biotechnol. 2010; 87:913–23. [PubMed: 20502894]

10. Do H, Kumaraswami M. Structural mechanisms of peptide recognition and allosteric modulation of gene regulation by the RRNPP family of quorum-sensing regulators. J Mol Biol. 2016; 428:2793–804. [PubMed: 27283781]

11. Perez-Pascual D, Monnet V, Gardan R. Bacterial cell-cell communication in the host via RRNPP peptide-binding regulators. Front Microbiol. 2016; 7:706. [PubMed: 27242728]

12. Dunny GM, Berntsson RP-A. Enterococcal sex pheromones: evolutionary pathways to complex, two-signal systems. J Bacteriol. 2016; 198:1556–62. [PubMed: 27021562]

13. Shi K, et al. Structure of peptide sex pheromone receptor PrgX and PrgX/pheromone complexes and regulation of conjugation in *Enterococcus faecalis*. Proc Natl Acad Sci USA. 2005; 102:18596–601. [PubMed: 16339309]

14. Lereclus D, Agaisse H, Gominet M, Salamitou S, Sanchis V. Identification of a *Bacillus thuringiensis* gene that positively regulates transcription of the phosphatidylinositol-specific phospholipase C gene at the onset of the stationary phase. J Bacteriol. 1996; 178:2749–56. [PubMed: 8631661]

15. Slamti L, Lereclus D. A cell-cell signaling peptide activates the PlcR virulence regulon in bacteria of the *Bacillus cereus* group. EMBO J. 2002; 21:4550–9. [PubMed: 12198157]

16. Declerck N, et al. Structure of PlcR: Insights into virulence regulation and evolution of quorum sensing in Gram-positive bacteria. Proc Natl Acad Sci USA. 2007; 104:18490–5. [PubMed: 17998541]

17. Dubois T, et al. Activity of the *Bacillus thuringiensis* NprR-NprX cell-cell communication system is co-ordinated to the physiological stage through a complex transcriptional regulation. Mol Microbiol. 2013; 88:48–63. [PubMed: 23388036]

18. Fleuchot B, et al. Rgg proteins associated with internalized small hydrophobic peptides: a new quorum-sensing mechanism in streptococci. Mol Microbiol. 2011; 80:1102–1119. [PubMed: 21435032]

19. Parashar V, Mirouze N, Dubnau DA, Neiditch MB. Structural basis of response regulator dephosphorylation by Rap phosphatases. PLoS Biol. 2011; 9:e1000589. [PubMed: 21346797]

20. Ishikawa S, Core L, Perego M. Biochemical characterization of aspartyl phosphate phosphatase interaction with a phosphorylated response regulator and its inhibition by a pentapeptide. J Biol Chem. 2002; 277:20483–9. [PubMed: 11923303]

21. Baker MD, Neiditch MB. Structural basis of response regulator inhibition by a bacterial anti-activator protein. PLoS Biol. 2011; 9:e1001226. [PubMed: 22215984]

22. Peters JM, et al. A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. Cell. 2016; 165:1493–506. [PubMed: 27238023]

23. Johnson CM, Grossman AD. Integrative and Conjugative Elements (ICEs): What they do and how they work. Annu Rev Genet. 2015; 49:577–601. [PubMed: 26473380]

24. Auchtung JM, Lee CA, Monson RE, Lehman AP, Grossman AD. Regulation of a *Bacillus subtilis* mobile genetic element by intercellular signaling and the global DNA damage response. Proc Natl Acad Sci USA. 2005; 102:12554–9. [PubMed: 16105942]

25. Hargreaves KR, Kropinski AM, Clokie MRJ. What does the talking?: quorum sensing signalling genes discovered in a bacteriophage genome. PLoS One. 2014; 9:e85131. [PubMed: 24475037]

26. Emanuelsson O, Brunak S, von Heijne G, Nielsen H. Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc. 2007; 2:953–971. [PubMed: 17446895]

27. Patrick JE, Kearns DB. Laboratory strains of *Bacillus subtilis* do not exhibit swarming motility. J Bacteriol. 2009; 191:7129–33. [PubMed: 19749039]

28. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 2008; 18:821–9. [PubMed: 18349386]

29. Salzberg SL, Delcher AL, Kasif S, White O. Microbial gene identification using interpolated Markov models. Nucleic Acids Res. 1998; 26:544–8. [PubMed: 9421513]

30. Goldfarb T, et al. BREX is a novel phage resistance system widespread in microbial genomes. EMBO J. 2015; 34:169–83. [PubMed: 25452498]

31. MacLean B, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics. 2010; 26:966–8. [PubMed: 20147306]

32. Erijman A, Dantes A, Bernheim R, Shifman JM, Peleg Y. Transfer-PCR (TPCR): A highway for DNA cloning and protein engineering. J Struct Biol. 2011; 175:171–177. [PubMed: 21515384]

33. Garber M, et al. A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. Mol Cell. 2012; 47:810–22. [PubMed: 22940246]

34. Dar D, et al. Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria. Science. 2016; 352:aad9822. [PubMed: 27120414]

35. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 2016; 44:W16–21. [PubMed: 27141966]

36. Lopes A, et al. Automated classification of tailed bacteriophages according to their neck organization. BMC Genomics. 2014; 15:1027. [PubMed: 25428721]
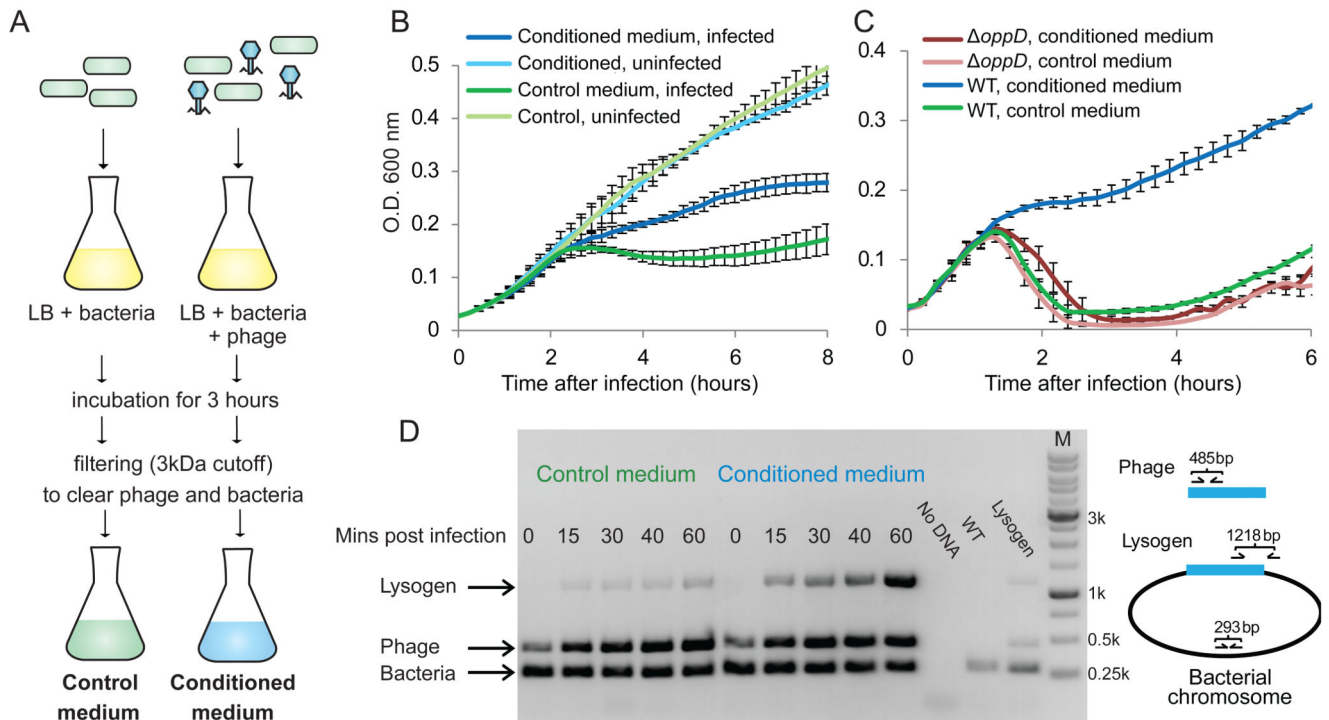
Figure 1. Effect of conditioned media on the infection dynamics of phage phi3T.
(A) Preparation protocol of control and conditioned media. (B) Growth curves of *B. subtilis* 168 infected by phi3T at MOI=0.1, in control and conditioned media. (C) Growth curves of *B. subtilis* strain 3610 (WT) and its derivative DS4979 (*oppD*::kan) infected by phi3T at MOI=0.1. For panels B-C, data represents average of 3 biological replicates, each with 3 technical replicates; error bars represent SE. (D) Semi-quantitative PCR assay for phage lysogeny during an infection time course of *B. subtilis* 168 with phi3T. "No DNA", control without DNA; "WT", DNA from uninfected culture; "Lysogen", genomic DNA of a phi3T lysogen .
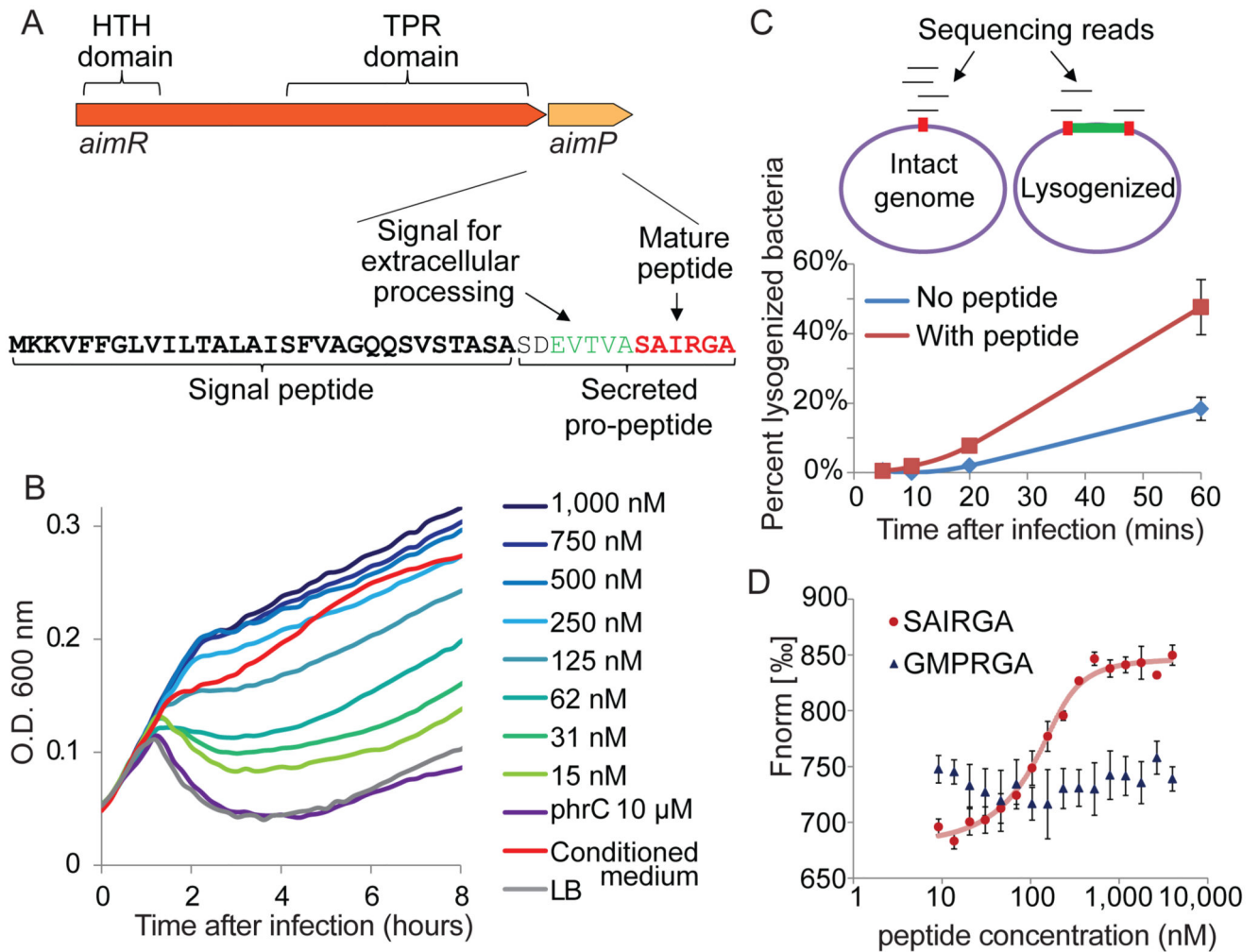
Figure 2. The *arbitrium* peptide and its receptor.

(A) The *arbitrium* locus in the phage genome. (B) Growth curves of *B. subtilis* 168 infected by phi3T at MOI=0.1, in LB media supplemented with synthesized SAIRGA peptide. Numbers represent peptide concentrations. Average of 3 biological replicates, each with 3 technical replicates. (C) Sequencing-based quantitative determination of the fraction of lysogenized bacteria. Top – schematics of the analysis. Percent lysogeny was calculated as the fraction of reads spanning the phage/bacteria integration junction out of total reads covering this junction. Integration junction is red; integrated phage is green. Bottom: percent lysogenized bacteria during infection of *B. subtilis* 168 with phi3T at MOI=2. Average of three biological replicates, error bars denote SE. Synthesized SAIRGA peptide was added at 1 µM. (D) Microscale thermophoresis (MST) analysis of the binding between purified AimR and synthesized SAIRGA or GMPRGA peptides. Average and SE of three replicates.
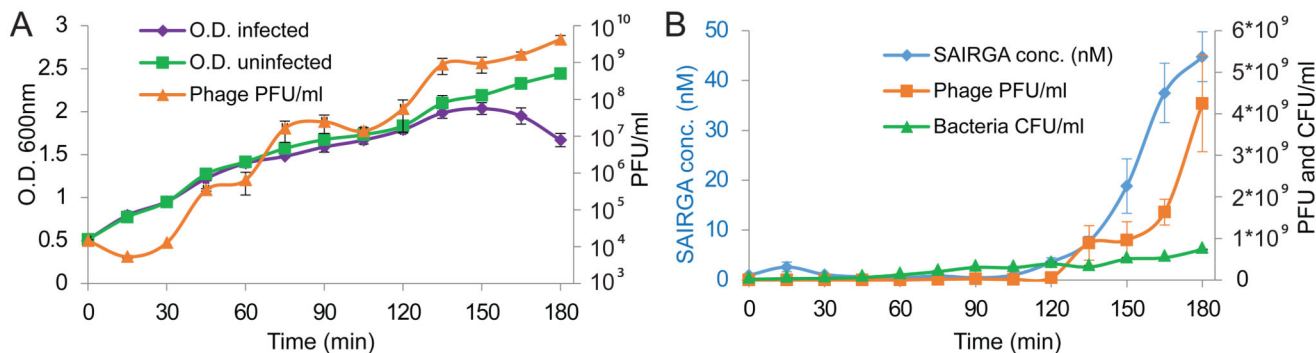
Figure 3. Accumulation of *arbitrium* peptide during an infection time course.
B. subtilis 168 culture was grown to OD=0.5 and then infected at t=0 by phi3T at MOI=0.001. PFU, plaque forming units; CFU, colony forming units, sampled from the uninfected control. For peptide, PFU and OD shown is average of triplicate (except for t=30 where it is duplicate) with error bars representing SE. CFU was measured in duplicates, and error bars represent the two measured points around the average. (A) Growth of infected and uninfected bacteria, and phage, during the infection time course. (B) Accumulation of peptide during the infection course as compared to bacteria and phage.
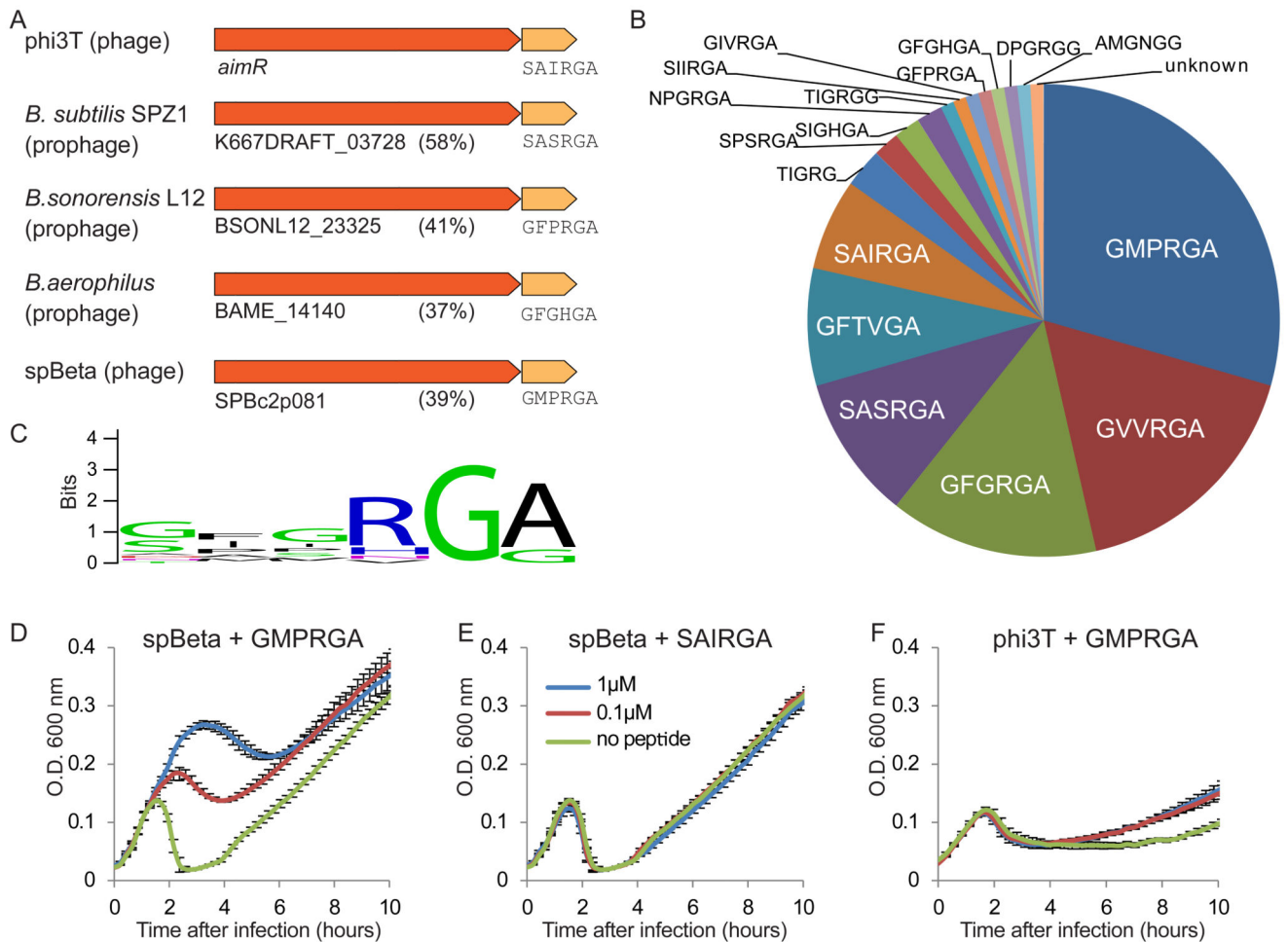
Figure 4. A peptide communication code guiding lysogeny in *Bacillus* phages.
(A) Selected instances of AimR homologs in sequenced genomes. Locus tags are indicated for AimR homologs, along with the percent amino acid sequence identity to the phi3T AimR. Mature *arbitrium* peptide is indicated below the AimP homolog. (B) Distribution of *arbitrium* peptides among 112 homologs of AimP. (C) Amino acid profile of *arbitrium* peptide types. (D) Growth curves of *B. subtilis* BEST7003 infected by spBeta at MOI=0.1, in LB media supplemented with synthesized GMPRGA peptide. The BEST7003 strain was used for spBeta infection as the *B. subtilis* 168 strain is naturally immune to spBeta. (E) Growth curves of *B. subtilis* BEST7003 infected by spBeta at MOI=0.1, in LB media supplemented with synthesized SAIRGA peptide. (F) Growth curves of *B. subtilis* 168 infected by phi3T at MOI=0.1, in LB media supplemented with synthesized GMPRGA peptide. Data in panels D-F represent average of 3 biological replicates, each with 3 technical replicates; error bars represent SE.
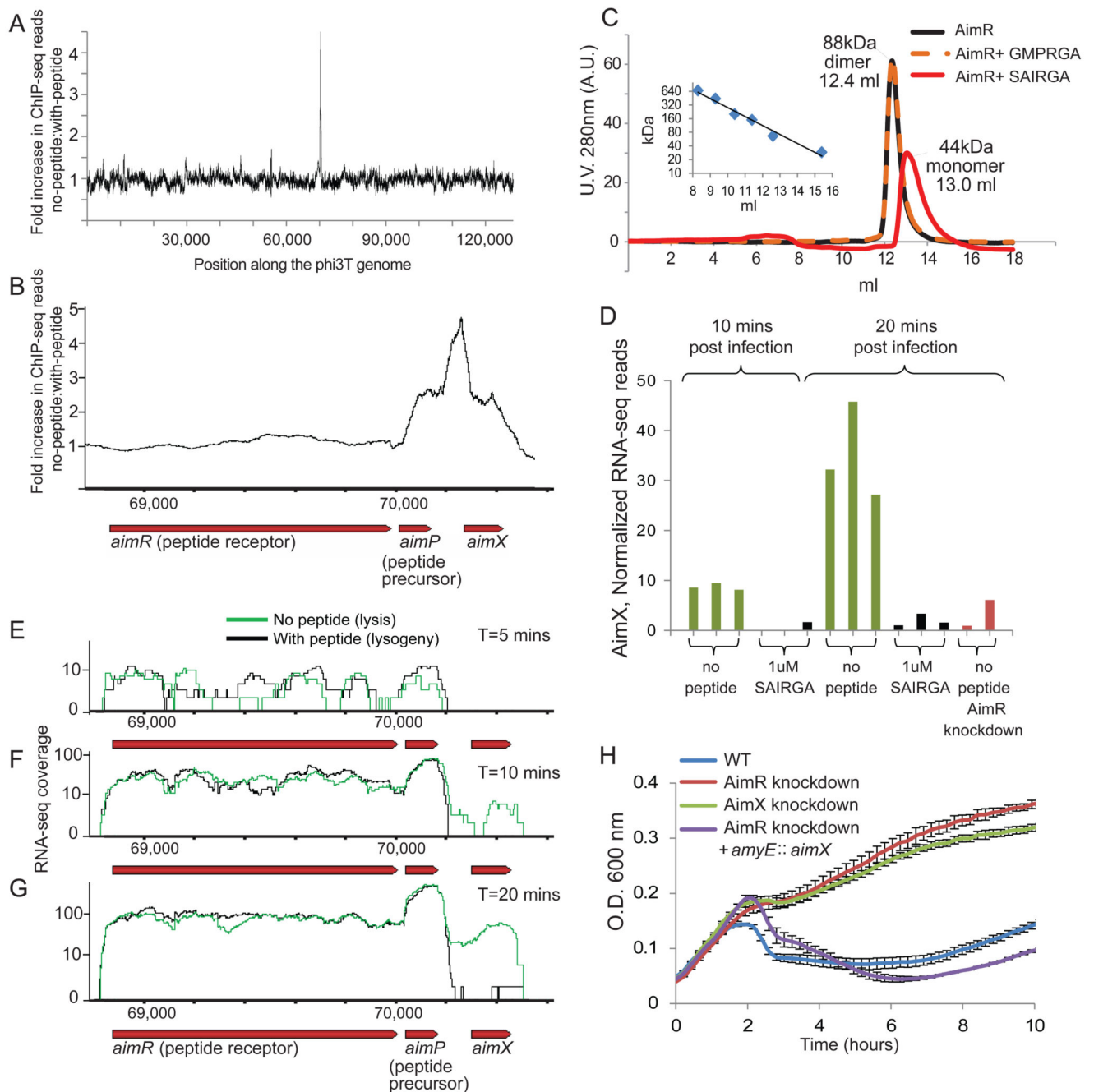
Figure 5. DNA binding and transcription regulation in the *arbitrium* system.
(A) ChIP-seq of His-tagged AimR 15 minutes post-infection with or without 1 μM of SAIRGA peptide. Shown is the ratio, along the phage genome, between sequenced pulled-down DNA during infection without the peptide and DNA pulled-down when the peptide was present in the medium. (B) Same as panel A, shown is a zoomed-in region in the phage genome. (C) Gel-filtration results of purified AimR with or without the presence of either SAIRGA or GMPRGA peptide. Inset presents a calibration curve for the gel filtration using proteins of known sizes. (D) Expression of the AimX gene during infection. Data presented

for individual biological replicates (E-G) RNA-seq coverage of the *arbitrium* locus at 5 minutes (E), 10 minutes (F) and 20 minutes (G) post infection. (H) Growth curves of WT and dCas9-silenced bacterial strains during phi3T-infection. Strains were infected at t=0 at MOI=0.1. Shown is average of 3 biological replicates, each with 3 technical replicates; error bars represent SE.
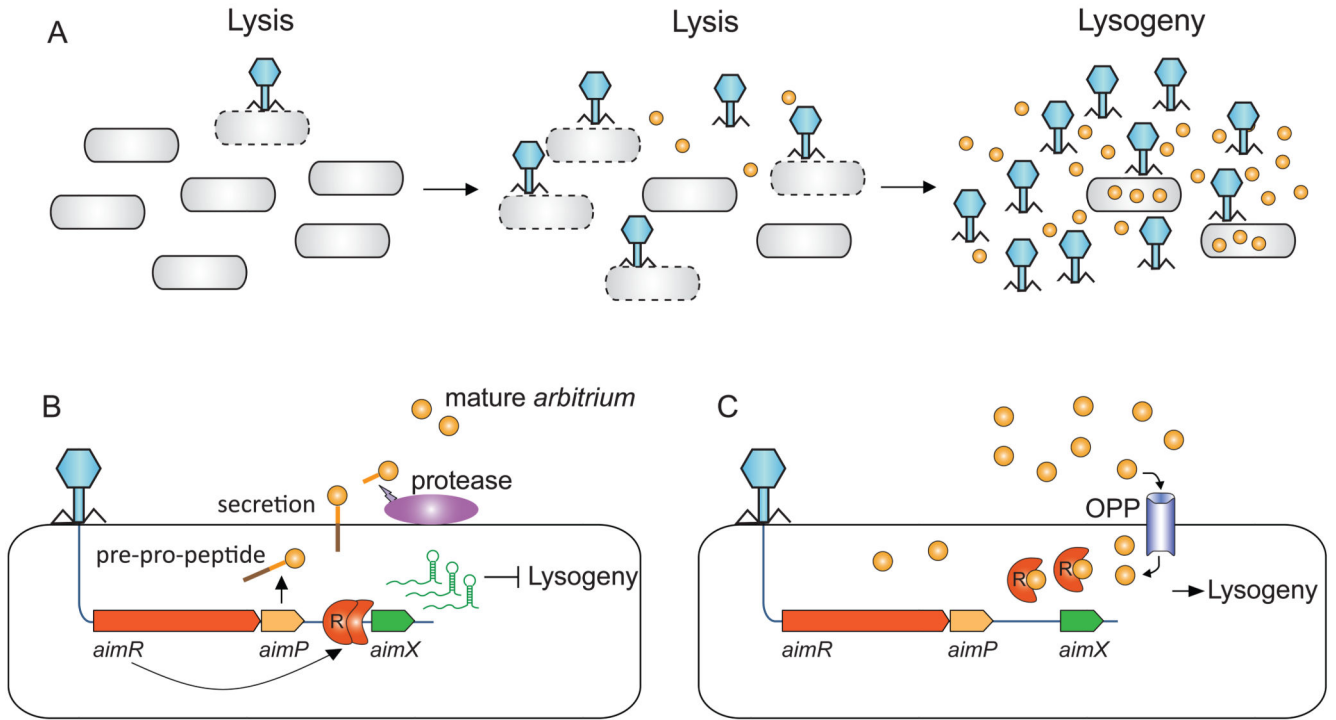
Figure 6. Mechanistic model for communication-based lysis-lysogeny decisions.
(A) Dynamics of *arbitrium* accumulation during infection of a bacterial culture by phage.
(B) At the first encounter of a phage with a bacterial population, the early genes *aimR* and *aimP* are expressed immediately upon infection. AimR, as a dimer activates AimX expression. AimX is an inhibitor of lysogeny, possibly as a regulatory ncRNA, directing the phage to a lytic cycle. At the same time AimP is expressed, secreted and processed extracellularly to produce the mature peptide. (C) At later stages of the infection dynamics, the *arbitrium* peptide accumulates in the medium and is internalized into the bacteria by the OPP transporter. Now when the phage infects the bacterium, the expressed AimR receptor binds the *arbitrium* molecules and cannot activate the expression of AimX, leading to lysogeny preference.