

Metagenome-inspired libraries to engineer phage M13 for targeted killing of Gram-negative bacterial species

Yanxi Yang^{1,2}, Dayeon Kang^{1,2}, Beatrice Mihalache^{1,2}, Shelby Vexler^{1,2}, Saumya Jain¹, Huan Peng³, Nasim Annabi^{1,4}, Shangxin Yang⁵, Irene A. Chen^{1,2,4,*}

¹Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA 90095, United States

²Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, United States

³Cellular Signaling Laboratory, International Research Center for Sensory Biology and Technology of MOST, Key Laboratory of Molecular Biophysics of MOE, College of Life Science and Technology, Huazhong University of Science and Technology, 430074 Wuhan, Hubei, China

⁴Department of Bioengineering, University of California, Los Angeles, CA 90095, United States

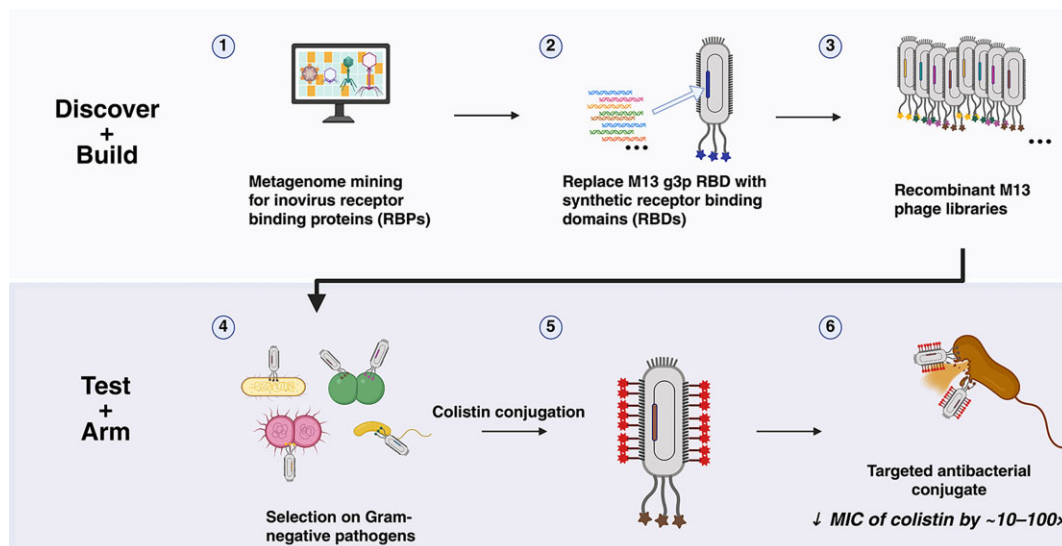
⁵Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, United States

*To whom correspondence should be addressed. irenechen@ucla.edu

Abstract

Given concerning trends in antibiotic resistance, phages have been increasingly explored as promising antimicrobial agents. However, a major problem with phage therapy is the overly high specificity of phages for their hosts, which is currently addressed by a personalized approach involving screening a bank of wild-type phages against each clinical isolate. To shorten this process, we propose that a focused library of synthetic phages could be rapidly selected for a member binding to a given clinical isolate. We created libraries of recombinant M13 phages expressing receptor-binding proteins based on the collective metagenome of inovirus phages, a diverse group whose members appear to infect nearly all bacterial phyla. Using two rounds of a pull-down selection, phage variants were identified against several Gram-negative pathogens, including a variant (M13^{PAB}) that bound to several *Pseudomonas aeruginosa* strains, including clinical isolates. To confer bactericidal activity to the nonlytic phage, a last-line but nephrotoxic lipopeptide, colistin, was cross-linked to the M13^{PAB} virions. The colistin-M13^{PAB} phage conjugate lowered the minimal inhibitory concentration of colistin by 1–2 orders of magnitude for multiple strains of *P. aeruginosa* and showed a lack of hemolytic or cytotoxic activity *in vitro*, suggesting high potency combined with low toxicity. Thus, a metagenome-inspired library displayed on the M13 phage scaffold, when subjected to a short selection for binding to a bacterial clinical isolate, could yield a phage variant that targets the specified strain. This approach may improve the speed, consistency, and cost-effectiveness of personalized phage therapy.

Graphical abstract



Received: March 24, 2025. Revised: September 7, 2025. Accepted: September 8, 2025

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Introduction

Antibiotic resistance in bacterial pathogens is one of the leading threats to global healthcare [1]. The CDC reported 2.8 million cases of antibiotic-resistant infections leading to 35 000 deaths in the U.S. in 2019, and the World Health Organization estimates that the number of deaths attributed to drug-resistant infections will grow to 10 million worldwide by 2050 [2–5]. Among bacteria, Gram-negative pathogens are particularly problematic due to numerous antimicrobial resistance mechanisms acquired vertically or horizontally [6–9]. Alternative antimicrobial strategies include bacteriophage (phage) therapy, based on naturally occurring viruses that specifically infect bacteria [10, 11]. Phages collectively show high diversity in terms of morphology, genetic content, and infection cycles [12–14]. Phage therapy typically employs lytic phages to infect and kill bacteria through replication [14]. Potential advantages of phages and phage-based products include high specificity, the ability to penetrate biofilms, relatively low manufacturing cost, and safety [11, 15].

However, phage therapy faces multiple challenges, including bacterial resistance to phage, the host immune response, and host cell specificity, each of which must be addressed to advance clinical applications [16, 17]. Regarding host cell specificity, high specificity can be advantageous in reducing impacts on the healthy microbiome compared to most antibiotics [18, 19]. However, the naturally high specificity of phages can be disadvantageous in requiring the identification, characterization, and manufacturing of a phage specific to the particular clinical strain of the patient. One approach to systematize this process is to build a pre-characterized library of phages [20, 21], as in PhageBank™, a continuously expanding phage library that has received FDA approval for clinical studies. However, collecting and characterizing each phage, comprising a highly diverse set of genomes, remains a resource-demanding process.

An alternative approach is to engineer a well-characterized phage, such as M13, to target various bacterial species [22–24]. During the attachment step of infection, phages achieve specificity through the receptor-binding domain (RBD) of the receptor-binding protein (RBP), which binds to the bacterial receptor. Thus, engineering the RBD can control bacterial targeting of this step. To avoid the need to engineer the full lytic cycle on the new host species, the phage may be further modified to exert antibacterial effect upon binding regardless of downstream infection. This approach therefore involves engineering both host specificity, through the RBD interaction, and mode of cell-killing, such as through attachment of bactericidal cargo. This combination of engineered specificity and controlled therapeutic effect has been called “controlled phage therapy” [22]. M13 phage, widely used for phage display [25], is a member of the *Inoviridae* family, whose members target various Gram-negative bacteria and establish chronic nonlytic infections [26]. The filamentous body of wild-type (WT) M13 virions is composed of ~2700 copies of the major capsid protein pVIII (g8p), five copies of minor capsid proteins pIII and pVI that cap one terminus of the virion, and pVII and pIX that cap the other [27]. The M13 minor capsid protein pIII (g3p) is the receptor-binding protein that allows WT M13 phages to specifically recognize host *Escherichia coli* (*E. coli*) cells. The g3p protein is composed of an N-terminal RBD [28] and a C-terminal domain anchored in the phage virion, with the domains connected by a glycine-rich linker [25, 29]. The M13

RBD itself consists of an N2 domain that binds the tip of the F pilus and an N1 domain that interacts with TolA, connected by a shorter glycine-rich linker (Scheme 1). Swapping the WT RBD with the RBD of other inoviruses (e.g. phage If1 or Pf1) creates recombinant phages with redirected targeting to other Gram-negative bacterial species [23, 30]. While the resulting recombinant phages bind their new target species, they may not complete a replicative cycle and are instead decorated with cargo effecting various functionalities. In particular, the major coat protein of M13 contains solvent-exposed primary amines and carboxylic acid functional groups that are amenable to bioconjugation for delivery of bactericidal agents [31, 32].

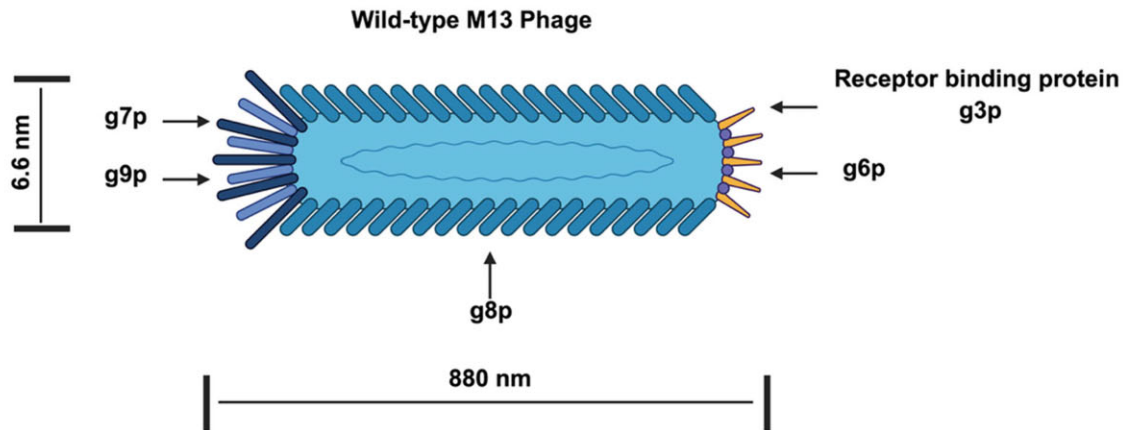
Here we present a strategy to build a library of recombinant phages with diverse RBDs for targeting various Gram-negative bacterial species. Phage display selection with a randomized peptide library is a possible approach, but an ideal library for repeated screening in a therapeutic setting should be as small as possible while still including members that could bind to a wide variety of Gram-negative species. In other words, an ideal library would only include functional RBDs covering a variety of species. Recent metagenomic analysis indicates that inoviruses are an extremely diverse group, encompassing ~10⁵ members with host species from nearly all bacterial phyla [33]. Given the close co-evolutionary relationship between phages and their hosts, we posited that a library based on the inovirus RBDs could be rapidly screened for phage variants capable of binding a given Gram-negative strain. Here we describe library construction and characterization, followed by 18 independent screens that identified M13 variants that bound to Gram-negative pathogens, including 2 strains of *E. coli*, 8 strains of *Pseudomonas aeruginosa* (including 6 clinical isolates), 5 strains of *Klebsiella pneumoniae* (including 4 clinical isolates), 1 strain of *Acinetobacter baumannii*, 1 strain of *Burkholderia cepacia*, and 1 strain of *Chronobacter sakazakii*. When loaded with bactericidal cargo (colistin), a phage conjugate killed the targeted species with high potency *in vitro*. Thus, synthetic RBD libraries based on the phage metagenome are a promising approach for rapidly identifying recombinant phages for controlled phage therapy.

Materials and methods

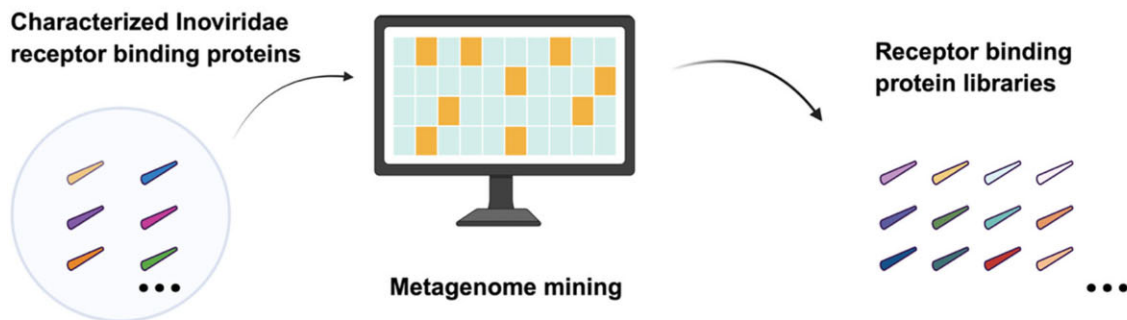
Materials

All chemical reagents used in this study, including colistin sulfate, were purchased from Millipore Sigma, if not specified otherwise. Oligonucleotides used for library construction, polymerase chain reaction (PCR), and quantitative PCR (qPCR) were ordered from Integrated DNA Technologies; sequences are listed in [Supplementary Table S1](#). Phagemid vector pADL-10b and helper phage CM13d3 were obtained from Antibody Design Labs. All restriction enzymes, Q5 PCR reagents, and DNA HiFi assembly kit reagents were purchased from New England Biolabs. DNA purifications from enzymatic reaction steps (restriction enzyme digestion and PCR) were performed using the QIAquick PCR Purification Kit (Qiagen). DNA purifications from agarose gel electrophoresis were performed using the QIAquick Gel Extraction Kit (Qiagen). Purification of colistin-M13^{PAB} was conducted using Slide-A-lyzer (20K MWCO, Thermo Fisher Scientific). Mix & Go! Competent Cells-JM109 were used for all transformations (Zymo Research).

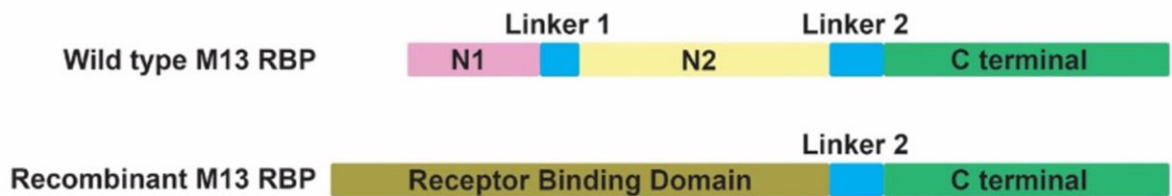
A.



B.



C.



Scheme 1. M13 phage and the metagenome-derived RBP libraries. **(A)** The WT M13 phage has a filamentous morphology that is ~880 nm in length and 6.6 nm in diameter [34]. The phage protein shell is composed of ~2700 copies of the major capsid protein g8p with one end capped by five copies of the minor capsid proteins g7p and g9p, and the other end with five copies of the minor capsid protein g6p and the RBP g3p. Physical dimensions are indicated. **(B)** Putative Inoviridae RBPs were identified computationally through metagenome mining using known RBPs of characterized Inoviridae family members. **(C)** WT g3p is composed of the N1, N2, and C terminal domains connected by glycine-rich linkers. In the recombinant libraries, the N1 and N2 domains and intervening linker were replaced with the putative RBD from the phage metagenome. Created in BioRender. Chen, I. (2025) <https://BioRender.com/ndq4658>.

Microbial strains

The microbial strains used in this study and their sources are given in Table 1. Bacterial and yeast strains were revived from glycerol stocks stored at -80°C by streaking on LB agar plates and culturing in a 37°C incubator. A single colony from each plate was inoculated into the corresponding liquid culture and incubated at 37°C with shaking at 220 rpm. To ensure bacteria were fully revived, the cells were subcultured >3 times by inoculating 100 μL of overnight culture into 10 mL of fresh media prior to use.

Clinical strains were obtained from routine aerobic bacterial cultures performed in the UCLA Clinical Microbiology Laboratory. All clinical isolates in this study were obtained from lower respiratory tract specimens, including sputum and bronchoalveolar lavage, and deidentified. Samples were inoculated on sheep blood agar plates (Becton Dickinson, MD, USA) and incubated at 35°C in 6%–10% CO_2 for 12 to 24 h. Isolates were identified to species level using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF, VITEK MS, BioMérieux, NC, USA).

Table 1. List of species, strains, and sources for microorganisms used in this work

Species	Strain name	Source
<i>Escherichia coli</i>	ATCC 700927	American Type Culture Collection (ATCC)
<i>Escherichia coli</i>	DH5 α	Zymo Research (Irvine, CA)
<i>Escherichia coli</i>	ATCC 25922	ATCC
<i>Pseudomonas aeruginosa</i>	ATCC 25102	ATCC
<i>Pseudomonas aeruginosa</i>	PAKpmrB6	Jian Li, Monash University [35, 36]
<i>Pseudomonas aeruginosa</i>	14 clinical isolates, designated A through M, and O	UCLA Clinical Microbiology Laboratory
<i>Pseudomonas aeruginosa</i>	Clinical Strain 320	UCLA Clinical Microbiology Laboratory
<i>Klebsiella pneumoniae</i>	ATCC 700603	ATCC
<i>Klebsiella pneumoniae</i>	4 clinical isolates, designated A through D	UCLA Clinical Microbiology Laboratory
<i>Acinetobacter baumannii</i>	ATCC 19606	ATCC
<i>Burkholderia cepacia</i>	ATCC 25416	ATCC
<i>Chronobacter sakazakii</i>	ATCC 25944	ATCC
<i>Saccharomyces cerevisiae</i>	ATCC 204508 / S288c (Baker's yeast)	ATCC

RBP library design from metagenomic data

To create a phagemid library expressing potential phage RBDs fused to the C-terminal domain of M13 g3p, we used a published metagenomic database containing 10 293 inovirus-like genome sequences identified using machine learning [33]. The inovirus-like genomes, organized in GenBank files with annotations, contain 75 325 open reading frames in total. A local database containing these open reading frames (ORFs) (putative inovirus protein or “PIP” database) was created (PIP database, [Supplementary Data File 1](#)). The PIP database was used to design two different phagemid libraries using characterized inovirus RBPs as reference sequences. Putative RBD sequences from the PIP database were included based on similarity to reference sequences. The two libraries were designed as follows.

The first library (“Pa-RBD” library) was designed to target *P. aeruginosa*. The PIP database was searched for sequences similar to the RBPs from *P. aeruginosa* filamentous phage Pf1 [37, 38] (reference sequences: GenBank protein ID: AAQ94686.1 and NP_039604.1; [Supplementary Fig. S1](#) [39, 40]). Protein sequences in the PIP database with high similarity to each reference sequence were identified using the Basic Local Alignment Search Tool (NCBI BLAST+ 2.11.0; default settings) [41, 42] and sequence hits with *E*-value >0.002 were removed. Given limitations on pool synthesis (described below) and the length of the two reference RBPs (437 and 439 amino acids), we removed sequences <300 or >550 amino acids long (Code 2). To reduce redundancy, a pairwise alignment score was calculated among the remaining hits using the Biopython package “pairwise2” module [43]. For any pair of hits having a pairwise alignment score ≥ 320 , the hit with a greater *E*-value was removed (Code 3). The remaining hits for the two reference RBPs were combined and any duplicates were removed (Code 4). A starting Met was added to any sequences that did not begin with Met (Code 5). Since signal peptides [44] are expected on true RBPs, SignalP 5.0 (Version 5.0b, Platform: Linux) was used to identify signal peptides, and sequences lacking them were removed [45]. Since the phagemid vector contains a signal peptide for phage production in *E. coli*, the native predicted signal peptides were removed from the sequence designs (Code 6). The RBPs of known inoviruses, including M13, If1, and both Pf1 phage genomes, contain a glycine-rich linker separating the C-terminal domain, which integrates into the capsid, from the solvent-exposed N-terminal RBD [30]. To identify putative glycine-rich linkers in sequence designs, a sliding win-

dow of 30 amino acids was applied to identify the region with the highest combined abundance of linker-prone amino acids (Gly, Glu, Asn, Thr, Asp, Ser, and Pro) in each design [46]. The region N-terminal to the putative linker was designated as the putative RBD for each sequence. A hit was removed if it lacked a putative linker (i.e. had no window containing >9 linker-prone amino acids) (Code 7). Given that the reference RBDs have lengths of 265 and 257 amino acids and that length homogeneity is important for synthesis, putative RBDs <200 amino acids in length were removed, yielding 61 sequences designed to comprise the Pa-RBD library (Pa-RBD designs, [Supplementary Data File 1](#)). Each putative RBD was manually inspected to ensure that no flexible linker domain remained in the RBD. The amino acid sequences were reverse-translated using the Python library “DNA Chisel” following a previously described codon bias ([Supplementary Table S2](#)) [47] and avoiding restriction sites for enzymes MlyI, NcoI, and AgeI, as well as consecutively repetitive 9-bp stretches (Code 8). Flanking primer sequences (Gibson_upstream and Gibson_downstream) were added to the beginning and end of each hit design for later PCR amplification, resulting in 61 DNA sequences.

The DNA sequences were designed to be assembled from synthetic oligonucleotides of fixed length (260 bp), with the following structure from 5' to 3': a primer sequence Oligo_upstream (20 bp), followed by a designed Pa-RBD sequence (220 bp), and ending with the primer sequence Oligo_downstream (20 bp). The two primer sequences were used to amplify the oligonucleotides by PCR and were then removed using the type IIS restriction enzyme MlyI. For Gibson assembly to synthesize library sequences, adjacent oligonucleotides were designed to have 30 bp overlaps. Therefore, assembly of four oligonucleotides could give an RBD up to 229 amino acids. We designed two assembly reactions to create the Pa-RBD library: one comprising the 43 hits that could be assembled from four oligonucleotides (Pa4 assembly) and another comprising the 18 sequences that required five oligonucleotides (Pa5 assembly). Oligonucleotides were designed from the reverse-translated sequences by breaking up the desired RBD sequence into 220 bp-long sequences with 30 bp overlap between adjacent oligonucleotides. The last fragment of the putative RBD was not necessarily 220 bp. If its length was ≤ 64 bp, the sequence was designed to be 64 bp long using the fill_up dummy sequence ([Supplementary Table S1](#)) at matching positions. Then, any last fragments under 220 bp were designed to be 220 bp long by adding a series of T

nucleotides (260 bp total including the PCR primers, Code 9). The sets of 172 (Pa4 oligos, [Supplementary Data File 1](#)) and 90 (Pa5 oligos, [Supplementary Data File 1](#)) oligonucleotides were synthesized (Integrated DNA Technologies). The design process of the Pa-RBD library is summarized in [Supplementary Fig. S2](#).

The second library (“Ino-200RBD” library) created a more diverse set by sampling a larger fraction of the ino-virus metagenome RBDs, using a process summarized in [Supplementary Fig. S3](#). We followed the same strategy as outlined above for the Pa-RBD library, with the following changes. We used 39 characterized ino-virus RBDs (all known ino-virus RBDs as of Sep. 2022; Ino-200RBD ref seqs, [Supplementary Data File 1](#)) as reference sequences for BLAST. No hits were removed based on *E*-values. Based on the relatively poor assembly efficiency we observed from Pa5 compared to Pa4 (see the “Results” section), we constrained the RBD design in Ino-200RBD to the N-terminal 229 amino acid residues of each RBD, which would be amenable to four-fragment assembly (Code 11). RBDs with protein sequences containing “GGG” were removed to avoid glycine-rich domains (Code 12). To limit the number of library members for feasible assembly (~50 library members per pool), we ranked candidate RBDs by a similarity score. Each PIP database member could match ≥ 1 of the 39 reference sequences. A similarity score was calculated as the sum of the reciprocals of *E*-values for matching reference sequences, and the top-scoring 200 candidates were identified (Ino-200RBD designs, [Supplementary Data File 1](#), Code 10). Oligonucleotides were designed and synthesized as described for the Pa-RBD library (Ino-RBD oligos, [Supplementary Data File 1](#)), noting that the dummy sequence was not needed. These 200 candidates were split into four pools (InoN, where $N = 1$ to 4) for assembly (200 oligonucleotides per pool).

Construction of phagemid libraries

The Pa4, Pa5, and four InoN assembly reactions were conducted as follows. Two nanograms of the oligonucleotide pool ($\sim 7 \times 10^7$ copies of each oligonucleotide) were amplified by 15 PCR cycles using primers Oligo_upstream and Oligo_downstream and purified using a QIAquick PCR purification kit (Qiagen). The purified pool was digested by MlyI and purified using a QIAquick PCR purification kit. Each pool was assembled using the NEBuilder® HiFi DNA Assembly Master Mix (New England Biolabs) following the manufacturer’s protocol. In brief, 10 μ L of water containing 0.2 pmol of each digested and purified oligonucleotide was mixed with the reaction master mix and incubated at 50°C for 1 h. The reaction products were amplified by 20 PCR cycles, using primers InsertApm_upstream and InsertApm_downstream, and run on a 1% agarose gel. DNA from the desired band was purified using the QIAquick Gel Extraction Kit.

An AgeI restriction site was introduced on the phagemid vector pADL-10b at the beginning of the pIII glycine-rich linker by mutating amino acid Ser239 to Thr using site-directed mutagenesis as follows. The vector pADL-10b was amplified for 15 PCR cycles using mutagenic primers S239T_Forward and S239T_Reverse, treated with KLD Enzyme Mix, and transformed into chemically competent cells. The cells were grown and plasmid was purified using a miniprep kit. The resulting phagemid vector pADL10b-S239T

and each assembly pool (Pa4, Pa5, and four InoN assembly pools) were double-digested with restriction enzymes NotI and AgeI, purified, ligated using ElectroLigase, and transformed into competent cells. The phagemid libraries were amplified by growth in *E. coli* and collected from cells using the QIAprep miniprep kit (Qiagen). The cells containing phagemid libraries for the four InoN assembly reactions were pooled, making the final Ino-200RBD phagemid library. The Pa4, Pa5, and Ino-200RBD phagemid libraries were linearized by single digestion with NcoI, purified, and characterized by PacBio sequencing. The library preparation and sequencing were conducted by CD Genomics (A Division of CD Biosciences, Inc., NY) Complete Plasmid Sequencing service using PacBio SMRT sequencing (single-molecule, ZMW-based, long-read) for real-time base detection and consensus-based accuracy. The Pa4 and Pa5 libraries were pooled together, creating the Pa-RBD phagemid library.

Characterization of phagemid libraries

To analyze the Pa4, Pa5, and Ino-200RBD libraries from PacBio sequencing data, all reads were first extracted into FASTA file format using the SMRT Link software (version 10.1.0.119588) [48]. Each read was screened by identifying the glycine-rich linker and signal peptide on phagemid vector pADL10b-S239T to isolate the intervening RBD sequence (Code 13). Only DNA sequences with both the signal peptide and the glycine linker with AgeI restriction site were considered valid phagemids from the library, such that an exact match to 33 nucleotides was required. Given an observed fractional yield of valid reads (y), this requirement allows calculation of an upper bound on the effective sequencing error rate (μ); namely, $(1 - \mu)^{33} < y$. For $y = 80\%$, $\mu < 0.67\%$. Reads in which the glycine-rich linker was found as a reverse complement were converted to the forward strand. The sequences with codons in-frame were translated *in silico*, and sequences that were out-of-frame were removed from the libraries (Code 13). Sequencing read databases (SR databases Pa4_read, Pa5_read, and Ino200_read) were generated using BLAST+, containing the translated putative RBDs from PacBio reads of each library. The number of times each unique protein sequence was found in each SR database was determined.

To match each sequence read in the SR databases to sequences in the original design, each read was compared to the corresponding designed phagemid library using BLAST+, and the three designed sequences with the lowest *E*-values were recorded (Code 14). As a control, the same analysis was performed on the set of reference sequences to establish the distribution expected whether the sequence reads conformed perfectly to the design. To represent the sequencing data, force-directed graphs were created using the Javascript D3 graphing library [49]. The force governing the placement of the nodes was related to the *E*-value between the reference and RBD sequences; specifically, *E*-values were scaled by dividing by the mode *E*-value in the dataset followed by a log-transformation, and the minimum value was used as a baseline, according to the following equation (Code 17–21).

$$force \propto \log \left(\frac{E_{value}}{Mode\ E_{value}} \right) - \min \left(\log \left(\frac{E_{value}}{Mode\ E_{value}} \right) \right).$$

Production and purification of phage libraries

One milliliter of frozen glycerol stock of *E. coli* cells transformed with each phagemid library (5×10^8 cells/mL) was inoculated in 10 mL of 2×YT media with 100 µg/mL ampicillin and 1% w/v glucose, and grown at 37°C at 250 rpm until $OD_{600} = 2$. Then 1 mL of the culture was sub-cultured in 45 mL of the same media until $OD_{600} = 0.5$. Fifty microliters of helper phage CM13d3 at 1×10^{13} virions/mL were added to the culture for superinfection and cultured for 30 min, followed by supplementation of isopropyl β-D-1-thiogalactopyranoside (IPTG) and kanamycin to 200 µM and 50 µg/mL, respectively, to induce recombinant pIII expression to produce phages. The cells were cultured overnight at 25°C with shaking at 200 rpm and then pelleted by centrifugation at $7000 \times g$ for 10 min. The supernatant was filtered through a 0.22 µm nitrocellulose filter, and the filtrate was concentrated to 2 mL volume and buffer-exchanged to 1× PBS buffer using an Amicon Ultra-15 Centrifugal Filter 50 kDa MWCO (Millipore). The phage concentration was determined by spectrometry following the equation $\frac{(A_{269} - A_{320}) \times 6 \times 10^{16}}{\text{number of bases phage genome}} = \frac{\text{virions}}{\text{mL}}$ and confirmed by qPCR using primer pair qPCR_upstream and qPCR_downstream [50, 51]. The phages were diluted to 1×10^{13} virions/mL with 1× PBS buffer and stored at 4°C for short-term storage.

Phage selection for binding bacterial cells

To identify members of the Pa-RBD and Ino-200RBD phagemid libraries that bind target pathogens, two rounds of selection were performed as follows. Four milliliters of log-phase cell culture ($OD_{600} = 1$) of the target pathogens were grown in LB media at 37°C [*P. aeruginosa* strains: ATCC 25102, engineered strain PAK pmrB6 (Pm^r), and 15 clinical strains; *E. coli* strains: DH5α and ATCC 700927; *A. baumannii* strain ATCC 19606; *B. cepacia* strain: ATCC 25416; *C. sakazakii* strain: ATCC 25944; and *K. pneumoniae* strains: ATCC 700603, and four clinical strains] (Table 1). Cells were pelleted by centrifugation at $5000 \times g$ for 5 min, washed twice in 1× PBS buffer, and resuspended in 1× PBS at the original volume [36, 52]. One hundred microliters of purified phage at 1×10^{13} virions/mL were mixed with the cells and incubated on a rotator for 30 min at room temperature. The cells were spun down at $3000 \times g$ and washed with 5 mL of 1× PBS buffer with 1% w/v bovine serum albumin (BSA) three times. The cells and any bound phages were pelleted and DNA was extracted from the pellet using a Qiagen miniprep kit. The DNA was eluted in 25 µL of Milli-Q® water, and phage sequences were amplified by 15 PCR cycles using back-to-back primers Orf1F1longForward and Orf1F1longReverse. The PCR products were purified using a QIAquick PCR Purification Kit, circularized by ligation using KLD enzyme mix, and transformed into competent cells (*E. coli* strain TG1). The transformed cells were cultured to produce phages and the selection steps were repeated as described earlier. After the second round of selection, the transformed cells were plated on LB agar plates supplemented with 100 µg/mL ampicillin and incubated at 37°C overnight. Sixty colonies were picked from each library selection and sent for Sanger sequencing using primers Sequencing_upstream and Sequencing_downstream.

qPCR phage-bacteria binding assay

The quantity of phages binding to different bacterial targets in each sample and the concentration of phage stocks were

determined using qPCR. For stock solutions, dilution to 100 µL was used for the sample. To quantify the amount of phage binding to different bacterial cell targets, a known quantity of phages in 100 µL volume was mixed with 1 mL of bacterial cells in the exponential growth phase at a concentration of $OD_{600} = 1$ and incubated for 30 min at room temperature. The samples were centrifuged at $4000 \times g$ for 5 min, washed two times with 1× PBS buffer, and resuspended in 1× PBS buffer back to the original volume.

For qPCR, samples were processed by the QIAprep Spin Miniprep Kit (Qiagen; standard protocol) to lyse bacterial cells and/or phage capsids and collect phagemid DNA. DNA was eluted from the spin columns in 50 µL of MilliQ water. Phagemid in the eluted samples was quantified using qPCR with a standard curve made from a known concentration standard of purified phagemid vector pADL-10b. The qPCR assay was carried out in 10 µL reaction volumes with triplicates for samples and duplicates for standards using SYBR Green Master Mix (Bio-Rad Laboratories) with primer set qPCR-upstream and qPCR-downstream on a Bio-Rad C1000 PCR machine. PCR conditions: 95°C for 30 s, then 45 cycles of 30 s at 95°C and 15 s at 60°C, followed by melt curve analysis (10 s at each increment using 0.5°C increments up to 95°C).

M13^{PAB} deletion mutants

Based on the predicted structure of PAB protein by AlphaFold2, four regional deletion mutants of PAB were designed: N-terminus deletion mutant (dNT-PAB), α-helix deletion mutant (dα-PAB), β-sheet deletion mutant (dβ-PAB), and C-terminus deletion mutant (dCT-PAB). Primer pairs used for PCR mutagenesis of the pADL-10b derivative expressing PAB were: dNT-PAB-F with dNT-PAB-R, dα-PAB-F with dα-PAB-R, dβ-PAB-F with dβ-PAB-R, and dCT-PAB-F with dCT-PAB-R, respectively (Supplementary Table S1). PCR mutagenesis was performed using Q5 Hot Start High-Fidelity 2× Master Mix (New England Biolabs) and amplicons were circularized by a kinase, ligase, and DpnI (KLD) enzyme mix (New England Biolabs). Plasmids were verified by whole-plasmid sequencing and subsequently transformed into Mix & Go TG1 cells (Zymo Research), which were grown on ampicillin-containing LB agar plates. Cultures of TG1 cells transformed with each deletion mutant were stored as 10% glycerol stocks at −80°C until further use. Chimeric M13 phages expressing these deletion mutants of PAB as RBDs were produced and quantified as described earlier.

Transmission electron microscopy

Samples were prepared for transmission electron microscopy (TEM) of bacterial cells, phages, and phages bound to bacterial cells following a previously developed protocol [53]. Cell samples were prepared by inoculating an overnight culture into 10 mL fresh MH broth (1:20 v/v) and allowing the culture to grow to $OD_{600} = 0.5$. The cells were washed twice with 1× PBS buffer and resuspended to 1/5 of the original volume. Phage samples were prepared at 1×10^{13} virions/mL. To image phages bound to cells, 100 µL of phages were mixed with 100 µL of the bacterial suspension and incubated for 30 min, washed two times using 1× PBS buffer with 1% w/v bovine serum albumin (BSA), and resuspended to 100 µL volume. The TEM sample grids were prepared at room temperature by loading 8 µL of each sample on a Formvar/Carbon 200 Mesh, Ni TEM grid (Electron Microscopy Science) and incubating for 2 min. The sample grid was washed four times with

washing solution (1× PBS buffer with 1% BSA) and blocked with blocking solution (1× PBS buffer with 0.5% w/v gelatin) for 1 h at room temperature. To label the phages, the sample grid was incubated with primary anti-M13pVIII antibody (Thermo Fisher Scientific) for 1 h, washed with washing solution four times, incubated with gold nanoparticles (6 nm dia.) coated with donkey anti-mouse IgG (Abcam) for 1 h, and washed with water five times. The sample was negatively stained with 8 µL of 1% uranyl acetate for 1.5 min, dried with a Kimwipe, and imaged on an FEI Tecnai T12 transmission electron microscope.

Synthesis of colistin-M13^{PAB}

Primary amines on M13^{PAB} phage were blocked by mixing 1 mL of phage at 1×10^{13} virions/mL in 1× PBS buffer with 10 mM Sulfo-NHS-Acetate on a rotator for 2 h at room temperature. The sample was dialyzed for 2 h three times using 1× PBS buffer and two times with 0.1 M MES (pH 5.5) buffer at 4°C, using a Slide-A-Lyzer dialysis cassette (20K MWCO, Thermo Fisher Scientific). The phage product was then incubated with a mixture of EDC (20 mM) and Sulfo-NHS (50 mM) in 100 µL of water for 20 min at room temperature for activation. The sample was dialyzed three times in 1× PBS (pH 7.2) buffer for 40 min at 4°C and was mixed with 1× PBS to reach 1×10^{12} virions/mL concentration. One hundred microliters of colistin sulfate (0.75 mg/mL) was mixed with 1 mL of activated phage and incubated at room temperature for 2 h on a rotator. The product, colistin-M13^{PAB}, was purified using dialysis for 4 h two times with 1× TBS buffer and three times with 1× PBS buffer at 4°C, then filtered through a 0.22 µm nitrocellulose membrane filter. The concentration of colistin-M13^{PAB} was determined using UV spectrometry as described earlier, and the sample was stored at 4°C for short-term storage up to 12 weeks. For amino acid composition analysis, 1×10^{13} virions of colistin-M13^{PAB} in total were synthesized in four separate batches and pooled together into one sample tube for analysis by Creative Proteomics (Shirley, New York).

Minimal inhibitory concentration and minimal bactericidal concentration assay

To characterize the colistin and colistin-M13^{PAB} minimal inhibitory concentration (MIC) and minimal bactericidal concentration (MBC), we followed an established microdilution protocol [54, 55]. In a nontreated, U-shaped bottom 96-well plate (Fisher Scientific), 5×10^5 CFU of log-phase bacterial cells in Mueller–Hinton (MH) broth were added to each well except wells in the last column. Colistin or colistin-M13^{PAB}, serially diluted in MH broth in two-fold dilutions, was added to each well ($N = 3$ replicates for each reagent at each concentration). In the second to last column, 50 µL of MH broth without antibiotics was added to wells as a growth control ($N = 3$). In the last column, 100 µL of MH broth was added to wells as sterility controls ($N = 3$). The sterile covered plates were incubated overnight without shaking at 37°C. The cell growth in each well was measured by optical density at 600 nm on a plate reader (Tecan Group Ltd). A lack of visible cell pellet was observed to always correspond to a <0.2 increase of OD₆₀₀ value in these experiments, and these observations were interpreted as no cell growth in a well. The well containing the lowest concentration of each antibiotic resulting in no growth indicated the MIC. To determine the MBC, 50 µL of

the culture at the MIC and higher reagent concentrations were spread onto MH agar plates and incubated overnight at 37°C. The plates without colony growth indicated the reagent was bactericidal at that concentration, and the lowest bactericidal concentration gave the MBC.

In vitro hemolysis assay

To characterize the hemolysis activity of colistin and colistin-M13^{PAB} at serially diluted concentrations, an established hemolysis assay protocol was followed [55, 56]. Five milliliters of 10% sheep red blood cells (MP Biomedicals) were washed five times with 1× PBS buffer by spinning down at $500 \times g$ for 5 min on a benchtop centrifuge, discarding supernatant, and gently resuspending to the original volume in 1× PBS buffer. After the last washing step, the sample was diluted to a 4% red blood cell solution and aliquoted into 1.5 mL Eppendorf tubes (100 µL in each tube). One hundred microliters of colistin or colistin-M13^{PAB} at 2× the desired final concentration, 1× PBS for negative control, or 2% Triton X-100 for positive control, was added to each tube ($N = 3$). The mixtures were incubated for 1 h at 37°C. After incubation, samples were spun down at $500 \times g$ for 5 min, and 100 µL of supernatant from each tube was transferred into a 96-well plate. Absorbance at 405 nm was measured using a plate reader and the hemolytic activity was determined using the following equation:

$$\% \text{ hemolysis} = (A_{\text{sample}} - A_{\text{PBS}}) / (A_{\text{Triton X-100}} - A_{\text{PBS}}),$$

where A_{sample} , A_{PBS} , and $A_{\text{Triton X-100}}$ are absorbance values at 405 nm of the experimental sample, negative control, and positive control, respectively.

In vitro mammalian cell toxicity assay

The toxicity of colistin and colistin-M13^{PAB} on mammalian cell cultures was assessed using human embryonic kidney (HEK-293) cells (ATCC CRL-1573) following a previously established protocol [57, 58]. The cells were revived from liquid N₂ storage (90% FBS, 10% dimethyl sulfoxide (DMSO)) by thawing in a 37°C water bath, pelleting, and resuspending in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% antibiotics (penicillin/streptomycin) in Falcon® polystyrene tissue culture flasks with ventilated caps (Corning). The cells were incubated at 37°C with 5% CO₂ until 80%–90% confluency. Then, cells were detached from the culture flask using trypsin-EDTA and seeded at a density of 1×10^4 cells/well in 24-well plates (Corning). The cells were then treated with DMEM containing 1% antibiotics (penicillin/streptomycin) and supplemented with either colistin sulfate (2, 8, or 64 µg/mL) or colistin-M13^{PAB} (5×10^9 , 1×10^{10} , or 2×10^{10} virions/mL). HEK-293 cells were also cultured in DMEM with or without penicillin/streptomycin as controls ($N = 3$). The media was refreshed every 2 days.

HEK-293 cell viability, indicated by metabolic activity, was quantified using 3-(4,5-dimethylthiazolyl-2)-2,5-diphenyltetrazolium bromide MTT assay (Millipore Sigma) following the manufacturer's protocol. After 1, 3, and 5 days of culture, 100 µL of 5 mg/mL MTT stock solution was added to the wells, incubated for 4 h at 37°C, and mixed with 1 mL of solubilization solution and incubated at 37°C overnight. The absorbance of the samples was read at 550 nm (reference

wavelength at 650 nm) using a Tecan Infinite M200 PRO plate reader (Tecan Group Ltd).

HEK-293 cell spreading and morphology were observed by staining the cells with Alexa Fluor 594 phalloidin and 4',6-diamidino-2-phenylindole dihydrochloride (DAPI) (Invitrogen) staining at room temperature. After 1, 3, and 5 days of culture, the cells were fixed using 4% (v/v) paraformaldehyde for 10 min, washed with 0.3% (v/v) Triton-X for 10 min, and blocked with 1× PBS buffer containing 1% BSA for 30 min at room temperature. A light-sensitive mixture of 0.1% (v/v) Alexa Fluor 594 phalloidin and 0.05% (v/v) DAPI in 1× PBS was added to stain the cells for 20 min. All procedures were conducted at room temperature. After washing with 1× PBS for two times, the cells were imaged with an AxioObserver Z1 inverted microscope (Zeiss) ($N = 2$).

Results

Construction of phagemid libraries

To construct libraries of recombinant M13 phages with RBD sequences from the *Inoviridae* family, we searched for sequences with similarity to known *Inovirus* RBDs in a putative *Inovirus* protein database (“PIP” database). The PIP database was created from protein sequences encoded by 10 295 *Inovirus*-like genomes previously identified using machine learning from published metagenomic databases [33]. Using two *P. aeruginosa* filamentous phage RBPs as references, we designed a “Pa-RBD” library theoretically consisting of 61 protein sequences. Using all known *Inovirus* RBPs (39 sequences as of Sep. 2022; [Supplementary Data File 1](#)) as reference sequences, we identified the top 200 candidates for an “Ino-200RBD” library design.

To synthesize the sequences, we conducted several pooled oligonucleotide assembly reactions (Scheme 2). Each reaction contained gene fragments with 30 bp overlapping regions between adjacent fragments that would be assembled into multiple full-length genes using Gibson assembly [59]. Two assembly reactions were used to synthesize the Pa-RBD library: one for the 43 sequences that could be assembled from four oligonucleotides each (Pa4 library) and one for the 18 sequences that required five oligonucleotides each (Pa5 library). The larger Ino-200RBD library was synthesized using only four-oligonucleotide assembly, truncating sequences past the first 229 amino acids of each RBD, using four reaction pools with each pool theoretically assembling 50 sequences. The Gibson assembly reactions resulted in DNA of approximately the expected lengths ([Supplementary Fig. S4A](#)). The assembled products were amplified by PCR, digested, cloned into phagemid vector pADL10b-S239T, and transformed into *E. coli* competent cells.

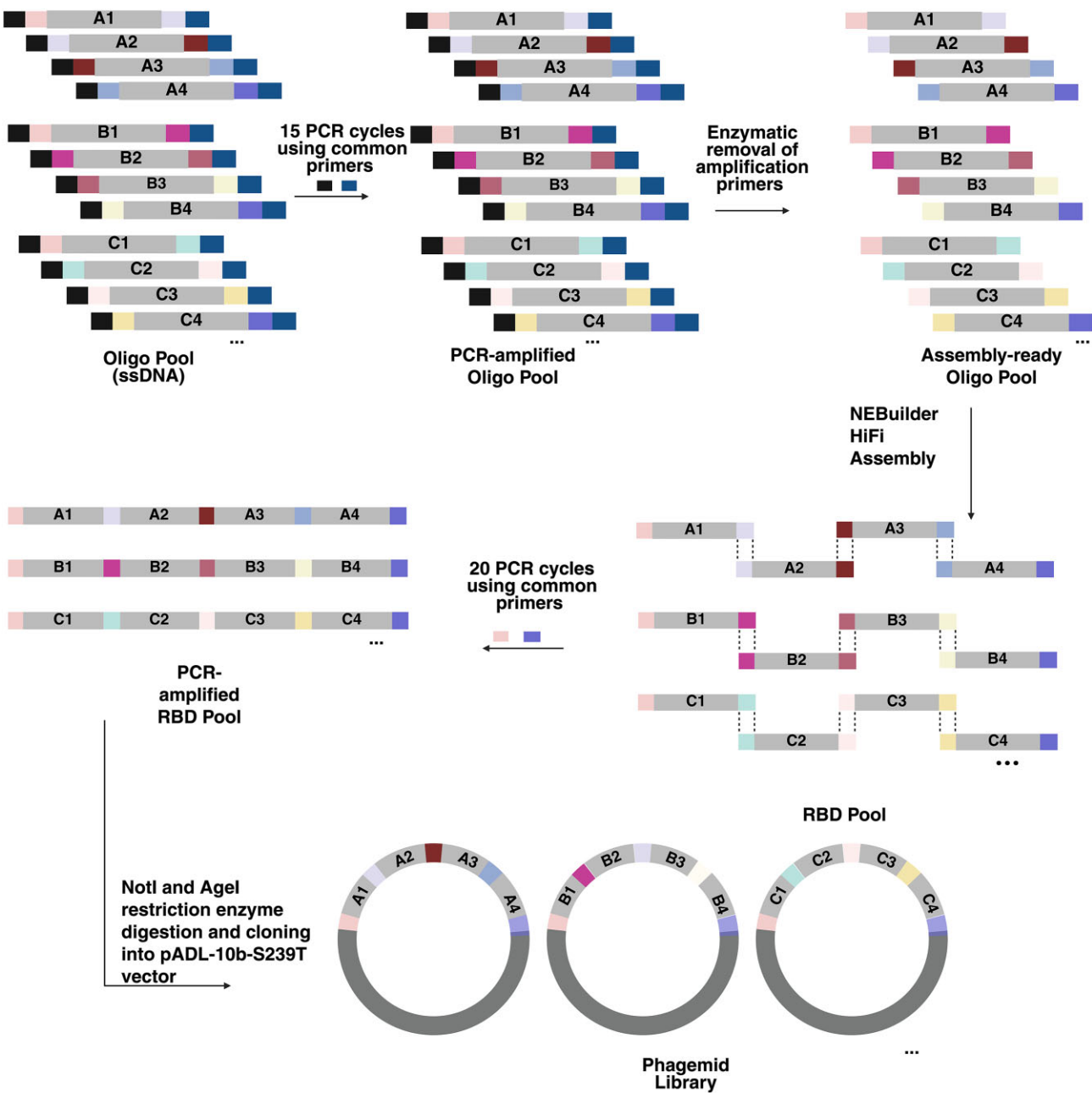
DNA from the Pa4, Pa5, and Ino-200RBD phagemid libraries was characterized using PacBio sequencing (see the “Data Availability” section). Valid reads were defined as reads demonstrating an in-frame protein placed between the signal peptide and the glycine-rich linker of the phagemid vector. While the Pa4 and Ino-200RBD libraries gave a high yield of valid reads (~80%), the Pa5 library gave a lower yield (7.5%) ([Supplementary Fig. S5](#)), suggesting inefficient assembly of five fragments.

The length distribution of valid reads roughly matched the design for the three libraries ([Supplementary Fig. S6](#)). To determine whether the designed sequences were assembled within

the libraries, the Levenshtein distance was calculated for each read to the reference sequences. For the Pa4, Pa5, and Ino-200RBD libraries, 83.7%, 66.7%, and 34.5% of the designed sequences had a Levenshtein distance of <100 bp to a sequence read in the corresponding library, respectively. However, the majority of reads did not exactly match a designed sequence, indicating a substantial amount of unintended fragment combinations and/or mutations (Fig. 1); indeed, 0/18, 26/43, and 30/200 of the designed sequences were observed in the PacBio sampling of the Pa4, Pa5, and Ino-200RBD libraries, respectively. Of the valid reads, the vast majority were unique (>99.9%; 29 047 out of 29 056 reads) in the Pa5 library, compared to 55.7% (16 956 out of 30 420 reads) in the Pa4 library and 23.8% (15 380 out of 64 583 reads) in the Ino-200RBD library. Overall, library synthesis using pooled assembly resulted in an unexpectedly diversified set of sequences. While the library design was based on the *Inovirus* metagenome, the synthesized library contained many additional recombinants. While some of the designed sequences were observed in the library, at the same time, most of the synthesized sequences increased the library complexity beyond the original design.

In vitro selection of recombinant phage binding to *P. aeruginosa* strains

Pa4 and Pa5 were combined into a single library, Pa-RBD. To identify *P. aeruginosa*-binding clones, phages were produced from the Pa-RBD library (1×10^{12} virions) and incubated with $\sim 4 \times 10^8$ CFU of bacteria from each of 15 different *P. aeruginosa* strains (ATCC 25102 and 14 clinical isolates; Table 1). Cells were spun down and washed, and unbound phages were discarded. Phagemid DNA from the pellet was extracted, amplified, and transformed into competent *E. coli* cells to produce phages for a second round of selection. After the second round, phages binding to *P. aeruginosa* strain ATCC 25102 and nine of the clinical isolate strains (A, B, C, E, F, G, J, K, L) were detected by PCR amplifying the phagemid ([Supplementary Fig. S4B](#)). However, five clinical isolates (D, H, I, M, and O) did not yield a visible PCR band, indicating no detectable phage was bound to these cells after two rounds of selection. For each of the 10 strains exhibiting bound phages, we randomly sampled 60 phagemid clones through Sanger sequencing. A single member of the Pa-RBD library (phage designated “M13^{PAB}”) dominated the bound phages across strains, constituting >90% of sampled clones, indicating that M13^{PAB} could recognize multiple strains of *P. aeruginosa*. M13^{PAB} showed similarity to two Pa-RBD designs (24 937 and 50 842 in [Supplementary Data File 1](#); Fig. 1B) in the Pa5 library, suggesting that it arose as a recombinant of these two designs (Fig. 2A and B). M13^{PAB} phages were produced, purified, and verified for binding by qPCR to the six *P. aeruginosa* strains that had yielded M13^{PAB} during the selection (ATCC 25102 and clinical strains A, B, C, E, and F). To test the breadth of activity, M13^{PAB} was also assayed for binding to the polymyxin-resistant strain PAK *pmrB6* and to clinical strain M, against which M13^{PAB} was not initially selected (Fig. 2C). M13^{PAB} phage exhibited >100-fold more binding to each *P. aeruginosa* strain compared to a negative control (*Saccharomyces cerevisiae*). TEM images confirmed that M13^{PAB} phage was bound to the bacterial cell wall through one terminus of the virion structure and not bound via the sidewall, consistent with binding mediated by g3p (Fig. 3).



Scheme 2. Construction strategy of phage RBD libraries. Colored blocks represent matching sequences among the oligonucleotides. Each oligo in the initial single-stranded DNA (ssDNA) pool contained five regions from 5' to 3': an upstream PCR amplification domain (black), an upstream Gibson assembly overlap region (various colors), a nonoverlapping RBD-encoding region (gray, with RBD “A” designed to be assembled in order from segments A1 to A4, and so forth), a downstream Gibson assembly overlap region (various colors), and a downstream PCR amplification domain (navy). The ssDNA oligo pool was amplified by PCR using primers targeting the 5' (black) and 3' (navy) regions. The resulting products were digested using type IIS restriction enzymes to remove the PCR primer regions, leaving blunt ends. The fragments were then assembled by Gibson assembly (NEBuilder HiFi DNA Assembly), which in theory would produce the RBDs as illustrated. The assemblies were PCR-amplified using primers targeting the 5' and 3' ends (pink and purple), digested with NotI and AgeI, and cloned into the phagemid vector pADL-10b-S239T for transformation into *E. coli*. Triple dots represent the presence of more RBD sequences progressing through the depicted workflow. The schematic shows four-oligo assembly per RBD, as used in Pa4 and the Ino-200 libraries, while the Pa5 library used five-oligo assembly. Created in BioRender. Chen, I. (2025) <https://BioRender.com/kyjhhqhk>.

Effect of sequence deletions on M13^{PAB} phage binding activity

The RBD of M13^{PAB} was a 92-mer (MW: 10 kDa) peptide termed the “PAB” protein here, which is expected to be negatively charged (theoretical pI = 4.16). Using AlphaFold2, PAB was predicted to be composed of four domains: an unstructured N-terminal domain (residues 1–20), an α -helix (residues

21–35), a β -sheet (41–73), and a disordered C-terminal region (residues 74–92) [60–62] (Fig. 4A and [Supplementary Table S6](#)). To identify the region important for *P. aeruginosa* binding, four deletion mutants were created, corresponding to the deletion of each domain: M13^{PAB}-dNT, M13^{PAB}-d α , M13^{PAB}-d β , and M13^{PAB}-dCT, respectively. The production of phage virions from each mutant was confirmed by UV-

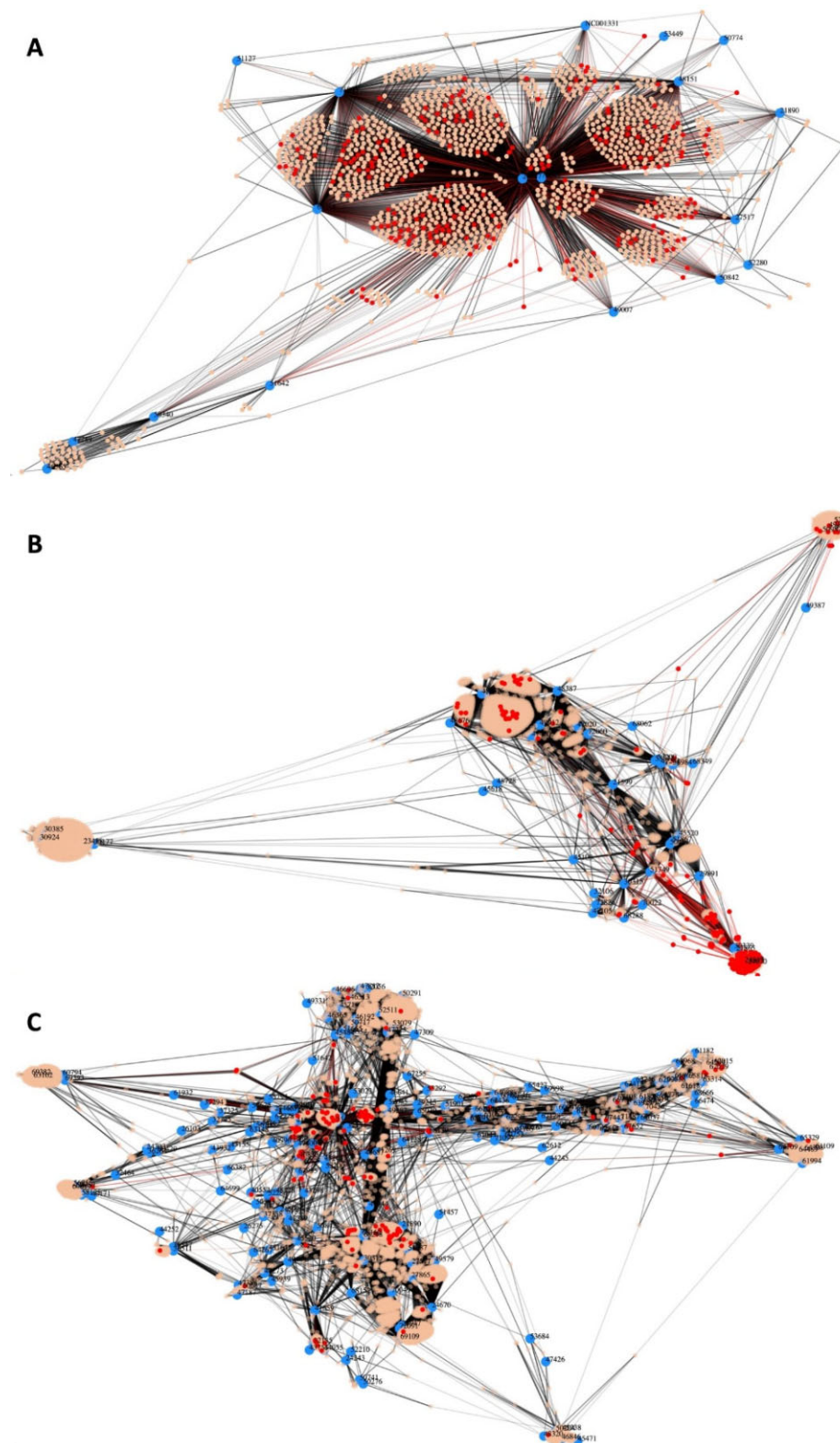


Figure 1. Representation of sequence space for the designs in the Pa4 (A), Pa5 (B), and Ino-200RBD (C) libraries co-registered with the PacBio sequencing reads. Force-directed graphs were based on sequence similarity. Blue nodes (with number labels referencing the PIP database; [Supplementary Data File 1](#)) represent the designed sequences; tan nodes represent a randomly selected subsample of sequencing reads [5%, 25%, and 25% of the valid reads in panels (A–C), respectively]. Each read was matched to the three most similarly designed RBD sequences using BLAST (lines; darker = more similar). Red nodes represent sequences similar to sequences selected for binding. M13^{PAB} was the only sequence selected out of the Pa4 (A) and Pa5 (B) libraries. No exact match of M13^{PAB} was found in the Pa4 and Pa5 reads. The BLAST e-value between each read and M13^{PAB} was calculated, and reads with e-values in the lowest quartile were highlighted as red nodes. The red nodes are concentrated around designs 24 937 and 24 670 in the Pa5 library, suggesting an origin of M13^{PAB} as a recombinant of these two designs. For the Ino-200RBD library (C), multiple sequences were selected for binding, and sequences exactly matching the selected sequences were highlighted using red nodes.

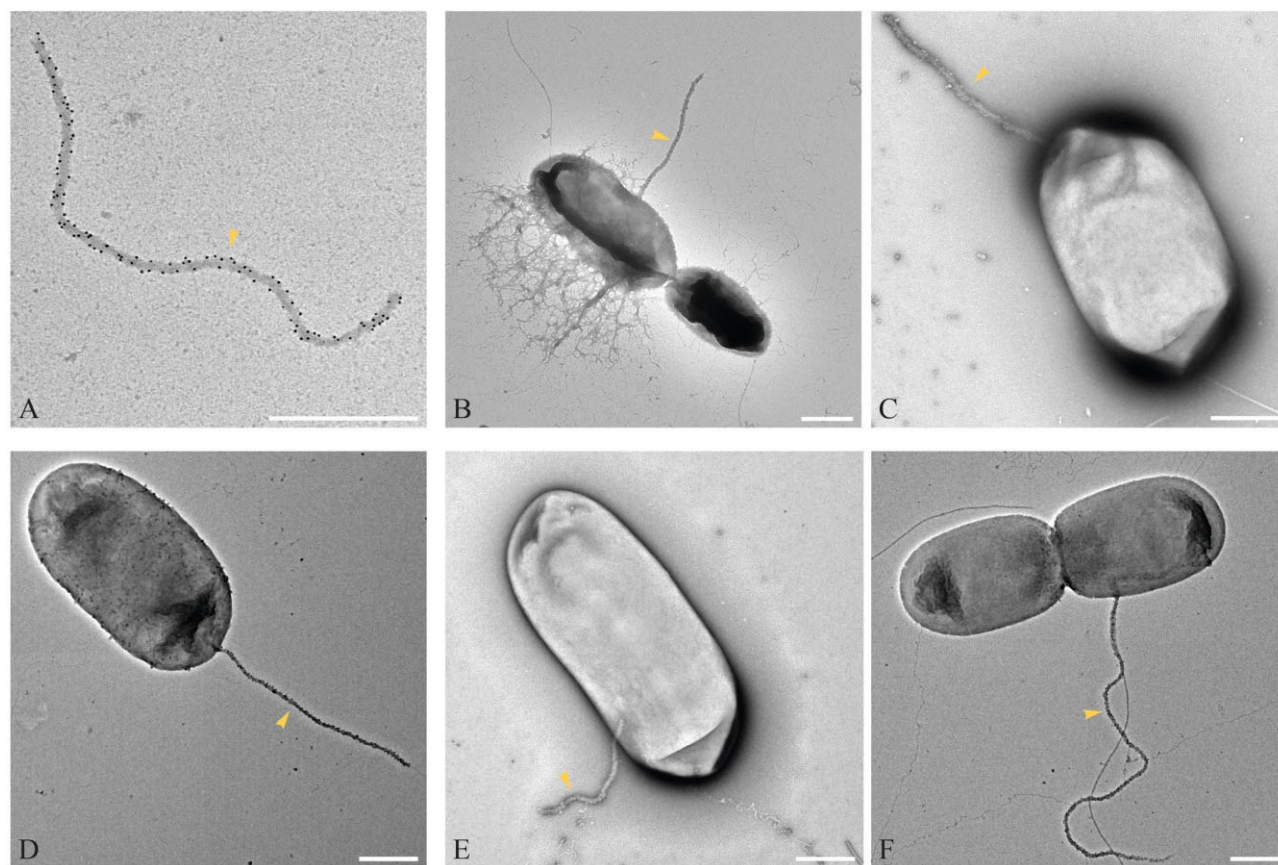


Figure 3. M13^{PAB} phage bound to different *P. aeruginosa* strains visualized by TEM (negative stain). Since filamentous phages have similar morphology to bacterial structures such as pili, phages were specifically labeled with gold nanoparticles (dark spheres) coated with antibodies against pVIII, allowing phages (yellow darts) to be easily differentiated from other structures. Gold-labeled M13^{PAB} phages are shown (A) without bacteria; or incubated with *P. aeruginosa* strain: (B) ATCC 25102; (C) clinical strain A; (D) clinical strain B; (E) clinical strain C; and (F) polymyxin-resistant mutant PAK *pmrB6*. Note that phage width is increased by the labeling reagents. Scale bars = 500 nm.

vis absorption spectrometry. Binding activity of WT and mutant M13^{PAB} phages to *P. aeruginosa* strain ATCC 25102 was measured by incubating cells with phages, centrifuging and washing, and quantifying bound phage by qPCR. M13^{PAB}-d α , M13^{PAB}-d β , and M13^{PAB}-dCT demonstrated similar binding as WT M13^{PAB}, but M13^{PAB}-dNT binding was reduced by >10-fold, indicating that the N-terminal domain was important for binding the *P. aeruginosa* cells (Fig. 4B).

Conjugation of colistin to M13^{PAB} phage

To test the ability of M13^{PAB} phage to deliver antimicrobial cargo to *P. aeruginosa*, we conjugated colistin to synthesize “colistin-M13^{PAB}.” Colistin contains primary amines from five 2,4-diaminobutyric acid (Dab) residues, while the major coat protein of M13 and M13^{PAB} phage, g8p, has three solvent-accessible carboxyl groups. M13^{PAB} phage was pretreated to block reactive amines to prevent phage cross-linking and then conjugated to colistin via EDC chemistry. The colistin-M13^{PAB} product was purified by dialysis and sterile filtration.

The amount of colistin conjugated per virion in colistin-M13^{PAB} was determined using amino acid composition analysis to measure the mole percentage of Thr and Leu residues, which are part of the colistin lipopeptide. The measured composition was compared to expectations calculated for varying stoichiometry of colistin molecules conjugated per g8p

protein (Supplementary Fig. S7, Supplementary Table S3, and Supplementary Text S1) [27]. The analysis indicated that on average, ~2.5 colistin molecules were conjugated per copy of pVIII, or 6750 colistin molecules per virion (1.3×10^{-11} μ g of colistin per virion). This measured ratio was used to calculate the concentration of colistin-M13^{PAB} given the phage concentration.

Antimicrobial activity of colistin-M13^{PAB} on *P. aeruginosa*

The antimicrobial effect of colistin-M13^{PAB} was tested on several strains of *P. aeruginosa*, including seven clinical isolate strains, and compared to colistin sulfate. MIC was determined as the lowest concentration of colistin or colistin-M13^{PAB} that prevented visible cell growth. From each culture lacking observable cell growth, 50 μ L were plated to determine the MBC. To validate the MIC test, a reference strain of *E. coli* (ATCC 25922) was tested and found to have an MIC of 0.5 μ g/mL, which is within the acceptable range (0.25–1 μ g/mL) [63, 64]. For colistin, strains with MIC $\leq$ 2 μ g/mL are considered sensitive, and strains with MIC $\geq$ 4 μ g/mL are considered resistant [64, 65].

The MIC measurements showed that the colistin-sensitive *P. aeruginosa* strains were also sensitive to colistin-M13^{PAB} (Table 2). Moreover, colistin-M13^{PAB} reduced the MIC as well as the MBC (Supplementary Table S4) by >1–2 orders of mag-

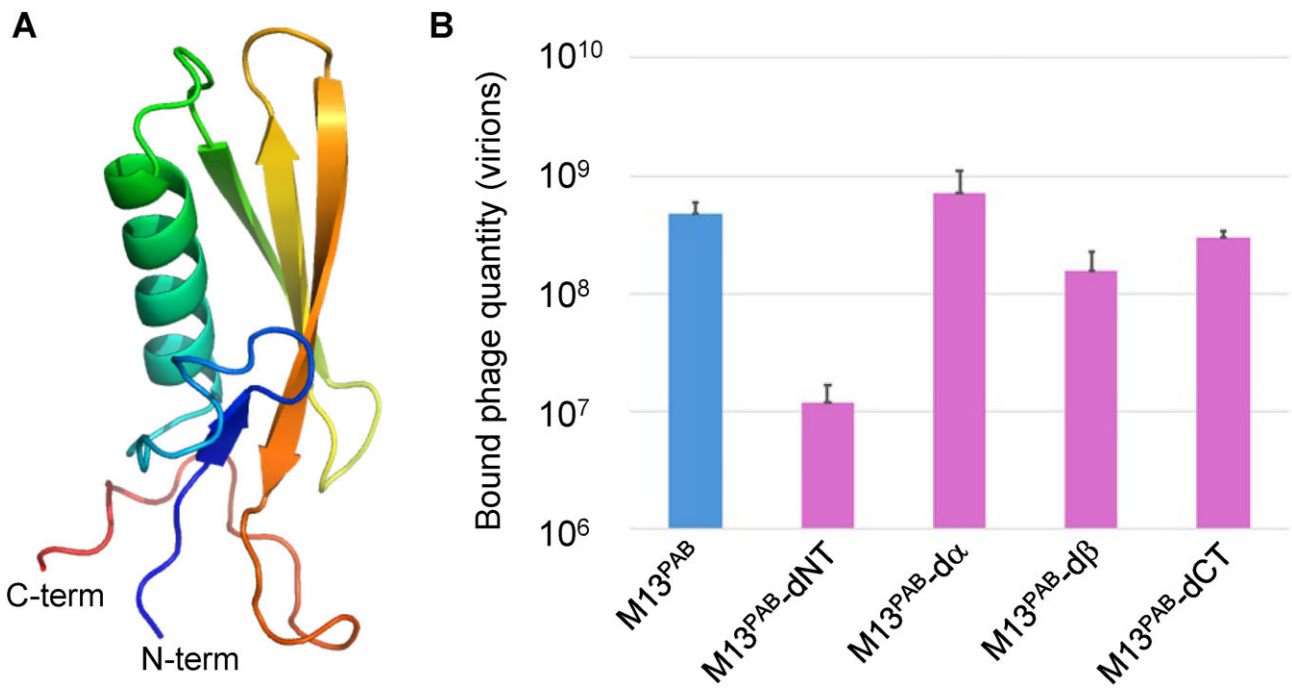


Figure 4. PAB domain effects on binding to *P. aeruginosa*. **(A)** The highest-ranked predicted structure of PAB protein by AlphaFold2, with the N-terminus and C-terminus indicated. **(B)** Binding of M13^{PAB} and mutants M13^{PAB}-dNT (deleting N-terminal domain), M13^{PAB}-dα (deleting alpha helix), M13^{PAB}-dβ (deleting beta sheet), and M13^{PAB}-dCT (deleting C-terminal domain) to *P. aeruginosa* ATCC 25102 when mixing 1×10^{11} virions of phages with 1 mL of cells at OD₆₀₀ = 1 (OD₆₀₀ = 1 corresponds to $\sim 10^8$ cells/mL) and incubating for 30 min at room temperature on a rotator. The amount of bound phage was measured by quantitative PCR. Bar height is the average of technical quadruplicate data on experimental duplicates, and error bars show 1 standard deviation. Significant loss of binding is caused by deletion of the N-terminal domain ($P < .0001$, two-sided Welch *t*-test).

nitude. For the colistin-resistant strain PAK*pmrB*6, colistin-M13^{PAB} decreased the MIC by 63-fold, bringing the MIC value into the sensitive range. Thus, colistin-M13^{PAB} substantially reduces the amount of colistin required to inhibit and kill *P. aeruginosa*.

In vitro hemolysis and cytotoxicity assays

Hemolysis activity was tested for colistin (2 to 64 μg/mL) and colistin-M13^{PAB} (2.5×10^9 to 8×10^{10} virions/mL) by incubation with sheep red blood cells. Materials exhibiting <2% hemolysis are considered nonhemolytic [66]. Similar to previous reports [67, 68], we observed no hemolytic activity for colistin at the concentrations tested, although hemolysis approached 2% at the highest concentration of colistin (64 μg/mL) (Supplementary Fig. S8A). For colistin-M13^{PAB}, no hemolytic activity (<0.4%) was observed over the concentration range, including the MIC_{CMP} for the colistin-resistant strain tested (PAK*pmrB*6) (Supplementary Fig. S8B).

The effect of colistin and colistin-M13^{PAB} on mammalian cell viability was tested by the MTT assay using human embryonic kidney cells (HEK-293; ATCC CRL-1573). Cell viability >80% in the MTT assay is an accepted threshold for a condition to be considered nontoxic [69]. No detectable *in vitro* toxicity was observed up to 64 μg/mL colistin or up to 2×10^{10} virions/mL of colistin-M13^{PAB}, compared to the negative control, up to day 5 (Supplementary Fig. S8C and D). These results indicate compatibility of colistin-M13^{PAB} with mammalian cells *in vitro*.

Stability of colistin-M13^{PAB} in refrigerated storage

The MIC of the same batch of colistin-M13^{PAB}, stored in 1× PBS buffer at 4°C, was tested weekly on *P. aeruginosa*

strain ATCC 25102 for 12 consecutive weeks. The MIC stayed within a factor of two over the 12 weeks (>50% activity retained) (Supplementary Fig. S9).

In vitro selection of recombinant phages binding to several Gram-negative species and clinical isolates

To test the generalizability of phage selection from a metagenome-inspired library, we conducted two rounds of *in vitro* selection from the Ino-200RBD library for binding to each of the following 11 Gram-negative strains: *E. coli* DH5α, *E. coli* (ATCC 700927), *P. aeruginosa* (ATCC 25102), *A. baumannii* (ATCC 19606), *B. cepacia* (ATCC 25416), *C. sakazakii* (ATCC 25944), *K. pneumoniae* (ATCC 700603), and *K. pneumoniae* clinical strains (A, B, C, and D). Bound phages were isolated for each strain (Fig. 5; Supplementary Data File 1). The lengths of the RBDs of the selected phages ranged from 23 to 251 amino acids, with most RBDs being around 100 amino acids long (Supplementary Fig. S10). Multiple sequence alignments were conducted using COBALT [70] on RBDs of all selected phages from both libraries (Pa-RBD and Ino-200RBD) to identify homology across the selected phages (Supplementary Data File 2). A plurality 435-residue consensus sequence was obtained in Jalview [71] (most frequent residue per column, ignoring gaps; Supplementary Data File 3). Within this, a 128-residue fragment (residues 181 to 308) showed homology by BLASTP (nonredundant protein sequences database) to *P. aeruginosa* attachment proteins (Supplementary Data File 4). Some phage RBDs, such as IB1 and IB4, were isolated from selections against multiple strains (Fig. 5), suggesting a broad binding range among Gram-negative organisms. At the same time, other phages were only isolated from a single selection (e.g. phages IB2 and IB3), suggesting higher specificity.

Table 2. Minimum inhibitory concentration (MIC) of M13^{PAB} phage (MIC_{MP}), colistin (MIC_{col}), and colistin-M13^{PAB} (MIC_{CMP})

Species	Strain	MIC _{MP} (virions/mL)	MIC _{col} (μg/mL)	MIC _{CMP}			MIC _{col} MIC _{CMP} (n-fold reduction in MIC)
				(virions/mL)	(μg/mL) ^a		
<i>E. coli</i>	ATCC 25922		0.5				
<i>P. aeruginosa</i>	ATCC 25102	> 1 × 10 ¹²	1	1.25 × 10 ⁹	0.016	63	
<i>P. aeruginosa</i>	PAKpmrB6	> 1 × 10 ¹²	64	8 × 10 ¹⁰	1.036	63	
<i>P. aeruginosa</i>	Clinical Strain A	> 1 × 10 ¹²	0.5	1.25 × 10 ⁹	0.016	31	
<i>P. aeruginosa</i>	Clinical Strain B	> 1 × 10 ¹²	0.5	6.25 × 10 ⁸	0.008	63	
<i>P. aeruginosa</i>	Clinical Strain C	> 1 × 10 ¹²	0.5	6.25 × 10 ⁸	0.008	63	
<i>P. aeruginosa</i>	Clinical Strain E	> 1 × 10 ¹²	0.5	6.25 × 10 ⁸	0.008	63	
<i>P. aeruginosa</i>	Clinical Strain F	> 1 × 10 ¹²	0.5	6.25 × 10 ⁸	0.008	63	
<i>P. aeruginosa</i>	Clinical Strain G	> 1 × 10 ¹²	0.5	6.25 × 10 ⁸	0.008	63	
<i>P. aeruginosa</i>	Clinical Strain L	> 1 × 10 ¹²	0.5	1.25 × 10 ⁹	0.016	31	
<i>P. aeruginosa</i>	Clinical Strain 320	> 1 × 10 ¹²	0.5	6.25 × 10 ⁸	0.008	125	

^a calculated using the measured conversion factor of 1.3 × 10⁻¹¹ μg of colistin per virion. MICs were determined in triplicate samples as the concentration where no cell growth was observed in any of the three samples. Plate reader raw data is provided in [Supplementary Table S5](#).

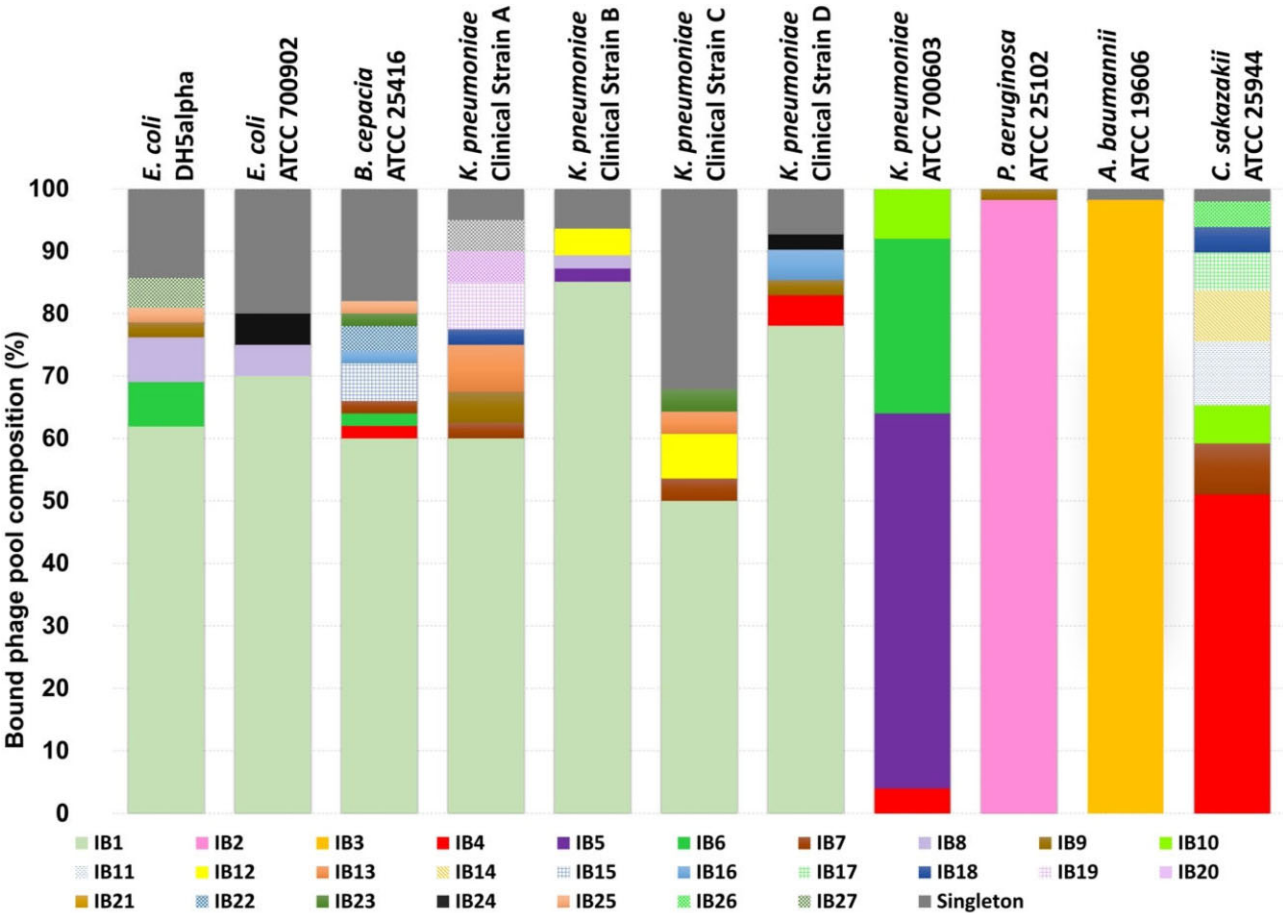


Figure 5. RBDs selected from the Ino-200RBD library. Pool composition of the Ino-200RBD library after two rounds of *in vitro* selection on each of 11 pathogens listed at top, measured by Sanger sequencing of 60 randomly selected clones per selection. Each colored/patterned block represents a unique phage RBD (labeled IB1 to IB27). Singletons (i.e. sequences found only once among all sets) are shown collectively in gray at the top of each bar. RBD sequences found at >50% in a selection (e.g. IB2) or in multiple pathogen selections (e.g. IB1) are shown as solid colors. Other RBD sequences are shown as patterns.

Discussion

Given the challenge of antibiotic resistance, phage-based therapeutics have been increasingly explored. However, the overly high specificity of most WT phages necessitates a lengthy screening process to identify lytic phages against a particular clinical isolate. In addition to identification, this process requires genome sequencing and analysis to ensure that the WT phages do not contain antibiotic resistance or toxin-encoding genes, as well as growth of the phage on a nonmodel organism and purification from bacterial cell lysate. In comparison, the M13 genome has been well-characterized and lacks antibiotic resistance or toxin genes, and M13 grows on standard laboratory strains of *E. coli* and is extruded through the cell wall into the media rather than released through lysis. These significant advantages suggest a promising platform for therapy using genetically and/or chemically modified M13.

In this work, we propose an approach to mine the RBDs of phages based on metagenomic data. A library of potential RBDs was created by pooled assembly PCR of oligonucleotide sequences from inovirus RBDs, generating a highly diverse set of recombinant sequences displayed on an M13 platform. As previously observed [72], gene synthesis efficiency decreased as the number of fragments to be assembled increased. In our case, Pa4 and Ino200-RBD assembly were relatively efficient compared to Pa5 assembly. With four-fragment designs, simultaneous pooled assembly of 200 designs yielded a library from which hits could be selected against each strain tested. The critical limitation on library size is likely to be synthetic (i.e. cost of oligonucleotides), since multiple pooled assembly reactions could be performed in parallel to achieve greater library sizes, and the selection procedure itself would not be greatly affected by a modest increase (1–2 orders of magnitude) in library size. Indeed, the Ino-200RBD library design was down-selected to 200 designs from a more comprehensive metagenomic library design (that might have included all predicted inovirus RBDs) due to the cost of oligonucleotide synthesis. In principle, a library of phages engineered to express RBDs from the inovirus metagenome should have a high probability of yielding members that bind to the wide variety of Gram-negative organisms, which serve as natural hosts of inoviruses.

Using this method, we prepared and characterized two phage libraries, a Pa-RBD library that was designed to target *P. aeruginosa* strains and an Ino-200RBD library to target a broader range of Gram-negative bacteria. In theory, a library encompassing all inovirus RBDs from the metagenomic database would have been an ideal starting point. However, the cost of oligonucleotide synthesis on such a scale was prohibitive for this study, leading to down-selection of the library design to a maximum of 200 designed RBDs. These RBD designs were chosen by sequence similarity to known inovirus RBDs to maximize the frequency of functional RBDs. However, this approach necessarily results in loss of sequence diversity. Inoviruses differ in receptors, even for the same bacterial host species. For example, the primary receptor for M13 is the F pilus, while the primary receptor of another filamentous coliphage, If1, is the I pilus, and the molecular mechanism of infection differs between M13 and If1 despite sharing the same secondary receptor (TolA) [30]. Similarly, Pf1 and Pf3, which both infect *P. aeruginosa*, use different primary receptors, namely the type IV PAK pilus or the conjugative RP4 pilus (not the type IV PAO pilus), respectively [73].

Based on the diversity of host receptors, high diversity in the metagenome-based library would be theoretically desirable. Interestingly, one-pot Gibson assembly created many recombinants (Figs 1 and 2B), expanding on the original library design. Using two rounds of selection for binding, one member (M13^{PAB}) of the Pa-RBD library was isolated, showing binding activity toward multiple strains of *P. aeruginosa*. Similarly, several members from the Ino-200RBD library were isolated individually for binding to *E. coli*, *P. aeruginosa*, *A. baumannii*, *B. cepacia*, *C. sakazakii*, and *K. pneumoniae*, including clinical isolates. Results from the parallel Ino-200RBD library selections indicated the library contains members having a variety of broadness in host binding range, illustrating the potential for selecting phages with high or low specificity depending on the application.

Compared to lytic phage banks used for phage therapy, the recombinant phages developed in this work do not have intrinsic bactericidal activity. While M13 is naturally nonlytic, applying an analogous approach to lytic phages is also likely to yield nonlytic phages due to a lack of compatibility of infection and replication mechanisms across strains and species. Therefore, in general for this approach, antimicrobial activity should be conferred through an independent mechanism, such as delivery of antibiotic molecules (in this case, colistin). Interestingly, the rate of bacterial cell-killing by delivered antibiotics could be significantly faster than that of a lytic phage [74]. Colistin is a membrane-active, cationic lipopeptide that is currently used as a last-line therapy to treat systemic infections against multidrug-resistant Gram-negative pathogens such as *P. aeruginosa* [75–77]. The positively charged peptide ring of colistin interacts with the negatively charged bacterial cell surface, and insertion of a hydrophobic chain into the outer membrane causes membrane disruption and cell death [78]. Although effective, clinical use of colistin is limited by its high rate (36.2%) of associated nephrotoxicity [79], which is caused by low selectivity for bacterial versus mammalian cell membranes. Administration of the prodrug, colistin methanesulfonate, reduces toxicity but also reduces antimicrobial activity [80]. In this work, the low-specificity interaction between colistin and bacterial cells is supplanted by a higher-specificity interaction between the recombinant phage RBD and cell surface antigens. Colistin was nonspecifically cross-linked to M13^{PAB} phage by carbodiimide chemistry. Although nonspecific conjugation raises a general concern of inhibiting binding activity, prior literature indicates that nonspecific conjugation of cargo does not necessarily hinder binding of recombinant M13 phages to their targets [23, 31, 81]. Indeed, colistin-M13^{PAB} was found to lower the MIC by ~30–125-fold across multiple *P. aeruginosa* strains tested, including reducing the MIC for a colistin-resistant strain to below the breakpoint value. In addition, colistin-M13^{PAB} was nonhemolytic and nontoxic to mammalian cells *in vitro*, suggesting that the phage conjugate could widen the therapeutic window of colistin.

While the library-based phage therapy approach should substantially shorten the time between isolation of the causative organism and purified phage product, several considerations remain, such as immunogenicity resulting in neutralization of phages. Interestingly, drug-conjugated M13 elicits ~4-fold lower anti-phage antibodies compared to unconjugated M13 phages *in vivo* [82]. Use of a phage platform like M13 rather than WT phages also raises the possibility that the phage could be engineered to reduce immunogenicity by

genetic or chemical modification of the major coat protein. Additional issues include biodistribution and pharmacokinetics of the engineered phage, as well as the frequency of bacterial resistance, and the presence of microbiome community members that may interfere with binding specificity for specific applications. Nevertheless, engineering a scaffold phage like M13 allows these parameters to be optimized in advance compared to using a WT phage.

Acknowledgements

The authors acknowledge use of the Electron Imaging Center for Nanomachines (EICN) at the UCLA California NanoSystems Institute (CNSI). The graphical abstract was created in BioRender. Chen, I. (2025) <https://BioRender.com/tek6wd1>.

Author contributions: Yanxi Yang (Conceptualization [equal], Data curation [lead], Formal analysis [lead], Investigation [lead], Methodology [lead], Software [lead], Visualization [lead], Writing—original draft [lead], Writing—review & editing [equal]), Dayeon Kang (Investigation [equal], Methodology [equal], Writing—review & editing [equal]), Beatrice Mihalache (Formal analysis [equal], Methodology [equal], Software [equal], Visualization [equal], Writing—review & editing [supporting]), Shelby Vexler (Investigation [equal], Methodology [equal], Writing—review & editing [equal]), Saumya Jain (Investigation [equal], Methodology [equal], Writing—review & editing [equal]), Huan Peng (Methodology [equal], Supervision [supporting], Writing—review & editing [supporting]), Nasim Annabi (Project administration [supporting], Supervision [equal], Writing—review & editing [supporting]), Shangxin Yang (Methodology [supporting], Resources [equal], Writing—review & editing [supporting]), and Irene A. Chen (Conceptualization [equal], Funding acquisition [lead], Project administration [lead], Supervision [lead], Writing—review & editing [equal]).

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

U.S. provisional patent application 63/597 619; PCT/US24/55193. Y.Y. and I.A.C. are co-founders of Paralos Bioscience.

Funding

This work was supported by the NIH National Institute of General Medical Sciences (Award Number R35GM148249); National Science Foundation (Award Number DMR-1933487); NIH National Institute of General Medical Sciences (Award Number T32GM145388); and NIH/National Center for Advancing Translational Science (NCATS) UCLA CTSI Core Voucher Award (Award Number UL1TR001881). H.P. would like to thank the National Natural Science Foundation of China (32201100) for financial support. Bridge funding to support costs, including the Open Access publication charges for this article, during a federal funding freeze was provided by UCLA internal funds.

Data availability

Supplementary Data Files are included for PIP database and library designs; multiple sequence alignment from the Ino-200RBD selections; Jalview consensus; BLASTP results of the consensus fragment; PacBio sequencing files and process analysis files on valid reads for the Pa4, Pa5, and Ino-200RBD phagemid libraries; and amino acid composition analysis. All data files are available on DRYAD (<https://doi.org/10.5061/dryad.6m905qg7g>) (Link for peer reviewers: <https://datadryad.org/share/maBZDApILuETFcB4pEFUuauW6iffeX--HsYS7C8jaU8>). Custom code is given in the Supplementary Information or Dryad, as well as through Github (<https://github.com/StevenYang1997/Receptor-binding-protein-library-design-from-metagenomic-data>; <https://github.com/ichen-lab-ucsb/Pacbio-sequencing-library-analysis>).

References

- Murray CJ, Ikuta KS, Sharara F *et al.* Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* 2022;399:629–55. [https://doi.org/10.1016/S0140-6736\(21\)02724-0](https://doi.org/10.1016/S0140-6736(21)02724-0)
- Kadri SS. Key takeaways from the U.S. CDC's 2019 antibiotic resistance threats report for frontline providers. *Crit Care Med* 2020;48:939–45. <https://doi.org/10.1097/CCM.00000000000004371>
- Walsh TR, Gales AC, Laxminarayan R *et al.* Antimicrobial resistance: addressing a global threat to humanity. *PLoS Med* 2023;20:e1004264. <https://doi.org/10.1371/journal.pmed.1004264>
- Gordillo Altamirano FL, Barr JJ. Phage therapy in the postantibiotic era. *Clin Microbiol Rev* 2019;32:e00066–18. <https://doi.org/10.1128/CMR.00066-18>
- Lessa FC, Sievert DM. Antibiotic resistance: a global problem and the need to do more. *Clin Infect Dis* 2023;77:S1–S3. <https://doi.org/10.1093/cid/ciad226>
- Breijyeh Z, Jubeh B, Karaman R. Resistance of Gram-negative bacteria to current antibacterial agents and approaches to resolve it. *Molecules* 2020;25:1340. <https://doi.org/10.3390/molecules25061340>
- Miller SI. Antibiotic resistance and regulation of the Gram-negative bacterial outer membrane barrier by host innate immune molecules. *mBio* 2016;7:1–3. <https://doi.org/10.1128/mBio.01541-16>
- Haque M, Sartelli M, McKimm J *et al.* Health care-associated infections—an overview. *IDR* 2018;11:2321–33. <https://doi.org/10.2147/IDR.S177247>
- Dutescu IA, Hillier SA. Encouraging the development of new antibiotics: are financial incentives the right way forward? a systematic review and case study. *IDR* 2021;14:415–34. <https://doi.org/10.2147/IDR.S287792>
- Gupta R, Sharma S. Role of alternatives to antibiotics in mitigating the antimicrobial resistance crisis. *Indian J Med Res* 2022;156:464–77. https://doi.org/10.4103/ijmr.IJMR_3514_20
- El-Atrees DM, El-Kased RF, Abbas AM *et al.* Characterization and anti-biofilm activity of bacteriophages against urinary tract *Enterococcus faecalis* isolates. *Sci Rep* 2022;12:1–11. <https://doi.org/10.1038/s41598-022-17275-z>
- Dion MB, Oechslin F, Moineau S. Phage diversity, genomics and phylogeny. *Nat Rev Microbiol* 2020 18:125–38. <https://doi.org/10.1038/s41579-019-0311-5>
- Stacey HJ, De Soir S, Jones JD. The safety and efficacy of phage therapy: a systematic review of clinical and safety trials. *Antibiotics* 2022;11:1340. <https://doi.org/10.3390/antibiotics11101340>

14. Hitchcock NM, Nunes DDG, Shiach J *et al.* Current clinical landscape and global potential of bacteriophage therapy. *Viruses* 2023;15:1020. <https://doi.org/10.3390/v15041020>
15. Liu D, Van Belleghem JD, de Vries CR *et al.* The safety and toxicity of phage therapy: a review of animal and clinical studies. *Viruses* 2021;13:1268. <https://doi.org/10.3390/v13071268>
16. Oechslin F. Resistance development to bacteriophages occurring during bacteriophage therapy. *Viruses* 2018;10:351. <https://doi.org/10.3390/v10070351>
17. Champagne-Jorgensen K, Luong T, Darby T *et al.* Immunogenicity of bacteriophages. *Trends Microbiol* 2023;31:1058–71. <https://doi.org/10.1016/j.tim.2023.04.008>
18. Loc-Carrillo C, Abedon ST. Pros and cons of phage therapy. *Bacteriophage* 2011;1:111–4. <https://doi.org/10.4161/bact.1.2.14590>
19. Green SI, Clark JR, Santos HH *et al.* A retrospective, observational study of 12 cases of expanded-access customized phage therapy: production, characteristics, and clinical outcomes. *Clin Infect Dis* 2023;77:1079–91. <https://doi.org/10.1093/cid/ciad335>
20. Fletcher J, Manley R, Fitch C *et al.* The citizen phage library: rapid isolation of phages for the treatment of antibiotic resistant infections in the UK. *Microorganisms* 2024;12:253. <https://doi.org/10.3390/microorganisms12020253>
21. Nagel T, Musila L, Muthoni M *et al.* Phage banks as potential tools to rapidly and cost-effectively manage antimicrobial resistance in the developing world. *Curr Opin Virol* 2022;53:101208. <https://doi.org/10.1016/j.coviro.2022.101208>
22. Peng H, Borg RE, Dow LP *et al.* Controlled phage therapy by photothermal ablation of specific bacterial species using gold nanorods targeted by chimeric phages. *Proc Natl Acad Sci USA* 2020;117:1951–61. <https://doi.org/10.1073/pnas.1913234117>
23. Peng H, Chen IA. Rapid colorimetric detection of bacterial species through the capture of gold nanoparticles by chimeric phages. *ACS Nano* 2019;13:1244–52.
24. Guo D, Chen J, Zhao X *et al.* Genetic and chemical engineering of phages for controlling multidrug-resistant bacteria. *Antibiotics* 2021;10:202. <https://doi.org/10.3390/antibiotics10020202>
25. Ledsgaard L, Kilstrup M, Karatt-Vellatt A *et al.* Basics of antibody phage display technology. *Toxins* 2018;10:236. <https://doi.org/10.3390/toxins10060236>
26. Knezevic P, Adriaenssens EM, Consortium IR. ICTV Virus Taxonomy Profile: inoviridae. *J Gen Virol* 2021;102:001614.
27. Jia Q, Xiang Y. Cryo-EM structure of a bacteriophage M13 mini variant. *Nat Commun* 2023;14:1–14.
28. Hofschneider PH. Untersuchungen über „kleine“ *E. coli* K 12 Bakteriophagen: 1. und 2. Mitteilung. *Zeitschrift Für Naturforschung B* 1963;18:203–10. <https://doi.org/10.1515/znb-1963-0306>
29. Kleinbeck F, Kuhn A. Membrane insertion of the M13 minor coat protein G3p is dependent on YidC and the SecAYEG translocase. *Viruses* 2021;13:1414. <https://doi.org/10.3390/v13071414>
30. Lorenz SH, Jakob RP, Weininger U *et al.* The filamentous phages fd and IF1 use different mechanisms to infect *Escherichia coli*. *J Mol Biol* 2011;405:989–1003. <https://doi.org/10.1016/j.jmb.2010.11.030>
31. Wang R, Li HD, Cao Y *et al.* M13 phage: a versatile building block for a highly specific analysis platform. *Anal Bioanal Chem* 2023;415:3927–44. <https://doi.org/10.1007/s00216-023-04606-w>
32. Chang C, Guo W, Yu X *et al.* Engineered M13 phage as a novel therapeutic bionanomaterial for clinical applications: from tissue regeneration to cancer therapy. *Materials Today Bio* 2023;20:100612. <https://doi.org/10.1016/j.mtbio.2023.100612>
33. Roux S, Krupovic M, Daly RA *et al.* Cryptic inoviruses revealed as pervasive in bacteria and archaea across Earth's biomes. *Nat Microbiol* 2019;4:1895–906. <https://doi.org/10.1038/s41564-019-0510-x>
34. Chung W-J, Lee D-Y, Yoo SY. Chemical modulation of M13 bacteriophage and its functional opportunities for nanomedicine. *Int J Nanomedicine* 2014;9:5825–36.
35. Gutu AD, Sgambati N, Strasbourger P *et al.* Polymyxin resistance of *Pseudomonas aeruginosa* mutants is dependent on additional two-component regulatory systems. *Antimicrob Agents Chemother* 2013;57:2204–15. <https://doi.org/10.1128/AAC.02353-12>
36. Miller AK, Brannon MK, Stevens L *et al.* PhoQ mutations promote lipid A modification and polymyxin resistance of *Pseudomonas aeruginosa* found in colistin-treated cystic fibrosis patients. *Antimicrob Agents Chemother* 2011;55:5761–9. <https://doi.org/10.1128/AAC.05391-11>
37. Fiedoruk K, Zakrzewska M, Daniluk T *et al.* Two lineages of *Pseudomonas aeruginosa* filamentous phages: structural uniformity over integration preferences. *Genome Biol Evol* 2020;12:1765–81. <https://doi.org/10.1093/gbe/evaa146>
38. Secor PR, Elizabeth BB, Kinnersley M *et al.* Pf bacteriophage and their impact on pseudomonas virulence, mammalian immunity, and chronic infections. *Front Immunol* 2020;11:244. <https://doi.org/10.3389/fimmu.2020.00244>
39. Kim SH, Lee KB, Lee JS *et al.* Genome diversification by phage-derived genomic islands in *Pseudomonas aeruginosa*. *J Microbiol Biotechnol* 2003;13:783–8.
40. Hill DF, Short NJ, Perham RN *et al.* DNA sequence of the filamentous bacteriophage Pf1. *J Mol Biol* 1991;218:349–64. [https://doi.org/10.1016/0022-2836\(91\)90717-K](https://doi.org/10.1016/0022-2836(91)90717-K)
41. Altschul SF, Gish W, Miller W *et al.* Basic local alignment search tool. *J Mol Biol* 1990;215:403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
42. Camacho C, Coulouris G, Avagyan V *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* 2009;10:421. <https://doi.org/10.1186/1471-2105-10-421>
43. Cock PJA, Antao T, Chang JT *et al.* Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 2009;25:1422–3. <https://doi.org/10.1093/bioinformatics/btp163>
44. Ahmadi Z, Farajnia S, Farajzadeh D *et al.* Optimized signal peptide for secretory expression of human recombinant somatropin in *E. coli*. *Adv Pharm Bull* 2023;13:339–49. <https://doi.org/10.34172/apb.2023.037>
45. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK *et al.* SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* 2019;37:420–3. <https://doi.org/10.1038/s41587-019-0036-z>
46. Reddy Chichili VP, Kumar V, Sivaraman J. Linkers in the structural biology of protein–protein interactions. *Protein Sci* 2013;22:153–67. <https://doi.org/10.1002/pro.2206>
47. Zulkower V, Rosser S. DNA Chisel, a versatile sequence optimizer. *Bioinformatics* 2020;36:4508–9. <https://doi.org/10.1093/bioinformatics/btaa558>
48. Wenger AM, Peluso P, Rowell WJ *et al.* Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nat Biotechnol* 2019;37:1155–62. <https://doi.org/10.1038/s41587-019-0217-9>
49. Bostock M, Ogievetsky V, Heer J. D3: data-driven documents. *IEEE Trans Vis Comput Graph* 2011;17:2301–9. <https://doi.org/10.1109/TVCG.2011.185>
50. Passaretti P, Khan I, Dafforn TR *et al.* Improvements in the production of purified M13 bacteriophage bio-nanoparticle. *Sci Rep* 2020;10:1–9. <https://doi.org/10.1038/s41598-020-75205-3>
51. Lee B-YY, Lee J, Ahn DJ *et al.* Optimizing protein V untranslated region sequence in M13 phage for increased production of single-stranded DNA for origami. *Nucleic Acids Res* 2021;49:6596–603. <https://doi.org/10.1093/nar/gkab455>
52. Moskowitz SM, Ernst RK, Miller SI. PmrAB, a two-component regulatory system of *Pseudomonas aeruginosa* that modulates resistance to cationic antimicrobial peptides and addition of aminoarabinose to lipid A. *J Bacteriol* 2004;186:575–9. <https://doi.org/10.1128/JB.186.2.575-579.2004>

53. Yang Y, Chen IA. Visualization of engineered M13 phages bound to bacterial targets by transmission electron microscopy. *Methods Mol Biol* 2024;2793:175–83.
54. Wiegand I, Hilpert K, Hancock REW. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat Protoc* 2008;3:163–75. <https://doi.org/10.1038/nprot.2007.521>
55. Salas-Ambrosio P, Vexler S, Rajalakshmi PS *et al.* Caffeine and cationic copolymers with antimicrobial properties. *ACS Bio Med Chem Au* 2023;3:189–200. <https://doi.org/10.1021/acsbioimedchemau.2c00077>
56. Evans BC, Nelson CE, Yu SS *et al.* *Ex vivo* red blood cell hemolysis assay for the evaluation of pH-responsive endosomolytic agents for cytosolic delivery of biomacromolecular drugs. *J Vis Exp* 2013;73:e50166.
57. Bang S, Jung U-W, Noh I. Synthesis and biocompatibility characterizations of *in situ* chondroitin sulfate-gelatin hydrogel for tissue engineering. *Tissue Eng Regen Med* 2018;15:25–35. <https://doi.org/10.1007/s13770-017-0089-3>
58. Kiliç H, Ceylan Tuncaboylu D, Argun A *et al.* Design of biocompatible multifunctional hydrogels with stearyl methacrylate and vinylpyrrolidone. *ACS Appl Polym Mater* 2022;4:1717–27.
59. Thomas S, Maynard ND, Gill J. DNA library construction using Gibson Assembly®. *Nat Methods* 2015;12:i–ii. <https://doi.org/10.1038/nmeth.f.384>
60. Baek M, Dimaio F, Anishchenko I *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* 2021;373:871–6. <https://doi.org/10.1126/science.abj8754>
61. Jumper J, Evans R, Pritzel A *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;596:583–9. <https://doi.org/10.1038/s41586-021-03819-2>
62. Mirdita M, Schütze K, Mirowski Y *et al.* ColabFold: making protein folding accessible to all. *Nat Methods* 2022;19:679–82. <https://doi.org/10.1038/s41592-022-01488-1>
63. Chew KL, La M-V, Lin RTP *et al.* Colistin and polymyxin B susceptibility testing for carbapenem-resistant and mcr-positive Enterobacteriaceae: comparison of Sensititre, Microscan, Vitek 2, and Etest with broth microdilution. *J Clin Microbiol* 2017;55:2609–16. <https://doi.org/10.1128/JCM.00268-17>
64. Satlin MJ, Lewis JS, Weinstein MP *et al.* Clinical and Laboratory Standards Institute and European Committee on Antimicrobial Susceptibility Testing position statements on polymyxin B and colistin clinical breakpoints. *Clin Infect Dis* 2020;71:e523–9. <https://doi.org/10.1093/cid/ciaa121>
65. CLSI. *CLSI M100-ED29: 2021 Performance Standards for Antimicrobial Susceptibility Testing*. 30th Edition. CLSI, Malvern, PA, 2020.
66. Lowe G *et al.* Nursing blood specimen collection techniques and hemolysis rates in an emergency department: analysis of venipuncture versus intravenous catheter collection techniques. *J Emerg Nurs* 2008;34:26–32. <https://doi.org/10.1016/j.jen.2007.02.006>
67. Witherell KS, Price J, Bandaranayake AD *et al.* *In vitro* activity of antimicrobial peptide CDP-B11 alone and in combination with colistin against colistin-resistant and multidrug-resistant *Escherichia coli*. *Sci Rep* 2021;11:1–10. <https://doi.org/10.1038/s41598-021-81140-8>
68. Naghmouchi K *et al.* Synergistic effect between colistin and bacteriocins in controlling Gram-negative pathogens and their potential to reduce antibiotic toxicity in mammalian epithelial cells. *Antimicrob Agents Chemother* 2013;57:2719–25. <https://doi.org/10.1128/AAC.02328-12>
69. Iwasawa A, Ayaki M, Niwano Y. Cell viability score (CVS) as a good indicator of critical concentration of benzalkonium chloride for toxicity in cultured ocular surface cell lines. *Regul Toxicol Pharm* 2013;66:177–83. <https://doi.org/10.1016/j.yrtph.2013.03.014>
70. Papadopoulos JS, Agarwala R. COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics* 2007;23:1073–9. <https://doi.org/10.1093/bioinformatics/btm076>
71. Waterhouse AM, Procter JB, Martin DM *et al.* Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 2009;25:1189–91. <https://doi.org/10.1093/bioinformatics/btp033>
72. Gibson DG, Young L, Chuang RY *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 2009;6:343–5. <https://doi.org/10.1038/nmeth.1318>
73. Holland SJ, Sanz C, Perham RN. Identification and specificity of pilus adsorption proteins of filamentous bacteriophages infecting *Pseudomonas aeruginosa*. *Virology* 2006;345:540–8. <https://doi.org/10.1016/j.virol.2005.10.020>
74. Nilsson AS. Pharmacological limitations of phage therapy. *Ups J Med Sci* 2019;124:218–27. <https://doi.org/10.1080/03009734.2019.1688433>
75. Martis N, Leroy S, Blanc V. Colistin in multi-drug resistant *Pseudomonas aeruginosa* blood-stream infections: a narrative review for the clinician. *J Infect* 2014;69:1–12. <https://doi.org/10.1016/j.jinf.2014.03.001>
76. Ledger EVK, Sabnis A, Edwards AM. Polymyxin and lipopeptide antibiotics: membrane-targeting drugs of last resort. *Microbiology* 2022;168:1–20.
77. Maranhão RC, Vital CG, Tavoni TM *et al.* Clinical experience with drug delivery systems as tools to decrease the toxicity of anticancer chemotherapeutic agents. *Expert Opin Drug Deliv* 2017;14:1217–26. <https://doi.org/10.1080/17425247.2017.1276560>
78. Andrade FF, Silva D, Rodrigues A *et al.* Colistin update on its mechanism of action and resistance, present and future challenges. *Microorganisms* 2020;8:1716. <https://doi.org/10.3390/microorganisms8111716>
79. Eljaaly K *et al.* Colistin nephrotoxicity: meta-analysis of randomized controlled trials. *Open Forum Infect Dis* 2021;8:ofab026. <https://doi.org/10.1093/ofid/ofab026>
80. Bergen PJ, Li J, Rayner CR *et al.* Colistin methanesulfonate is an inactive prodrug of colistin against *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother* 2006;50:1953–8. <https://doi.org/10.1128/AAC.00035-06>
81. Chung W-J, Lee D-Y, Yoo SY. Chemical modulation of M13 bacteriophage and its functional opportunities for nanomedicine. *Int J Nanomedicine* 2014;9:5825–36.
82. Vaks L, Benhar I. *In vivo* characteristics of targeted drug-carrying filamentous bacteriophage nanomedicines. *J Nanobiotechnology* 2011;9:1–10. <https://doi.org/10.1186/1477-3155-9-58>