



OPEN ACCESS

EDITED BY

Alicja Wegrzyn,
University of Gdansk, Poland

REVIEWED BY

Gunaraj Dhungana,
Meharry Medical College, United States
Hasan Arsin,
University of Bergen, Norway
Becky Hess,
Pacific Northwest National Laboratory (DOE),
United States

*CORRESPONDENCE

Aditi Singh

✉ asingh3@lko.amity.edu;
✉ singh.aditi00@gmail.com

RECEIVED 23 December 2024

ACCEPTED 04 June 2025

PUBLISHED 09 July 2025

CITATION

Saeed H, Padmesh S, Singh A, Singh SP,
Siddiqui MH, Sen M, Hussain I and
Beg MMA (2025) Characterization of holins,
the membrane proteins of coliphage
ASEC2201: a genomewide *in silico* approach.
Front. Microbiol. 16:1550594.
doi: 10.3389/fmicb.2025.1550594

COPYRIGHT

© 2025 Saeed, Padmesh, Singh, Singh,
Siddiqui, Sen, Hussain and Beg. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Characterization of holins, the membrane proteins of coliphage ASEC2201: a genomewide *in silico* approach

Humaira Saeed¹, Sudhaker Padmesh¹, Aditi Singh^{1*},
Sujeet Pratap Singh¹, Mohammed Haris Siddiqui²,
Manodeep Sen³, Imran Hussain⁴ and Mirza Masroor Ali Beg⁵

¹Amity Institute of Biotechnology, Amity University Uttar Pradesh, Lucknow Campus, Lucknow, India,

²Dr. Ram Manohar Lohia Institute of Medical Sciences, Lucknow, India, ³Integral Institute of
Agricultural Science and Technology, Integral University, Lucknow, India, ⁴Department of
Bioengineering, Faculty of Engineering and Information Technology, Integral University, Lucknow,
India, ⁵Faculty of Medicine, Alatoo International University, Bishkek, Kyrgyzstan

Drug-resistant *Escherichia coli* poses a significant healthcare burden, driving the search for novel antimicrobials. We have previously done the isolation and whole-genome sequencing of ASEC2201, a novel coliphage derived from multidrug-resistant clinical *E. coli* strains. Here, we report the identification and characterization of phage enzyme, holin by *in silico* approaches. Genome annotation using Prokka identified three putative holin genes (PROKKA_03659, PROKKA_04292, and PROKKA_04422) belonging to the Phage_holin_2_1 superfamily. Upstream promoter prediction revealed active regulatory elements at positions 112, 177, and 186 for these genes, indicating robust transcriptional activity. Transmembrane topology analysis using DeepTMHMM confirmed the presence of two to three α -helical membrane-spanning domains in each holin, essential for pore formation. Homology modeling with SWISS-MODEL yielded high-confidence three-dimensional structures characterized by conserved membrane-anchoring motifs, as supported by QMEAN and GMQE quality scores. *In silico* identification of cell-penetrating peptide motifs within the holin sequences suggests potential for enhanced intracellular delivery in CPP-fusion therapeutic constructs. Overall, our *in-depth* analysis elucidates the structural and functional properties of ASEC2201 holins, underscoring their biotechnological significance as scaffolds for developing novel antimicrobial strategies against MDR *E. coli*. It gives us an understanding on how the holins, with their inherent membrane-disrupting functions, can be explored in detail for future use as lysis modules in programmable bacterial systems, while their identified CPP motifs offer additional potential for engineering targeted therapeutic delivery vehicles. This study also demonstrates the potential of integrative *in silico* approaches in developing a comprehensive foundation for future experimental validation for proteins with no prior functional annotation.

KEYWORDS

transmembrane domain, phage-mediated lysis, ASEC2201, protein 3D structure, functional annotation, holins, bacteriophage, bacteriophage-encoded enzymes

Introduction

Holins and endolysins are crucial for host lysis in bacteriophage infections (Wang et al., 2000). Endolysins are muralytic enzymes that degrade the bacterial cell wall and accumulate in the cytosol, fully folded, during the vegetative cycle (Liu et al., 2023). Holins, small membrane proteins, accumulate in the bacterial membrane and, at a programmed time, create pores, allowing endolysins to access the cell wall (Abeysekera et al., 2022). This leads to the destruction of the bacterial murein and cell bursting. Holins regulate the timing of phage infection, which is under intense evolutionary pressure for optimal lysis timing. Their activity is controlled by protein inhibitors. Holins represent a highly diverse protein group, with over 100 known sequences forming more than 30 ortholog groups. Holins are classified into three classes based on their membrane topology. Class I holins, the largest, contain over 95 amino acids and form three transmembrane domains. Class II holins are smaller, with 65–95 amino acids, and form only two transmembrane domains. Class III holins are the smallest, with just one transmembrane domain. Class I and II holins are more commonly found than class III. These structural differences contribute to the varying mechanisms of pore formation and function within bacterial membranes, influencing the timing and efficiency of bacterial cell lysis during bacteriophage infection (Grabowski et al., 2021). However, on the basis of functionality, there are majorly three types of holins found, i.e.: canonical holins, pinholins and antiholins (Cahill and Young, 2019). In the context of bacteriophage infection, “canonical holins,” “pinholins,” and “antiholins” are all proteins involved in the cell lysis process, with the key difference being the size of the holes they create in the bacterial membrane. Canonical holins form large membrane pores, enabling endolysin release to degrade the cell wall, leading to cell lysis. Pinholins, in contrast, create smaller pinholes to depolarize the membrane, often acting before canonical holins initiate full-scale lysis.

Holins function as transmembrane transporters for endolysins, since endolysins cannot cross the bacterial membrane alone. After phage infection, precise timing and host lysis disrupt the cell envelope—inner membrane, peptidoglycan, and outer membrane. In the phage lytic system, holins form pores in the inner membrane to release endolysins and set the infection cycle's endpoint. A threshold concentration of holins triggers pore formation, ensuring accurate lysis timing. By facilitating endolysin escape, holins accelerate and enhance host cell breakdown. Although phage-derived enzymes rarely face resistance due to extensive horizontal gene transfer, bacteria can resist exogenous endolysins via capsules, biofilms, and other outer defenses (Murray et al., 2021).

Antiholins regulate holin-mediated lysis by inhibiting pore formation until optimal timing is reached. In phage λ , dual translation of gene S yields holin S105 and antiholin S107; an extra N-terminal Lys residue hinders TMD1 membrane insertion, preventing early lysis, but membrane depolarization at trigger removes this barrier to activate lysis (White et al., 2010). In phage T4, cytoplasmic RIII binds the T holin's N-terminal domain to delay pore formation, and upon superinfection, RIII activation establishes lysis inhibition, extending the latent period and boosting progeny maturation (Chen and Young, 2016; Krieger et al., 2020). Antiholins thus finely tune lysis timing, ensuring synchronized bursting with phage maturation.

Thus, such enzymes are necessary components of the lytic phage life cycle and provide a possible alternative to antibiotics (Singh et al.,

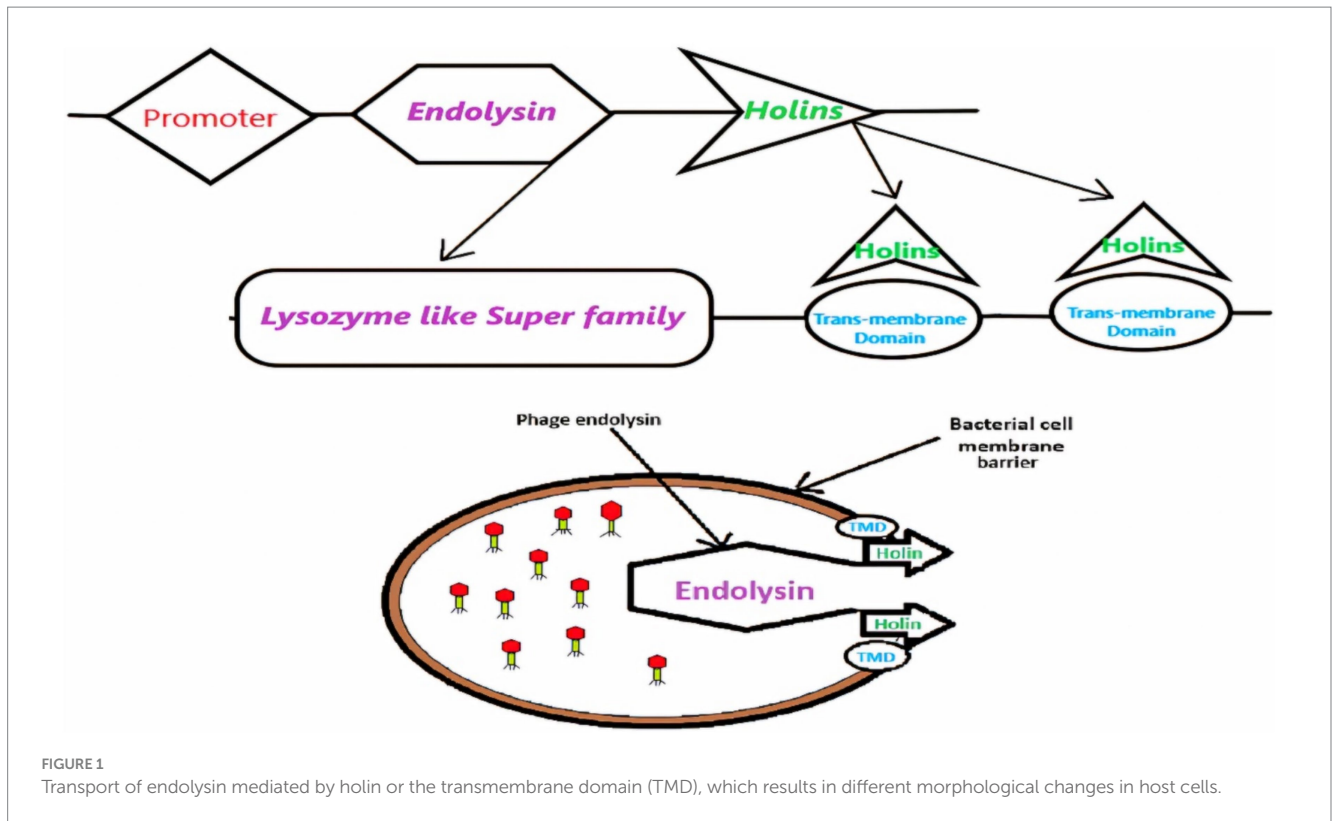
2022). Endolysins have gained popularity in recent years because of their wide lytic activity against both Gram-positive and Gram-negative bacterial cells (Masih and Singh, 2023; Nandy et al., 2022). Using *in silico* technologies, researchers may examine the three-dimensional structural conformation of holins (Verma et al., 2024), categorize new domains (Yuan et al., 2023), investigate specific pathways to acquire a better knowledge of our evolutionary tree (Turner et al., 2021), identify more clusters (Zhang et al., 2023), and assign functions to the proteins (Zhang et al., 2025). This information may also be utilized to design effective pharmacological approaches and help in the creation of novel medications to treat a wide range of ailments (Setten et al., 2019; Ali et al., 2020; Kumar et al., 2021; Shaath et al., 2024). Lauer et al. (2010) developed genetically engineered *Listeria monocytogenes* strains expressing holin proteins to enhance the delivery of heterologous antigens into host cell cytosol. This strategy facilitates robust immune responses, making these strains promising vectors for vaccines and immunotherapies targeting cancer and infectious diseases. The approach is detailed in US Patent US20100129406A1, which describes pharmaceutical compositions and methods utilizing these modified *Listeria* strains (Lauer et al., 2010). Cell-penetrating peptides (CPPs) are promising tools for delivering therapeutic molecules, such as nucleic acids, drugs, and imaging agents, into cells and tissues in a nontoxic manner (Du et al., 2025; Pancholi et al., 2025). They are simple to synthesize, functionalize, and characterize (Copolovici et al., 2014), enabling delivery of bioactive cargos, including small drugs and large plasmid DNA, primarily via endocytosis (Moreno-Vargas and Prada-Gracia, 2024). CPPs facilitate processes like gene expression, silencing, and tumor targeting, making them valuable for future drug and diagnostic development (Borrelli et al., 2018; Zorko et al., 2022). However, their passive and nonselective nature often necessitates functionalization or chemical modification to achieve specific targeting (Nam et al., 2023). Designing effective systemic delivery systems requires careful consideration of both CPPs and their delivered cargos (Gautam et al., 2013; Ndeboko et al., 2020).

In prior research, our laboratory identified two novel lytic bacteriophages, ASEC2201 and ASEC2202, from urban sewage, which effectively lysed 40–44% of *E. coli* clinical isolates and demonstrated resilience across diverse physicochemical conditions. Transmission electron microscopy classified them as tailed phages, and genomic analysis placed them within the Drexelviriidae family. In the current study, we conducted an *in silico* characterization of ASEC2201's holins, focusing on their structural and functional attributes. This analysis provided insights into their transmembrane domains and supported their role in facilitating endolysin-mediated host cell lysis (Figure 1).

Methodology

Phage isolation, characterization, and genomic analysis

In our previous study, we have isolated and characterized a novel coliphage designated as ASEC2201. It has shown promising results for extreme environmental resilience and biocontrol efficacy (Padmesh et al., 2024). The phage retained infectivity after exposure to thermal cycles between -20°C and 62°C and across a broad proton concentration window (pH 4–10). Moreover, ASEC2201 particles exhibited 40–94% viability when challenged with organic solvents and



sustained plaque-forming ability following treatment with anionic and cationic detergents. *In vitro* lytic assays demonstrated potent bactericidal activity against a laboratory reference *E. coli* strain and lysed 40–44% of multidrug-resistant clinical *E. coli* isolates.

After gradient purification, the phage was identified by Transmission electron microscopy, which characterized ASEC2201 as tailed phage. For genomic analysis, DNA from ASEC2201 was extracted following a modified protocol (Jakočiūnė and Moodley, 2018). High-quality DNA was used for Illumina NovaSeq 6,000 sequencing (RedCliff Labs, India). Libraries were prepared using PCR-based dual-indexing. Reads were quality-checked and assembled *de novo*; contigs were validated using BLAST. The whole-genome sequencing and comparative genomics assigned them to the genus of Escherichia phages within the family Drexelviridae (order Caudovirales). Resulting data were deposited into the National Center for Biotechnology Information (NCBI) databases, with assembled genomes submitted to BioProject and raw FASTQ files to the Sequence Read Archive (SRA, 2004).

Protein selection and sequence retrieval

Protein-coding gene annotation is typically a two-step process. Initially, Prodigal is employed to identify open reading frames (ORFs) by locating gene coordinates, but it does not infer gene function. To assign putative functions, Prokka performs hierarchical annotation by comparing candidate genes to curated protein databases. It begins with a user-supplied, high-confidence protein set, using BLAST+ for sequence similarity searches. If no match is found, it progresses to UniProt's verified bacterial proteins, covering ~16,000 sequences, and then optionally to RefSeq proteins specific to the organism's genus—capturing nomenclature consistency. When sequence-based annotation

fails, Prokka applies profile-based searches using HMMER's hmmscan to query against Pfam and TIGRFAMs databases. An e-value threshold of 10^{-6} is consistently applied to ensure significance. If no reliable match is found across all levels, the gene is designated as a “hypothetical protein.” This layered strategy maximizes annotation accuracy and functional insight across diverse bacterial genomes (Seemann, 2014).

The genome of coliphage ASEC2201 has been analyzed and three holin protein coding genes were selected. The sequences of all three holin proteins were retrieved from NCBI using accession no. SRX17770782 in the FASTA format. The sequence similarity search was performed via BLAST against the non-redundant database (Altschul et al., 1990).

Phylogenetic analysis of ASEC2201 holin system

The annotated genome of coliphage ASEC2201 was retrieved from the NCBI database and analyzed for similarity using BLASTn. This allowed the identification of regions with homology and potential functional similarities. A phylogenetic analysis using the neighbor-joining method was conducted, Mega XII was employed to build a phylogenetic tree among three key proteins, providing insights into their evolutionary relationship (Kumar et al., 2018).

Secondary structural properties and assessment

Secondary structure predictions were conducted using the following two step process. The amino acid sequences of the holin proteins were retrieved in FASTA format. SOPMA (Self-Optimized

Prediction Method with Alignment) was used to predict the proportion of α -helices, β -strands, and random coils for each protein. The SOPMA web server was accessed at: https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html and default parameters were used. In the next step, PSIPRED v4.0 (available at: <http://bioinf.cs.ucl.ac.uk/psipred/>), was used for secondary structure validation. Each protein sequence was submitted individually, and predictions were recorded based on the graphical output and raw data. Predicted percentages of helices, strands, and coils were compared across tools for consistency. No manual adjustments were made to the input data or output interpretation.

Three-dimensional structure prediction and validation

Swiss Model and HHpred were utilized to predict the 3D (tertiary) structure of holins from phage ASEC2201 (Zimmermann et al., 2018; Söding et al., 2005). To validate the predicted structures, the Ramachandran plot analysis was conducted using the PROCHECK tool from the SAVES (v.6.0) program.

In silico analysis of the selected proteins

The amino acid sequence composition, instability index, aliphatic index, and GRAVY index of the three holin proteins from phage ASEC2201 were analyzed using the ExPASyProtParam tool (Nedvedova et al., 2023; Wilkins et al., 1999). The theoretical isoelectric point (pI) was determined using SMS Suite (v.2.0). These properties help predict protein stability, hydrophobicity, and overall behavior under physiological conditions, aiding in understanding their structural and functional characteristics (Stothard, 2000).

Functional annotation of the three holins

The conserved domains of the three holins from phage ASEC2201 were predicted using NCBI's CD-search tool. Protein motifs were identified with ExPASy's ScanProsite tool, revealing key functional patterns (de Castro et al., 2006). Additionally, the evolutionary relationships of these holins were analyzed using the SuperFamily program, which provides insights into the evolutionary history and relationships of the proteins by comparing them to known protein superfamilies (Wilson et al., 2009). These tools aid in understanding the structural and evolutionary aspects of these holins.

In silico genome-wide analysis of lysis system

ClustalW2 was used to align the amino acid sequences of phage ASEC2201 with homologous proteins. The 3D structure of ASEC2201 was modeled using Phyre2, while conserved domains were identified via the NCBI conserved domain database (Kelley et al., 2015). Promoter regions were predicted using Bprom, helping to map functional and regulatory regions within the phage genome for further research and application (Solovyev and Salamov, 2011).

Cell penetrating peptide identification

High-throughput techniques have advanced the discovery of cell-penetrating peptides (CPPs). CPP has been used to determine the penetrating peptides in the amino acid stretch of selected holins protein, combining *in silico* screening with experimental validation offers a cost-effective and practical approach, ensuring reliable, reproducible outcomes *in vitro*. CellPPD, a standalone tool for predicting and designing CPPs, has been utilized in this study to evaluate the cell-penetrating capabilities of selected peptides, streamlining the identification process. Such approaches significantly enhance CPP research, enabling the development of efficient peptide-based delivery systems for therapeutic and diagnostic applications (Gautam et al., 2013).

Protein transmembrane domain recognition by deep transmembrane helices hidden Markov models (DeepTMHMM)

DeepTMHMM, a protein language model-based algorithm that can detect and predict the topology of both alpha helical and beta barrels proteins with unprecedented accuracy, has been used in the study. It scales to proteomes and covers all domains of life, which makes it ideal for metagenomic analyses (Hallgren et al., 2022). DeepTMHMM outperforms traditional HMM-based methods by integrating deep neural networks with biological sequence understanding and accurately predicts cytoplasmic vs. extracellular orientation of N- and C-termini. This is crucial for understanding holin activation, interaction with antiholins, and membrane-disruption mechanisms. The number and arrangement of predicted TMDs help classify holins into Class I, II, or III and is particularly effective in genome-wide mining for new holins or membrane proteins with limited sequence homology, using structure-based topology rather than relying solely on sequence alignment.

Results

Genome annotation of phage ASEC2201

The complete nucleic acid sequence of coliphage ASEC2201 was determined through next-generation sequencing (NGS) as described above. Subsequent *de novo* assembly of the raw sequencing data yielded the complete genome sequence. The genome was annotated with Prokka, which identified three holin protein-coding genes: PROKKA_03659, PROKKA_04292, and PROKKA_04422 (Supplementary Table 1).

Phylogenetic analysis of holin proteins in coliphage ASEC2201

The un-annotated holins named PROKKA_03659, PROKKA_04292 and PROKKA_04422 fall into two quite different parts of the Gammaproteobacteria/phage-holin subtree (Figure 2).

PROKKA_03659 is nested deep within a well-supported subclade (bootstrap 1.00) of Escherichia phage holins. Its closest neighbors in

that clade are experimentally characterized phage holins such as, YP_009901614.1 (holin from *E. coli* phage vB EcoS G29-2), UGL62530.1 (putative holin from *E. coli* phage JLBV09), QBZ71339.1 (holin from *E. coli* phage Jahat MG145) and WWD13012.1 (holin from *E. coli* phage WaterSpirit). Clustering supports that PROKKA_03659 is very closely related to known *Escherichia* phage holins.

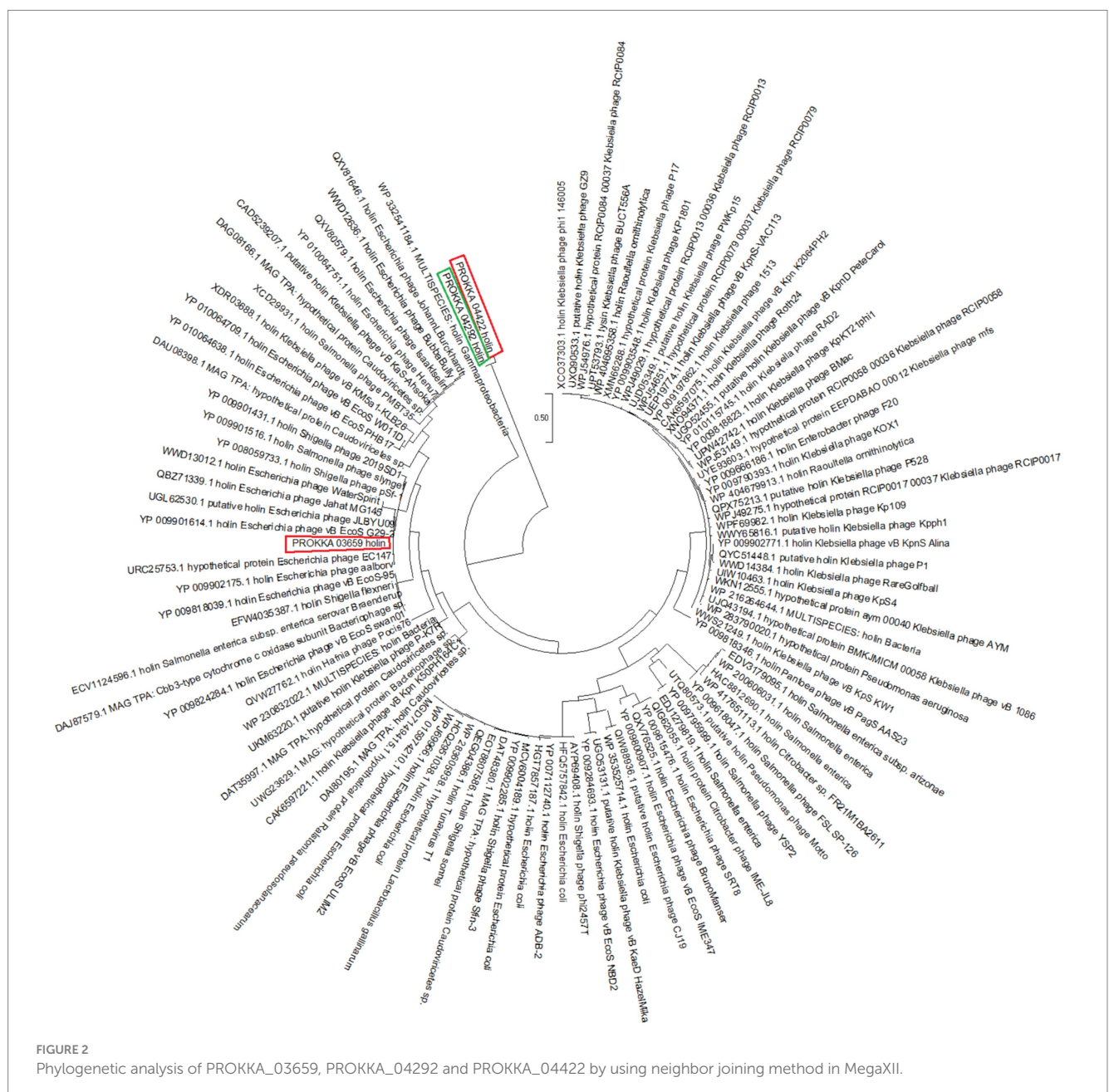
PROKKA_04292 and PROKKA_04422 actually form their own little sister-pair at the base of the Gammaproteobacteria holin radiation. The branch uniting them has only weak bootstrap support (~0.06), suggesting it is uncertain (i.e., it is a tentative placement). They sit immediately downstream of the broad “WP_332541184.1 holin Gammaproteobacteria” root, before the big cluster of Enterobacteria phage holins (including PROKKA_03659’s clade). In other words, PROKKA_04292 and PROKKA_04422 represent a

distinct lineage of gammaproteobacterial holins that is not part of the main *Escherichia* phage grouping (Supplementary Figures 1, 2).

PROKKA_03659 is almost certainly a canonical *Escherichia* phage-type holin, given its tight clustering and high support with those sequences. PROKKA_04292 and PROKKA_04422, by contrast, look like they belong to a more divergent and as yet uncharacterized branch of Gammaproteobacteria-associated holins. Their basal position and low bootstrap suggest they may have quite different sequence/functional features compared to the well-studied enterobacterial phage holins.

Secondary structure prediction by SOPMA

The secondary structure prediction of the holin proteins using SOPMA (Self-Optimized Prediction Method with Alignment)



software provides valuable insights into their structural composition. Each output line contained 70 amino acids, and the sequence length was noted (Supplementary Tables 2, 3). Additionally, the amino acid composition was analyzed using the ProtParam Tool from ExPASy, providing insights into the properties of these holins, revealed that 55–63% of the secondary structures are in alpha helix forms. Alpha-helices are known for their stability and robustness to mutations, making them a common structural motif in proteins. The presence of α -helices is often associated with specific functional roles, such as membrane spanning and protein–protein interactions (Langosch et al., 2011). Amino acid composition analysis revealed the hydrophobicity, charge distribution, and potential post-translational modifications of these holins. All of which are crucial for understanding the protein's behavior and interaction with other molecules.

Determination and validation of tertiary-structure of the holin

The target sequence of the holin proteins in FASTA format was inserted into the HHpred Template Selection tool as the input, and the most suitable template (A0A655MZF6.1. A, P77237.1. A and P0A9R2.1. A) was selected with a probability rate of 100%, an E-Value of 2.4×10^{-116} , a Cols of 342, and a target length of 372 and finally stored the tertiary modeled protein structure in PDB format, as predicted by Swiss model (Figure 3) (Gallego-Páramo et al., 2021). The Ramachandran plot by PROCHECK (Mondal et al., 2021) (Figure 4) was used to assess the holins's tertiary structure, which revealed that for sequence PROKKA_03659; the 90 percent of the total residues (46) were found in the core (A, B, L); 7.5 percent of residues were in the additional allowed regions (a, b, l, p); and 0.0 percent of residues were in the generously allowed regions (a, b, l, p). The total number of non-glycine and non-proline residues were 40; that of the end-residues (excluding Gly and Pro) was 2; that of the glycine and proline residues was 4 and 0, respectively, out of 46 total residues. For the sequence PROKKA_04292 96.9 percent of the total residues (38) were found in the core (A, B, L); no residues were found in the additional allowed

regions (a, b, l, p); and in the generously allowed regions (a, b, l, p). The total number of non-glycine and non-proline residues was 32 and 2 respectively; that of the end-residues (excluding Gly and Pro) was 2; that of the glycine and proline residues was 3 and 1, respectively, out of 38 total residues. For the sequence PROKKA_04422 91.2 percent of the total residues (31) were found in the core (A, B, L); 8.8 percent residues were found in the additional allowed regions (a, b, l, p); and no residues were found in the generously allowed regions (a, b, l, p). The total number of non-glycine and non-proline residues were 3 and 1 respectively; that of the end-residues (excluding Gly and Pro) was 2; that of the glycine and proline residues was 3 and 1, respectively, out of 40 total residues (Supplementary Table 4). This analysis confirmed the quality and accuracy of the predicted structure by checking the distribution of dihedral angles in protein conformations, ensuring structural reliability (Laskowski et al., 1993).

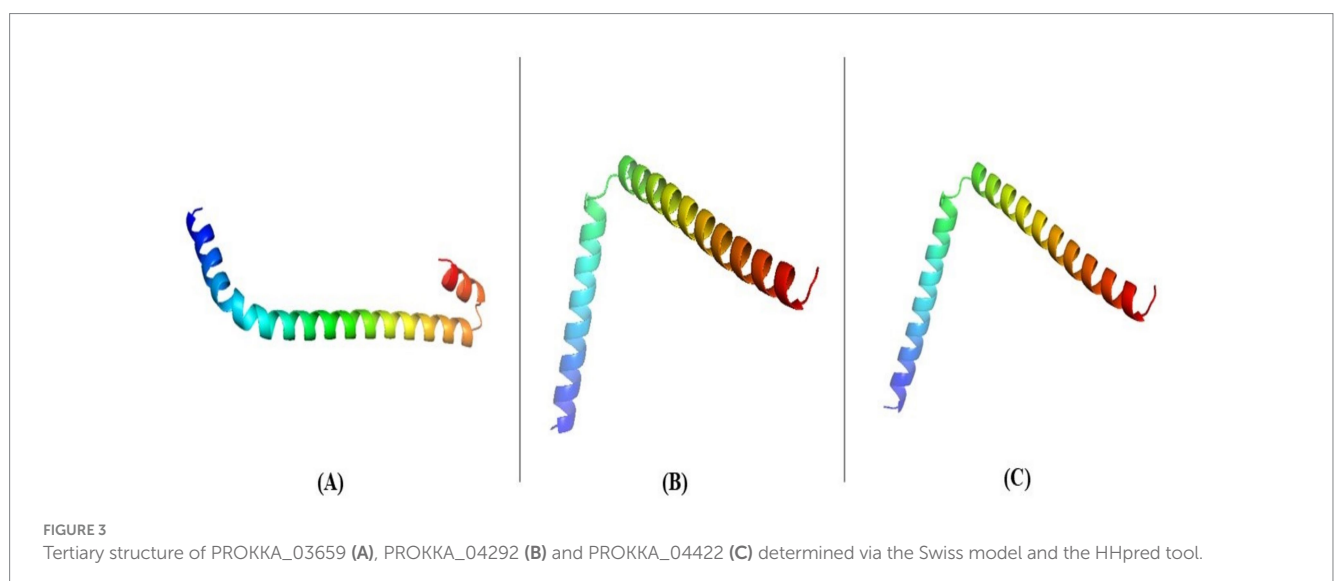
The structural stability and favorable conformations of these holins imply that they are capable of adopting the necessary membrane-insertion conformations to facilitate pore formation in bacterial membranes (Lella et al., 2015).

In silico analysis of the selected proteins

The physicochemical characterization of the three holins from coliphage ASEC2201 reveals essential insights into their stability and functional properties. The amino acid sequences were obtained in FASTA format and used for analysis. The stability indices of the holins indicated that all three proteins are stable, with PROKKA_03659 having an index of 28.95, PROKKA_04292 at 5.53, and PROKKA_04422 at 2.76, signifying their potential efficacy in therapeutic applications.

The theoretical isoelectric points (pI) for these proteins are in the basic range, with values of 9.43, 9.52, and 9.57, respectively, suggesting their roles may be influenced by pH conditions. Their aliphatic indices, above 66, indicate a high thermal stability, with PROKKA_03659, PROKKA_04292, and PROKKA_04422 values of 85.49, 97.65, and 92.11, respectively (Table 1).

The GRAVY index values, all >0 (0.183, 0.101, and 0.61), indicate these proteins are hydrophobic, suggesting a likelihood of less water



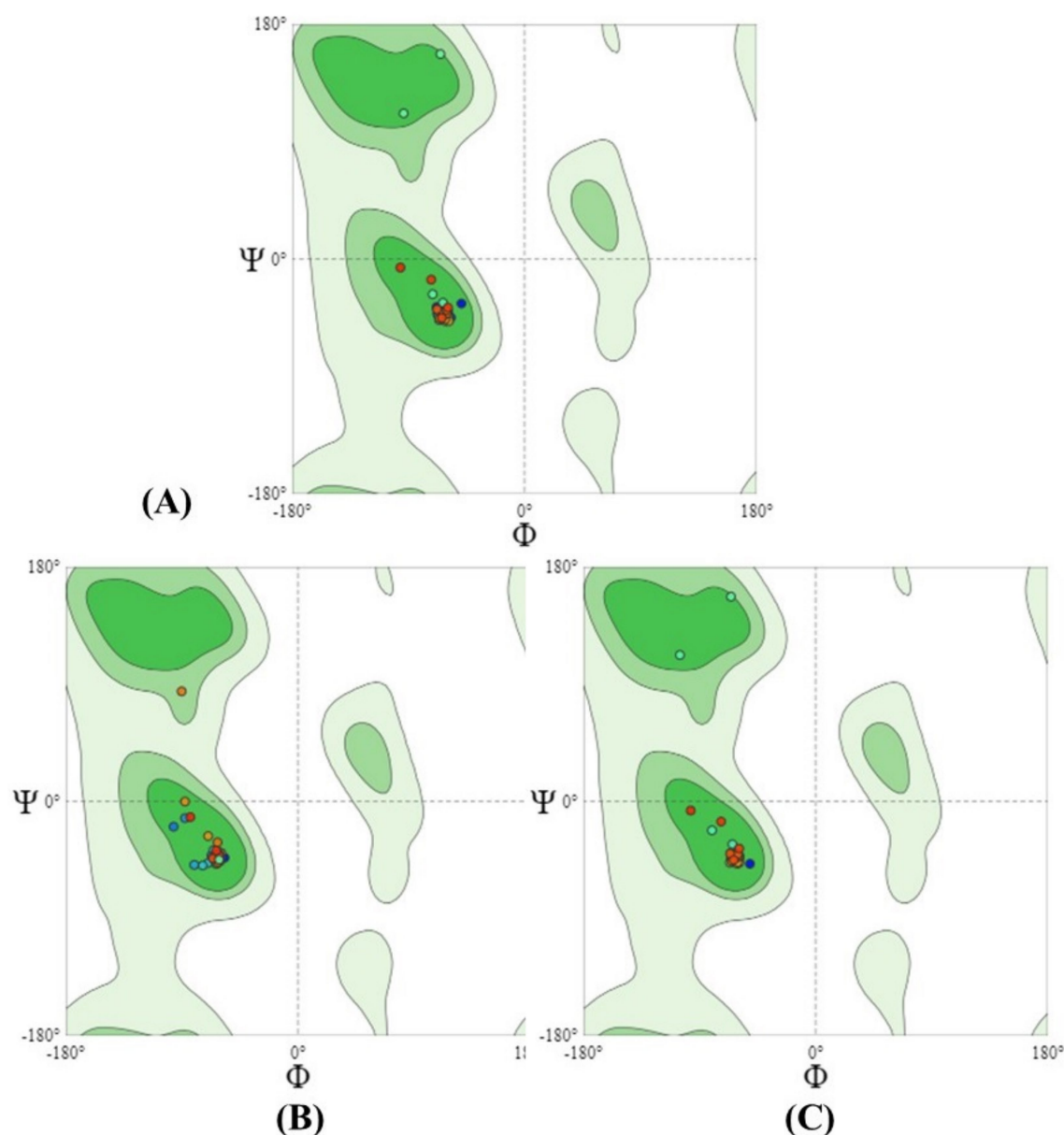


FIGURE 4

The PROCHECK program validated the Ramachandran plot statistics of the three-dimensional protein structure predicted by Swiss model. The graph's red areas represent the locations that are most permissible. Green, light green, and white fields are, respectively, demarcated as additional allowed, generously allowed, and restricted regions [in the figure above, (A) is PROKKA_03659, (B) is PROKKA_04292 and (C) is PROKKA_04422].

interaction and potential membrane association. These characteristics highlight the promising nature of these holins as candidates for further exploration in phage therapy and biocontrol applications (Singh et al., 2022).

Protein homology/analogy recognition

The structural analysis of the coliphage holin proteins, predicted using the Protein Homology/Analogy Recognition Engine - Phyre2, reveals significant insights into their functional characteristics (Supplementary Figure 3). Phyre2 modeling suggests that all three holins adopt predominantly helical conformations, indicative of

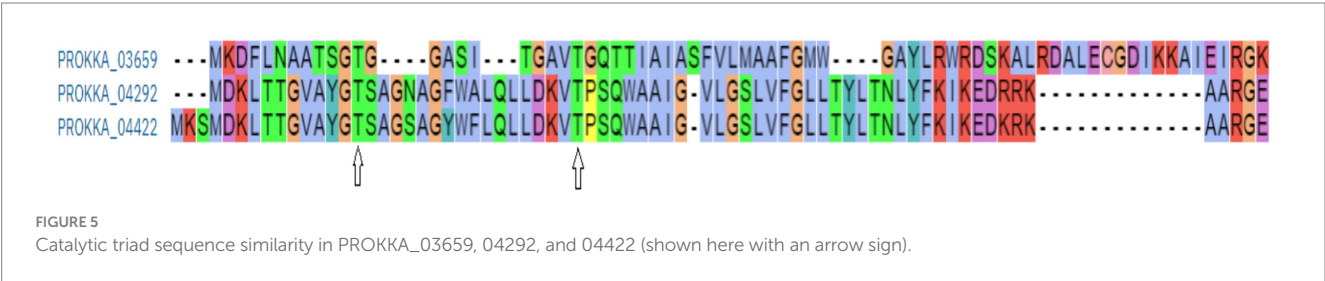
α -helical transmembrane domains. This structural feature is characteristic of holin proteins, which typically function by forming pores in bacterial membranes, facilitating the release of endolysins and subsequent cell lysis. Further comparative analysis indicates that these holins exhibit high structural homology to known proteins: PROKKA_03659 aligns with c8vxqB, PROKKA_04292 with c8jjrm, and PROKKA_04422 with c6hwhX. These homologous relationships suggest that the ASEC2201 holins share evolutionary lineage with established holin families, providing a framework for understanding their functional roles and potential applications in bacteriophage therapy. The identification of conserved residues, such as Glu-2, Asp-4, and Thr-6, forming a (Figure 5 and Table 2) analogous to that found in lysozymes, further underscores the functional significance of these

TABLE 1 Physico-chemical parameters of holins of ASEC2201 characterized by *in silico* methods.

Parameters	PROKKA_03659 holin	>PROKKA_04292 holin	>PROKKA_04422 holin
No. of amino acids	71	68	71
Molecular weight	7482.68	7394.56	7778.06
Theoretical Pi	9.43	9.52	9.57
Total number of negatively charged residues (Asp + Glu)	6	5	5
Total number of positively charged residues (Arg + Lys)	9	8	9
Formula	C ₃₃₁ H ₅₃₄ N ₉₂ O ₉₇ S ₄	C ₃₃₉ H ₅₃₄ N ₈₈ O ₉₅ S ₁	C ₃₅₈ H ₅₆₃ N ₈₉ O ₁₀₀ S ₂
Total number of atoms	1,058	1,057	1,112
The estimated half-life	30 h (mammalian reticulocytes, <i>in vitro</i>). >20 h (yeast, <i>in vivo</i>). >10 h (<i>Escherichia coli</i> , <i>in vivo</i>).	30 h (mammalian reticulocytes, <i>in vitro</i>). >20 h (yeast, <i>in vivo</i>). >10 h (<i>Escherichia coli</i> , <i>in vivo</i>).	30 h (mammalian reticulocytes, <i>in vitro</i>). >20 h (yeast, <i>in vivo</i>). >10 h (<i>Escherichia coli</i> , <i>in vivo</i>).
Instability index (II)	28.95 (Stable)	5.53(stable)	2.76 (Stable)
Aliphatic index	85.49	97.65	92.11
Grand average of hydropathicity (GRAVY)	0.183	0.101	0.061

TABLE 2 Catalytic triad of ASEC2201 holins.

Key residues	PROKKA_03659	PROKKA_04292	PROKKA_04422
Glu (E)	2	2	2
Asp (D)	4	3	3
Thr (T)	6	6	6



proteins. While holins are not enzymes themselves, the presence of such conserved motifs may indicate a regulatory mechanism or interaction with enzymatic partners, enhancing the efficiency of the lytic process.

Functional annotation of the selected holins

The identification of a conserved domain in the three holin proteins from coliphage ASEC2201 underscores their functional significance within the phage lifecycle. Utilizing the NCBI Conserved Domain Database (CDD) and CD-Search, which employs RPS-BLAST to compare sequences against position-specific scoring matrices derived from conserved domain alignments, a conserved domain was detected in these holins corresponding to the Phage_holin_2_1 superfamily (Pfam accession no. pfam04971). It has two transmembrane segments with both the N- and C-termini on the cytoplasmic side of the inner membrane in *E. coli*. Members of this family fall into the holin superfamily II, and Phage 21 S holin is the prototype for this

superfamily. Class II holin family protein which is similar to *E. coli* prophage lysis protein S homolog EssQ and prophage lysis protein S homolog EssD, meaning *E. coli*, the prophage-encoded proteins EssQ and EssD are homologous to Class II holins. EssQ is a predicted membrane protein, while EssD is a hypothetical protein associated with the DLP12 prophage (Srividhya and Krishnaswamy, 2012). These proteins are believed to function in the lysis process during prophage induction. The EssQ protein is predicted to have a role in the lysis of the host cell, similar to other holin-like proteins (Wang et al., 2023).

In silico genome-wide analysis of lysis system

The study of holins of coliphage ASEC2201 involved several bioinformatic tools to analyze its protein and genetic structure. ClustalW2 was employed to align its amino acid sequences with homologous proteins, providing insight into evolutionary relationships. For membrane topology and signal peptide predictions,

TABLE 3 Promoter predictions of the modeled holin proteins using BPROM.

Promoter parameters	PROKKA_03659	PROKKA_04292	PROKKA_04422
Sequence length	216	207	216
Number of predicted promoters	1	1	1
Promoter position	112, LDF 0.74 No such sites for promoter at 112	177, LDF 2.18 rpoD17: AATCTTTA at position 154 Score- 7 purR: ATTTCAAG at position 161 Score- 9	186, LDF 2.40 purR: ATTTCAAG at position 170 Score- 9 soxS: GATAAGCG at position 187 Score- 9

TABLE 4 Cell penetrating peptide predictions of selected holin proteins using CPP web server and their peptide stretches.

Holins' sequence	Peptide Sequence	SVM score	Prediction	Hydrophobicity	Hydropathicity	Hydrophilicity	Charge	Mol wt
PROKKA_03659	GMWGAYLRWR	0.03	CPP	−0.14	−0.54	−0.67	2.00	1295.67
	MWGAYLRWRD	0.03	CPP	−0.23	−0.85	−0.37	1.00	1353.70
	GAYLRWRDSK	0.26	CPP	−0.43	−1.42	0.43	2.00	1251.54
	YLRWRDSKAL	0.07	CPP	−0.39	−1.00	0.25	2.00	1307.65
	LRWRDSKALR	0.24	CPP	−0.57	−1.32	0.78	3.00	1300.66
	RWRDSKALRD	0.09	CPP	−0.69	−2.05	1.26	2.00	1302.58
>PROKKA_04292 holin	GFWALQLLDK	0.07	CPP	0.05	0.38	−0.56	0.00	1190.56
	YFKIKEDRRK	0.05	CPP	−0.68	−2.17	1.44	3.00	1382.76
	FKIKEDRRKA	0.21	CPP	−0.66	−1.86	1.62	3.00	1290.66
	KIKEDRRKAA	0.32	CPP	−0.69	−1.96	1.82	3.00	1214.56
	IKEDRRKAAR	0.09	CPP	−0.76	−2.02	1.82	3.00	1242.57
	KEDRRKAARG	0.07	CPP	−0.82	−2.51	2.00	3.00	1186.46
>PROKKA_04422 holin	YFKIKEDKRK	0.07	CPP	−0.61	−2.11	1.44	3.00	1354.75
	FKIKEDKRKA	0.22	CPP	−0.59	−1.80	1.62	3.00	1262.65
	KIKEDKRKAA	0.25	CPP	−0.63	−1.90	1.82	3.00	1186.55
	KEDKRKAARG	0.02	CPP	−0.75	−2.45	2.00	3.00	1158.45

TMHMM and the SignalP server were used, identifying potential transmembrane domains (Petersen et al., 2011). The 3D structure of holins of ASEC2201 has been predicted using Phyre2, a protein modeling tool. Conserved domains were detected via the NCBI Conserved Domain Database (CDD), which helped pinpoint functional elements within the proteins. Promoter regions were predicted using Bprom, aiding in understanding regulatory sequences (Table 3). Such a multi-layered analysis helps unravel phage biology and potential therapeutic applications (Wu et al., 2021).

Cell penetrating peptide predictions

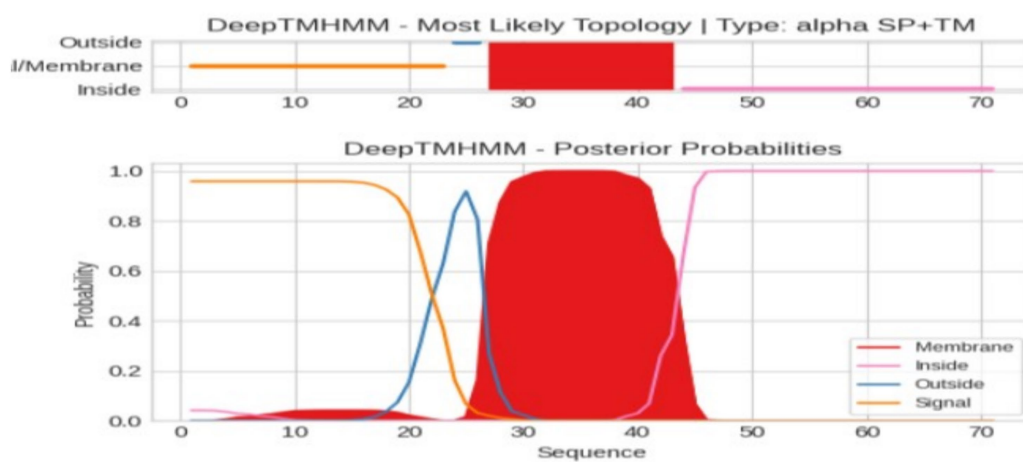
Cell penetrating peptides have been identified and selected using CPP web server. Six peptide stretches were identified in PROKKA_03659, PROKKA_04292 and four peptide stretches have been identified in PROKKA_04422 holin protein, respectively (Table 4). These peptides, located predominantly in the N-terminal bactericidal regions of the proteins, exhibit characteristics typical of CPPs, such as a high density of positively charged amino acids. This structural feature facilitates their ability to traverse cellular membranes efficiently.

The presence of CPPs in holin proteins suggests their potential as vectors for intracellular delivery of therapeutic agents. Due to their

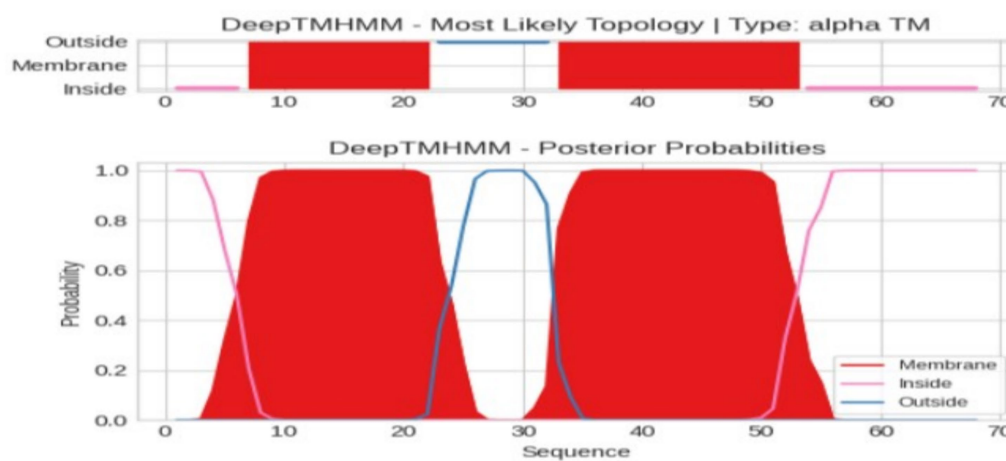
positive charge, these peptides can form complexes with negatively charged molecules like nucleic acids, enhancing the delivery of genetic materials into cells. Additionally, CPPs can be conjugated with various therapeutic biomolecules, including peptides, proteins, and antimicrobial drugs, thereby expanding their utility in targeted drug delivery systems.

Protein transmembrane domain recognition by deep transmembrane helices hidden Markov models (DeepTMHMM)

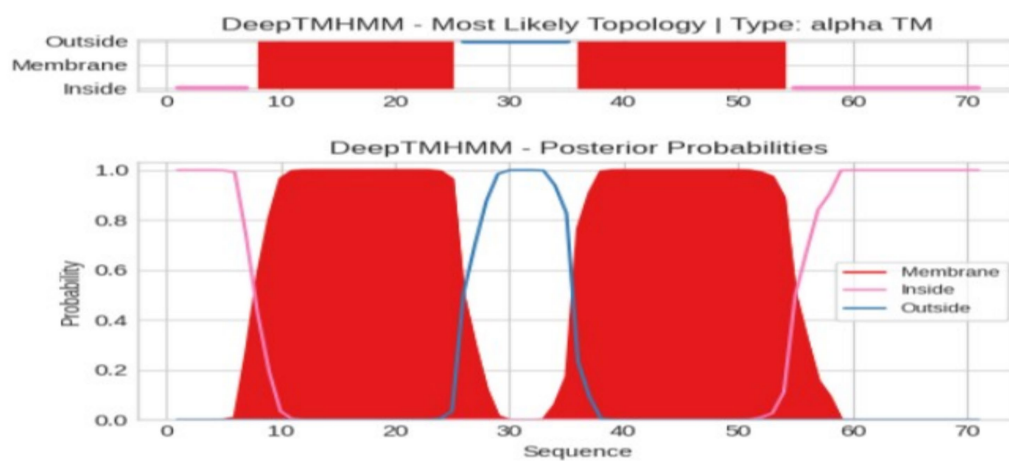
Transmembrane proteins span the lipid bilayer and are divided into two major structural classes, namely alpha helical and beta barrels. TMHMM analysis as represented in Figure 6 indicated that PROKKA_03659 contains a single transmembrane domain (TMD), while PROKKA_04292 and PROKKA_04422 each possess two putative TMDs. Holins are classified into three groups based on their TMD number and topology. Class I holins typically have three TMDs, class II holins contain two TMDs, and class III holins possess only one TMD. Most phages encode class I or II holins, whereas class III holins, with a single TMD, are less common, as reported by Moussa et al. (2014). The N-terminal of PROKKA_03659 was found to



(A)



(B)



(C)

FIGURE 6

Protein transmembrane domain recognition by deep transmembrane helices hidden Markov models (DeepTMHMM), where (A) represents the result of PROKKA_03659, (B) for PROKKA_04292 and (C) for PROKKA_04422, respectively.

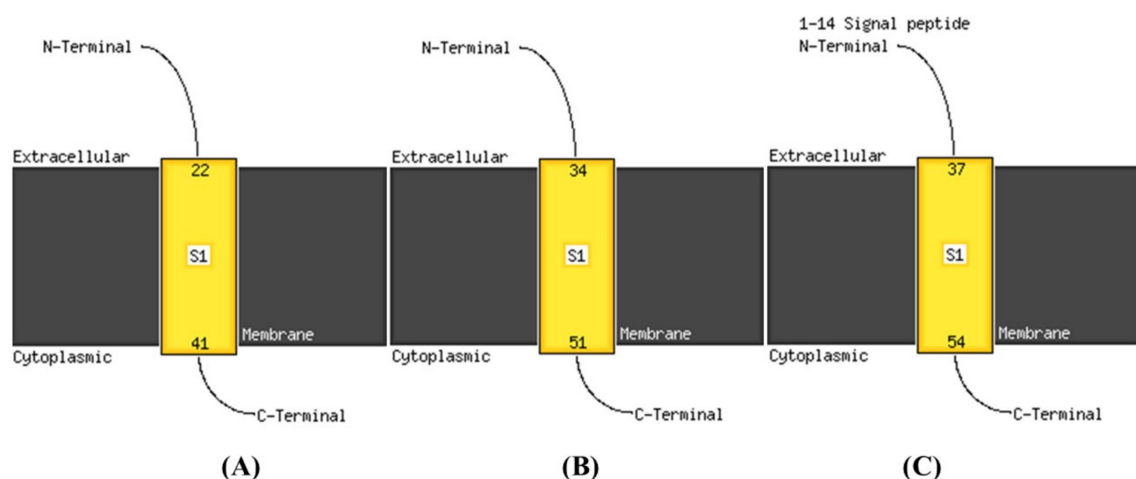


FIGURE 7
Predicted transmembrane N-terminal regions of annotated proteins PROKKA_03659 (A); PROKKA_04292 (B); PROKKA_04422 (C), via Phyre2 Modeling.

be positioned in extracellular region at 22nd amino acid and C-terminal positioned in the cytoplasmic region of 41st amino acid. Moreover, PROKKA_04292 and PROKKA_04422 belong to class II holin proteins with two transmembrane domains. Here N-terminal positioned in extracellular region at 34th amino acid and C-terminal positioned the cytoplasmic region of 51st amino acid of PROKKA_04292. However, in case of PROKKA_04422, the N-terminal position has been found in extracellular region at 37th amino acid and C-terminal was found in the cytoplasmic region of 54th amino acid (as shown in Figure 7). Based on their TMD composition, PROKKA_03659 likely belongs to class III, while PROKKA_04292 and PROKKA_04422 may align with class II holins.

Discussion

The ASEC2201 coliphage, cataloged under accession number SRX17770782, in NCBI SRA represents a significant addition to our understanding of bacteriophage biology, particularly in the context of antibiotic resistance. Isolated from drug-resistant clinical strains of *E. coli* ASEC2201 underwent comprehensive characterization and whole-genome sequencing, with the resulting genomic data made available in the NCBI database. This foundational work allows for further exploration of the phage's functional components, particularly its lytic enzymes, which play a crucial role in the phage's ability to lyse bacterial cells. Among the lytic enzymes identified in ASEC2201, three holin proteins PROKKA_03659, PROKKA_04292, and PROKKA_04422 were characterized. Holins are integral membrane proteins that facilitate the release of endolysins into the bacterial periplasm, ultimately leading to cell lysis. The presence of active promoters associated with these proteins indicates significant transcriptional activity, underscoring their importance in the bacteriophage-mediated lysis system. Specifically, the identified promoters at positions 112, 177, and 186 for each respective protein suggest a well-regulated expression system that enables the phage to effectively target and dismantle resistant bacterial strains. Stability indices were calculated for these holins, revealing that all three proteins exhibit stability, with values of 28.95 for PROKKA_03659, 5.53 for PROKKA_04292, and 2.76 for PROKKA_04422.

A higher stability index often correlates with the potential efficacy of proteins in therapeutic applications, as stable proteins are less likely to denature under physiological conditions (Heselpoth et al., 2015; Abdelghafar et al., 2023). This stability is crucial for their functionality, especially when considering the potential therapeutic applications of phage therapy against resistant bacterial infections (Zaki et al., 2023; Bagińska et al., 2024). The theoretical isoelectric points (pI) for these holins fall within the basic range, with values of 9.43, 9.52, and 9.57, respectively. This basicity suggests that the functional roles of these proteins may be influenced by pH conditions, which is particularly relevant in the variable microenvironments within the human body. The aliphatic indices of these holins, all exceeding 66, indicate a high degree of thermal stability, which enhances their potential as therapeutic agents. The specific values of 85.49 for PROKKA_03659, 97.65 for PROKKA_04292, and 92.11 for PROKKA_04422 reflect their structural resilience, making them promising candidates for further study. Wang et al. (2025) have reported that the holin protein from *Acinetobacter nosocomialis* phage XC1 (ORF: XC1_81) exhibited an aliphatic index of 109.12 and an instability index of 21.36, concluding high thermal stability and structural resilience of protein. Additionally, the GRAVY index values for these proteins, all greater than 0 (specifically, 0.183 for PROKKA_03659, 0.101 for PROKKA_04292, and 0.61 for PROKKA_04422), suggest that they are hydrophobic in nature. This hydrophobic characteristic indicates a likelihood of reduced water interaction and potential membrane association, which is crucial for their function in transmembrane lysis. The presence of a hydrophobic region in holin proteins is critical for their ability to interact with the inner bacterial membrane, ultimately leading to the formation of pores that allow endolysins to access the peptidoglycan layer. Through domain analysis, these holins were classified within the Phage_holin_2_1 superfamily (Reddy and Saier, 2013) characterized by small hydrophobic proteins that are essential for the lytic cycle of bacteriophages. The identification of this superfamily highlights the evolutionary conservation of these proteins and their critical roles in phage biology.

The use of DeepTMHMM in the transmembrane domain (TMD) analysis of holin proteins in bacteriophage ASEC2201 represents a significant advancement in accurately predicting membrane topology.

Previous researches have utilized various methods to predict holin TMDs. For instance, studies on *E. coli* phage ECP26 employed TMHMM to identify holin candidates based on TMD presence and length, leading to the identification of ORF151 as a class II holin. Similarly, investigations into Lactobacillus phages revealed a prevalence of class III holins with single TMDs, highlighting the diversity in holin structures across different phage species (Park et al., 2022). The findings from ASEC2201 holin analysis align with structural predictions of other phage-derived therapeutic proteins. For example, the mycobacteriophage D29 holin has been shown to possess two TMDs with a coiled-coil C-terminal region, essential for efficient host cell lysis. Similarly, the phage ϕ 21 holin S^Δ21 utilizes its second TMD to form small membrane pores, facilitating lysis (Bavda et al., 2021).

Integrating cell-penetrating peptides (CPPs) with holins enhances phage mediated lysis against intracellular bacterial infections. Zhao et al. (2023) engineered Salmonella phages displaying CPPs, improving their uptake by epithelial cells and effectively eradicating intracellular bacteria without harming host cells. Similarly, Wang et al. (2018) fused a CPP to the phage lysin Jdlys, facilitating its entry into keratinocytes and eliminating intracellular MRSA, thereby accelerating wound healing in mice. The three holins of ASEC2201 under study, possess specific N-terminal regions which exhibit CPP-like properties, suggesting potential for delivering holin activity directly to intracellular pathogens.

The structural characterization of the three identified holin proteins of ASEC2201 revealed that the holin sequences were aligned with high-confidence templates (e.g., A0A655MZF6.1. A, P77237.1. A, P0A9R2.1. A), yielding a 100% probability and an exceptionally low E-value of 2.4×10^{-116} . This indicates a robust structural similarity to known holin proteins, reinforcing the hypothesis that these sequences encode functional holins. Subsequent validation of Ramachandran plot analysis revealed high percentages in the favored regions underscore the structural integrity of the holins, suggesting that they are well-folded and likely to be functionally active. The comprehensive structural analysis of the ASEC2201 holins provides a solid foundation for their potential application in phage therapy, particularly against multidrug-resistant *E. coli* strains. Their well-defined tertiary structures and validation through Ramachandran plot analysis affirm their suitability as candidates for drug delivery vectors (Hou et al., 2018; Bavda et al., 2021; Pollenz et al., 2022; Li et al., 2023; Tang et al., 2025).

Conclusion

In summary, the detailed genomic and biophysical characterization of ASEC2201 holins (PROKKA_03659, PROKKA_04292, PROKKA_04422) underscores their utility as modular membrane-disruption devices in engineered antimicrobial platforms. Their confirmed transmembrane topologies, robust stability profiles, and CPP-like N-terminal signals position these proteins not only as precise lytic effectors in phage therapy against multidrug-resistant *E. coli*, but also as versatile carriers for intracellular cargo delivery. By leveraging strong, promoter-driven expression and the inherent specificity of holin-endolysin pairs, next-generation biotechnological applications could integrate holin modules into synthetic vesicles, fusion proteins, or conjugated nanocarriers, enabling targeted breaching of bacterial membranes or host-cell endosomal compartments. Ultimately, ASEC2201 holins represent a powerful toolkit for designing bespoke antimicrobial and drug-delivery systems that combine genetic programmability with membrane-permeabilizing precision.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found at: <https://www.ncbi.nlm.nih.gov/>, accession number: SRX17770783.

Author contributions

HS: Data curation, Formal analysis, Investigation, Methodology, Writing – original draft. SP: Conceptualization, Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft. AS: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review & editing. SPS: Supervision, Visualization, Writing – review & editing. MHS: Project administration, Resources, Supervision, Writing – review & editing. MS: Supervision, Writing – review & editing. IH: Supervision, Writing – review & editing. MB: Visualization, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The financial support received to Dr. I Kostova by the European Union-Next Generation EU, through the National Recovery and Resilience Plan of the Republic of Bulgaria, project No. BG-RRP-2.004-0004-C01 is greatly acknowledged.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2025.1550594/full#supplementary-material>

References

- Abdelghafar, A., El-Ganiny, A., Shaker, G., and Askoura, M. (2023). A novel lytic phage exhibiting a remarkable in vivo therapeutic potential and higher antibiofilm activity against *Pseudomonas aeruginosa*. *Eur. J. Clin. Microbiol. Infect. Dis.* 42, 1207–1234. doi: 10.1007/s10096-023-04649-y
- Abeyssekera, G. S., Love, M. J., Manners, S. H., Billington, C., and Dobson, R. C. J. (2022). Bacteriophage-encoded lethal membrane disruptors: advances in understanding and potential applications. *Front. Microbiol.* 13:1044143. doi: 10.3389/fmicb.2022.1044143
- Ali, A., Bandaranyake, R. M., Cai, Y., Davis, D. A., Amano, M., Li, B., et al. (2020). Structure-based design of highly potent HIV-1 protease inhibitors containing new tricyclic ring P 2-ligands: design, synthesis, biological, and X-ray structural studies. *J. Med. Chem.* 63, 6847–6860. doi: 10.1021/acs.jmedchem.0c00202
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Bagińska, N., Grygiel, I., Orwat, F., Harhala, M. A., Jędrusiak, A., Gębarowska, E., et al. (2024). Stability study in selected conditions and biofilm-reducing activity of phages active against drug-resistant *Acinetobacter baumannii*. *Sci. Rep.* 14:4285. doi: 10.1038/s41598-024-54469-z
- Bavda, V. R., Yadav, A., and Jain, V. (2021). Decoding the molecular properties of mycobacteriophage D29 holin provides insights into holin engineering. *J. Virol.* 95, 2173–2120. doi: 10.1128/JVI.02173-20
- Borrelli, A., Tornesello, A. L., Tornesello, M. L., and Buonaguro, F. M. (2018). Cell penetrating peptides as molecular carriers for anti-cancer agents. *Molecules* 23:295. doi: 10.3390/molecules23020295
- Cahill, J., and Young, R. (2019). Phage lysis: multiple genes for multiple barriers. *Adv. Virus Res.* 103, 33–70. doi: 10.1016/bs.aivir.2018.09.003
- Chen, Y., and Young, R. (2016). The last r locus unveiled: T4 RIII is a cytoplasmic Antiholin. *J. Bacteriol.* 198, 2448–2457. doi: 10.1128/JB.00294-16
- Copolovici, D. M., Langel, K., Eriste, E., and Langel, Ü. (2014). Cell-penetrating peptides: design, synthesis, and applications. *ACS Nano* 8, 1972–1994. doi: 10.1021/nn4057269
- de Castro, E., Sigrist, C. J., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P. S., Gastegger, E. (2006). Scan Prosite: detection of PROSITE signature matches and pro rule-associated functional and structural residues in proteins. *Nucleic Acids Res.* 34, W362–W365. doi: 10.1093/nar/gkl124
- Du, J. J., Zhang, R. Y., Jiang, S., Xiao, S., Liu, Y., Niu, Y., et al. (2025). Applications of cell penetrating peptide-based drug delivery system in immunotherapy. *Front. Immunol.* 16:1540192. doi: 10.3389/fimmu.2025.1540192
- Gallego-Páramo, C., Hernández-Ortiz, N., Buey, R. M., Rico-Lastres, P., García, G., Díaz, J. F., et al. (2021). Structural and functional insights into Skl and pal endolysins, two cysteine-amidases with anti-pneumococcal activity. Dithiothreitol (DTT) effect on lytic activity. *Front. Microbiol.* 12:740914. doi: 10.3389/fmicb.2021.740914
- Gautam, A., Chaudhary, K., Kumar, R., Sharma, A., Kapoor, P., Tyagi, A., et al. (2013). In silico approaches for designing highly effective cell penetrating peptides. *J. Transl. Med.* 11:74. doi: 10.1186/1479-5876-11-74
- Grabowski, L., Lepek, K., Stasiłojć, M., Kosznik-Kwaśnicka, K., Zdrojewska, K., Maciąg-Dorszyńska, M., et al. (2021). Bacteriophage-encoded enzymes destroying bacterial cell membranes and walls, and their potential use as antimicrobial agents. *Microbiol. Res.* 248:126746. doi: 10.1016/j.micres.2021.126746
- Hallgren, J., Tsiirigos, K. D., Pedersen, M. D., Armenteros, J. J. A., Marcatili, P., Nielsen, H., et al. (2022). Deep TMHMM predicts alpha and beta transmembrane proteins using deep neural networks. *bioRxiv* 4:487609. doi: 10.1101/2022.04.08.487609
- Heselpoth, R. D., Owens, J. M., and Nelson, D. C. (2015). Quantitative analysis of the thermal stability of the gamma phage endolysin Ply G: a biophysical and kinetic approach to assaying therapeutic potential. *Virology* 477, 125–132. doi: 10.1016/j.virol.2014.11.003
- Hou, R., Li, M., Tang, T., Chen, R., Ma, J., Gao, J., et al. (2018). Construction of *Lactobacillus casei* ghosts by holin-mediated inactivation and the potential as a safe and effective vehicle for the delivery of DNA vaccines. *BMC Microbiol.* 18:80. doi: 10.1186/s12866-018-1216-6
- Jakočič, D., and Moodley, A. (2018). A Rapid Bacteriophage DNA Extraction Method. *Protoc. J.* 27, doi: 10.3390/mps1030027
- Kelley, L., Mezulis, S., Yates, C., Kelley, L. A., Yates, C. M., Wass, M. N., et al. (2015). The phyre 2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858. doi: 10.1038/nprot.2015.053
- Krieger, I. V., Kuznetsov, V., Chang, J. Y., Zhang, J., Moussa, S. H., Young, R. F., et al. (2020). The structural basis of T4 phage lysis control: DNA as the signal for lysis inhibition. *J. Mol. Biol.* 432, 4623–4636. doi: 10.1016/j.jmb.2020.06.013
- Kumar, A., Purohit, R., Sheokand, N., Singh, A., Singh, R. P., and Bisht, D. (2021). Domain-wise differentiation of *Mycobacterium tuberculosis* H37Rv hypothetical proteins: a roadmap to discover bacterial survival potentials. *IUBMB Life* 73, 498–512. doi: 10.1002/iub.2436
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096
- Langosch, D., Herrmann, J. R., Unterreitmeier, S., Fuchs, A. (2011). “Helix-helix interaction patterns in membrane proteins” in Structural bioinformatics of membrane proteins. ed. D. Frishman (Vienna: Springer), 165–186.
- Laskowski, R. A., Mac Arthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291. doi: 10.1107/s0021889892009944
- Lauer, P. M., Dubensky, T. W., Luckett, W. S., and Hanson, W. G. (2010). Holin-enhanced vaccines and reagents, and methods of use thereof (US patent no. US 201001294.06A1). Alexandria, VA: U.S. Patent and Trademark Office.
- Lella, M., Kamilla, S., Jain, V., Mahalakshmi, R. (2015). Molecular mechanism of Holin transmembrane domain I in pore formation and bacterial cell death. *ACS Chem. Biol.* 11, 910–920. doi: 10.1021/acschembio.5b00875
- Li, X., Chen, Y., Wang, L., Zhao, H., Zhou, Y., Liu, J., et al. (2023). Synergistic action of phages and lytic proteins with antibiotics: a combination strategy to target bacteria and biofilms. *BMC Microbiol.* 23:149. doi: 10.1186/s12866-023-02881-2
- Liu, H., Hu, Z., Li, M., Yang, Y., Lu, S., and Rao, X. (2023). Therapeutic potential of bacteriophage endolysins for infections caused by gram-positive bacteria. *J. Biomed. Sci.* 30:29. doi: 10.1186/s12929-023-00919-1
- Masih, A., and Singh, A. (2023). Bacteriophage endolysins and their role in eradication of bacterial biofilms. *Nov. Res. Microbiol. J.* 7, 1995–2014. doi: 10.21608/nrmj.2023.304309
- Mondal, S. I., Akter, A., Draper, L. A., Ross, R. P., and Hill, C. (2021). Characterization of an Endolysin targeting Clostridioides difficile that affects spore outgrowth. *Int. J. Mol. Sci.* 22:5690. doi: 10.3390/ijms22115690
- Moreno-Vargas, L. M., and Prada-Gracia, D. (2024). Exploring the chemical features and biomedical relevance of cell-penetrating peptides. *Int. J. Mol. Sci.* 26:59. doi: 10.3390/ijms26010059
- Moussa, S. H., Lawler, J. L., and Young, R. (2014). Genetic dissection of T4 lysis. *J. Bacteriol.* 196, 2201–2209. doi: 10.1128/jb.01548-14
- Murray, E., Draper, L. A., Ross, R. P., and Hill, C. (2021). The advantages and challenges of using Endolysins in a clinical setting. *Viruses* 13:680. doi: 10.3390/v13040680
- Nam, S. H., Park, J., and Koo, H. (2023). Recent advances in selective and targeted drug/gene delivery systems using cell-penetrating peptides. *Arch. Pharm. Res.* 46, 18–34. doi: 10.1007/s12272-022-01425-y
- Nandy, A., Yadav, R., and Singh, A. (2022). Phage derived lytic peptides, a secret weapon against *Acinetobacter baumannii*-an in silico approach. *Front. Med.* 9:1047752. doi: 10.3389/fmed.2022.1047752
- Ndeboko, B., Omuouessi, S. T., Ongali, B., and Mouinga Ondémé, A. (2020). A cell penetrating peptides used in delivery of therapeutic oligonucleotides targeting hepatitis B virus. *Pharmaceuticals* 13:483. doi: 10.3390/ph13120483
- Nedvedova, S., Guillié, F., Miele, A. E., Cantrelle, F. X., Dvorak, J., Walker, O., et al. (2023). Divide, conquer and reconstruct: how to solve the 3D structure of recalcitrant Micro-exon gene (MEG) protein from *Schistosoma mansoni*. *PLoS One* 18:e0289444. doi: 10.1371/journal.pone.0289444
- Padmesh, S., Singh, A., Chopra, S., Sen, M., Habib, S., Shrivastava, D., et al. (2024). Isolation and characterization of novel lytic bacteriophages that infect multi drug resistant clinical strains of *Escherichia coli*. *Environ. Sci. Pollut. Res. Int.* 31, 57327–57337. doi: 10.1007/s11356-023-28081-z
- Pancholi, B., Choudhary, M. K., Kumar, M., Babu, R., Vora, L., Khatri, D., et al. (2025). Cell-penetrating proteins and peptides as a promising therapeutic agent for neurodegenerative disorder. *J. Drug Deliv. Sci. Technol.* 107:106816. doi: 10.1016/j.jddst.2025.106816
- Park, D. W., Lee, J. H., and Park, J. H. (2022). Prediction and identification of new type holin protein of *Escherichia coli* phage ECP26. *Food Sci. Biotechnol.* 31, 843–847. doi: 10.1007/s10068-022-01089-w
- Petersen, T. N., Brunak, G., von Heijne, J. H., and Nielsen, H. (2011). SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods.* 8, 785–786. doi: 10.1038/nmeth.1701
- Pollenz, R. S., Bland, J., and Pope, W. H. (2022). Bioinformatic characterization of endolysins and holin-like membrane proteins in the lysis cassette of phages that infect *Gordonia rubripertincta*. *PLoS ONE* 17:e0276603. doi: 10.1371/journal.pone.0276603
- Reddy, B. L., and Saier, M. H. Jr. (2013). Topological and phylogenetic analyses of bacterial holin families and superfamilies. *Biochim. Biophys. Acta* 1828, 2654–2671. doi: 10.1016/j.bbame.2013.07.004
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* Oxford, England 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Setten, R. L., Rossi, J. J., and Han, S. P. (2019). The current state and future directions of RNAi-based therapeutics. *Nat. Rev. Drug Discov.* 18, 421–446. doi: 10.1038/s41573-019-0017-4

- Shaath, H., Alfadda, A. A., Mohamed, G., Al-Khadairi, G., Al-Sheikh, Y. A., and Zayed, H. (2024). Identification of interpretable clusters and associated signatures in breast cancer single-cell data: a topic modeling approach. *Cancers* 16:1350. doi: 10.3390/cancers16071350
- Singh, A., Padmesh, S., Dwivedi, M., and Kostova, I. (2022). How good are bacteriophages as an alternative therapy to mitigate biofilms of nosocomial infections. *Infect Drug Resist.* 15, 503–532. doi: 10.2147/IDR.S348700
- Söding, J., Biegert, A., and Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 33, W244–W248. doi: 10.1093/nar/gki408
- Solovyev, V., and Salamov, A. (2011). “Automatic annotation of microbial genomes and metagenomic sequences” in *Metagenomics and its applications in agriculture, biomedicine and environmental studies*. ed. R. W. Li (Hauppauge, NY: Nova Science Publishers), 61–78.
- SRA. (2004). Available online at: <https://www.ncbi.nlm.nih.gov/sra/> (Accessed September 28, 2022).
- Srividhya, K. V., and Krishnaswamy, S. (2012). Characterization of DLP12 prophage membrane associated protein: holin GFP. *Indian J. Microbiol.* 52, 464–471. doi: 10.1007/s12088-012-0278-7
- Stothard, P. (2000). The sequence manipulation suite: Java script programs for analyzing and formatting protein and DNA sequences. *BioTechniques* 28, 1102–1104. doi: 10.2144/00286ir01
- Tang, W., Zhu, L., Chen, H., Li, Y., Zhang, Z., Wang, B., et al. (2025). Biological function identification of phage holin Hol-4086 and treatment of *Staphylococcus aureus* infection. *Front. Microbiol.* 16:1499566. doi: 10.3389/fmicb.2025.1499566
- Turner, D., Adriaenssens, E. M., Cresawn, S. G., Butler, A., Hunter, M., Smith, D., et al. (2021). A roadmap for genome-based phage taxonomy. *Viruses* 13:506. doi: 10.3390/v13030506
- Verma, S., Sharma, P., Kumar, S., Gupta, R., Yadav, R., Kumar, A., et al. (2024). Functional domains of *Acinetobacter* bacteriophage tail fibers. *Front. Microbiol.* 15:1230997. doi: 10.3389/fmicb.2024.1230997
- Wang, J., Chitsaz, F., Derbyshire, M. K., Gonzales, N. R., Gwadz, M., Lu, S., et al. (2023). The conserved domain database in 2023. *Nucleic Acids Res.* 51, D384–D388. doi: 10.1093/nar/gkac1096
- Wang, Z., Kong, L., Liu, Y., Fu, Q., Cui, Z., Wang, J., et al. (2018). A phage lysin fused to a cell-penetrating peptide kills intracellular methicillin-resistant *Staphylococcus aureus* in keratinocytes and has potential as a treatment for skin infections in mice. *Appl. Environ. Microbiol.* 84:e00380-18. doi: 10.1128/AEM.00380-18
- Wang, I. N., Smith, D. L., and Young, R. (2000). Holins: the protein clocks of bacteriophage infections. *Ann. Rev. Microbiol.* 54, 799–825. doi: 10.1146/annurev.micro.54.1.799
- Wang, C., Zhao, S., Jiang, H., Shi, H., Li, J., Zhao, C., et al. (2025). Biological characteristics and genomic analysis of *Acinetobacter nosocomialis* lytic phage XC1. *Curr. Issues Mol. Biol.* 47:335. doi: 10.3390/cimb47050335
- White, R., Tran, T. A., Dankenbring, C. A., Deaton, J., and Young, R. (2010). The N-terminal transmembrane domain of lambda S is required for holin but not antiholin function. *J. Bacteriol.* 192, 725–733. doi: 10.1128/JB.01263-09
- Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., et al. (1999). Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* 112, 531–552. doi: 10.1385/1-59259-584-7:531
- Wilson, D., Pethica, R., Zhou, Y., Talbot, C., Vogel, C., Madera, M., et al. (2009). SUPERFAMILY—sophisticated comparative genomics, data mining, visualization and phylogeny. *Nucleic Acids Res.* 37, D380–D386. doi: 10.1093/nar/gkn762
- Wu, Z., Zhang, Y., Xu, X., Ahmed, T., Yang, Y., Loh, B., et al. (2021). The Holin-Endolysin lysis system of the OP2-like phage X2 infecting *Xanthomonas oryzae* pv. *Oryzae*. *Viruses* 13:1949. doi: 10.3390/v13101949
- Yuan, Y., Liu, X., Wang, H., Zheng, J., Xie, L., Zhang, Y., et al. (2023). Pha GAA: an integrated web server platform for phage genome annotation and analysis. *Nucleic Acids Res.* 51, W548–W556. doi: 10.1093/nar/gkad293
- Zaki, B. M., Fahmy, N. A., Aziz, R. K., Samir, R., and El-Shibiny, A. (2023). Characterization and comprehensive genome analysis of novel bacteriophage, vB_Kpn_ZCKp20p, with lytic and anti-biofilm potential against clinical multidrug-resistant *Klebsiella pneumoniae*. *Front. Cell. Infect. Microbiol.* 13:1077995. doi: 10.3389/fcimb.2023.1077995
- Zhang, H., Li, Y., Chen, R., Wu, Z., Liu, J., Wang, Y., et al. (2023). Pha VIP: phage virion protein classification based on chaos game representation and vision transformer. *bioRxiv* 39, i30–i39. doi: 10.48550/arXiv.2301.12422
- Zhang, H., Wang, H., Zhang, S., Liu, Y., Zhou, Y., Wang, Y., et al. (2025). GOphe: protein function annotation for bacteriophages by integrating the genomic context. *Bioinformatics* 26:bbaf014. doi: 10.1093/bioinformatics/btae209
- Zhao, M., Tan, X., Liu, Z.-q., Dou, L., Liu, D., Pan, Y.-j., et al. (2023). Engineered phage with cell-penetrating peptides for intracellular bacterial infections. *mSystems* 8:e00646-23. doi: 10.1128/msystems.00646-23
- Zimmermann, L., Stephens, A., Nam, S. Z., Rau, D., Kübler, J., Lozajic, M., et al. (2018). A completely Reimplemented MPI bioinformatics toolkit with a new HHpred server at its Core. *J. Mol. Biol.* 430, 2237–2243. doi: 10.1016/j.jmb.2017.12.007
- Zorko, M., Jones, S., and Langel, Ü. (2022). Cell-penetrating peptides in protein mimicry and cancer therapeutics. *Adv. Drug Deliv. Rev.* 180:114044. doi: 10.1016/j.addr.2021.114044