

AI-Assisted Design of Antimicrobial Peptides Targeting *Klebsiella pneumoniae*

Yanling Lin¹

Received November 20, 2025

Accepted January 14, 2026

Electronic access January 31, 2026

Multidrug-resistant *Klebsiella pneumoniae* is a major driver of hospital mortality and healthcare costs, with a growing association with antibiotic treatment failures. Recent studies indicate that certain synthetic antimicrobial peptides can combat *K. pneumoniae* and represent a promising avenue for new therapeutics. Together, these findings underscore the urgent need for new antimicrobial development. This study aimed to computationally design novel antimicrobial peptides (AMPs) through two strategies using AI: (1) shortening amino acid length and (2) substituting polar uncharged or non-polar neutral residues and the structurally flexible glycine with positively charged or hydrophobic amino acids, while maintaining balanced hydrophobicity to improve predicted antimicrobial activity and reduce predicted toxicity. We chose a low-performing candidate with weak predicted AMP probability (predicted AMP probability=0.531) and high predicted toxicity (predicted toxicity hybrid score=0.5) in the GRAMPA database, and used AI-driven tools (CAMP3/4, AlphaFold, APD3, DBAASP, HeliQuest) to generate 243 variants via single-site substitutions and assess their properties. Top candidates were further optimized using ToxinPred3.0, generating 381 variants by introducing additional single amino acid substitutions in the highest-scoring sequence to reduce predicted toxicity while maintaining antimicrobial activity. The final peptide achieved a high predicted AMP probability score (predicted AMP probability=1) and zero predicted toxicity (predicted toxicity hybrid score=0) but its predicted reduced membrane-insertion potential indicates areas for future improvement. These findings demonstrate the potential and effectiveness of computational design in accelerating the development of peptide-based therapeutics. By enabling time- and cost-efficient AMP discovery, this approach may accelerate the development of viable antibiotic alternatives and help reduce the health burden posed by *K. pneumoniae* infections.

Keywords: Antimicrobial peptide, *Klebsiella pneumoniae*, antimicrobial activity, toxicity, computational design.

Introduction

Klebsiella pneumoniae is a Gram-negative, opportunistic pathogen that normally resides in the human gut but can cause severe infections when it spreads systemically. It poses a major threat in healthcare settings, particularly among vulnerable patients, and is one of the six highly drug-resistant ESKAPE pathogens—*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species—that are increasingly associated with antibiotic treatment failure¹. *K. pneumoniae* is responsible for a range of infections—including pneumonia, urinary tract infections, bacteremia, and liver abscesses—and was estimated to have caused 176,000 deaths globally from lower respiratory infections in 2021^{2,3}. A meta-analysis of nearly 38,000 patients reported 30-day and 90-day mortality rates of 29% and 34%, respectively, with higher risks in ICU and hospital-acquired cases⁴. Another meta-analysis across 14 countries found that 29% of *K. pneumoniae* infections involve carbapenem-resistant strains

(CRKP), with prevalence reaching 66% in low-income regions such as South Asia, compared to 14% in high-income areas like North America⁵. Misuse of antibiotics accelerates bacterial resistance, which spreads as bacteria evolve and share resistance genes via plasmids⁶.

Antimicrobial peptides (AMPs) offer a promising alternative due to their broad efficacy, reduced resistance risk, and immune-enhancing properties⁷. AMPs disrupt bacterial membranes and/or interfere with intracellular processes, causing cell rupture or increased permeability that enhances antibiotic effectiveness and making resistance development less likely⁸.

Recent research shows that certain synthetic antimicrobial peptides have potential for treating *K. pneumoniae* by reducing the abundance of proteins involved in DNA and protein metabolism, cytoskeleton and cell wall organization, redox metabolism, regulatory factors, ribosomal proteins, and antibiotic resistance—suggesting that AMPs may act through multiple mechanisms, help mitigate resistance development, and offer a potent strategy for new drug development against *K. pneumoniae* infections⁹. For example, antimicrobial peptide A20L has shown potential against *K.*

¹ Thomas Jefferson High School for Science and Technology, Alexandria, VA

pneumoniae by demonstrating in vitro and in vivo antibacterial and antibiofilm activity, membrane-disruptive mechanisms, low toxicity, and good stability under certain conditions¹⁰. Another example, FK5P—an ultra-short palmitoylated 5-mer phenylalanine:lysine random peptide mixture—effectively killed carbapenem-resistant *K. pneumoniae*, disrupted biofilms, reduced bacterial burden in mice, and exhibited strong safety and pharmacokinetic profiles, highlighting its promise as a therapeutic agent¹¹.

This study computationally investigates the design of novel antimicrobial peptides targeting *K. pneumoniae* using AI tools to optimize peptide structure. We hypothesize that amino acid length shortening and substitutions can increase predicted AMP probability and reduce predicted toxicity. Specifically, through two strategies — (1) shortening amino acid length, and (2) substituting polar uncharged or non-polar neutral residues and the structurally flexible glycine with positively charged or hydrophobic amino acids while maintaining a balanced hydrophobic/hydrophilic profile — we aim to computationally enhance antimicrobial efficacy by reducing hemolytic activity, improving selectivity, maintaining hydrophobic/hydrophilic balance, and increasing stability. Importantly, the results are based on computational predictions, and changes in biological efficacy or toxicity should not be implied without experimental validation.

This paper is structured as follows: Section II outlines the methods used to apply AI tools for amino acid length shortening and substitution in peptide design computationally; Section III presents the results showing improved predicted AMP probability and reduced predicted toxicity; Section IV discusses the significant implications and broader applicability of the research; and Section V concludes. While the findings are based on computational simulations, they may help inform future efforts to develop safer and more effective AMPs against *K. pneumoniae*, and guide the design of experimental validation studies.

Literature Review

Rising drug discovery costs and the need for new antibiotics have increased interest in antimicrobial peptides, with computational methods complementing experimental screening and enabling exploration of novel chemical spaces¹². Existing research shows that AI is accelerating antimicrobial peptide discovery by enabling novel designs and supporting preclinical validation¹³. Meanwhile, computer-aided drug design (CADD) is limited by model assumptions, incomplete data, high computational costs, and inability to fully predict pharmacokinetics or toxicity, requiring experimental validation¹⁴. Future efforts aimed at improving predictive model accuracy, mitigating biases in AI algorithms, and integrating diverse biological data could enhance CADD's effectiveness and its con-

tribution to antimicrobial peptide development¹⁵.

AI-designed antimicrobial peptides show potential as effective and safe alternatives to traditional antibiotics, with peptide modifications enhancing antimicrobial activity and yielding promising candidates for future therapeutic development. Large language model-based screening has been used to evaluate millions of peptides, identify promising candidates, and engineer more selective variants by increasing amphipathicity through targeted modifications, resulting in peptides with stronger antimicrobial activity but slightly reduced selectivity and demonstrating how AI can accelerate the discovery of new peptide antibiotics¹⁶. Similarly, AI-generated and optimized peptides have been shown to exhibit strong antimicrobial activity, with many displaying MICs below 10 μ M and some demonstrating anticancer activity, indicating that peptide modifications can enhance biological effectiveness while remaining safe¹⁷.

Among the many strategies for optimizing antimicrobial peptides, two approaches are shortening amino acid length and substituting specific amino acids to modify their properties. These approaches can improve peptide activity and safety. Shorter antimicrobial peptides often exhibit lower hemolytic activity and toxicity, and truncated versions have been developed that retain antimicrobial effectiveness. Some of these peptides show broad-spectrum antibacterial effects, while others are more selective against Gram-positive bacteria, and they generally remain safer and more cost-efficient than the original parent peptide¹⁸. Balancing hydrophobicity and charge distribution in cationic antimicrobial peptides can enhance their membrane-disrupting activity, providing a framework for designing effective therapeutic agents¹⁹. Using a model α -helical antimicrobial peptide as a framework, single amino acid substitutions were introduced to examine their effects on charge, hydrophobicity, and helicity, revealing that changes in hydrophobicity influenced biological activity without altering secondary structure and demonstrating a rational design strategy for enhancing antimicrobial peptide activity²⁰.

Computational approaches are increasingly being used to aid the discovery and design of antimicrobial peptides (AMPs) against drug-resistant bacteria such as *K. pneumoniae*. Machine learning screening, molecular docking, and molecular dynamics simulations were combined to identify potential antibiofilm AMPs, with several candidates showing stronger binding to the biofilm regulator MrkH than its native ligand, demonstrating the power of ML and simulation in discovering novel peptides²¹. Deep learning-based peptide language models were also used to design and test broad-spectrum AMPs, identifying peptides such as T2-9 that are highly effective against drug-resistant *K. pneumoniae*, *S. aureus*, and *P. aeruginosa*, with some also reducing resistance and treating infections in a mouse wound model²². Similarly, the Controlled Latent Attribute Space Sampling (CLaSS) computa-

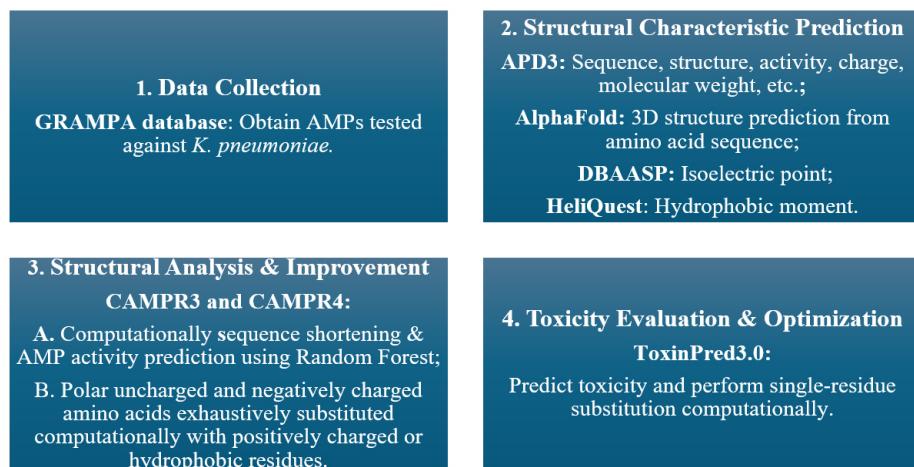


Fig. 1 Flowchart for computationally designing a high predicted-AMP-probability, low predicted-toxicity peptide.

tional framework was used to design non-toxic AMPs, leading to the identification of novel peptides with broad-spectrum potency, activity against drug-resistant *K. pneumoniae*, low toxicity, and resistance-mitigating properties²³.

Methods

AMPs are small proteins composed of amino acid chains—the fundamental building blocks of all proteins—and are considered promising alternatives to antibiotics due to their ability to disrupt bacterial membranes in ways that reduce the likelihood of resistance. Their activity is shaped by the chemical properties of the 20 standard amino acids, which vary in charge, polarity, and hydrophobicity. Among them, arginine (R), lysine (K), and histidine (H) are positively charged; aspartic acid (D) and glutamic acid (E) are negatively charged. The polar but uncharged amino acids include serine (S), threonine (T), asparagine (N), glutamine (Q), tyrosine (Y), and cysteine (C). The remaining nonpolar, hydrophobic residues—important for membrane interaction—are glycine (G), alanine (A), valine (V), leucine (L), isoleucine (I), methionine (M), phenylalanine (F), proline (P), and tryptophan (W). The composition and distribution of these residues determine an AMP’s antimicrobial activity, selectivity, and toxicity. Key features such as net charge, hydrophobicity, and the ability to form amphipathic structures are crucial for enabling AMPs to target bacterial membranes while sparing host cells⁸.

In this study, AMPs active against *K. pneumoniae* were retrieved from the Giant Repository of AMP Activities (GRAMPA) database and analyzed using various AI tools to evaluate their sequence, structure, activity, and toxicity, guiding the computational design of a new peptide with improved efficacy (Figure 1). GRAMPA is a database of antimicrobial

peptides that provides their amino acid sequences and corresponding minimum inhibitory concentrations (MICs) against various bacterial strains, including target species and strain-level details. From GRAMPA database, peptides tested against the bacterium *K. pneumoniae* were examined to identify patterns in activity and to guide the computational design of new and improved AMPs. The Antimicrobial Peptide Database (APD3), a searchable database of innate immune peptides, was used to gather detailed information—including sequence, structure, activity, source, and physicochemical properties like total net charge and molecular weight computationally. AlphaFold, an AI system developed by DeepMind, was used to predict the three-dimensional (3D) structure of proteins based on input amino acid sequences. The Database of Antimicrobial Activity and Structure of Peptides (DBAASP), a manually curated resource developed to support the computational design of antimicrobial compounds with a high therapeutic index, was used to obtain the isoelectric point of peptides. HeliQuest, a bioinformatics tool for analyzing the amphipathic and structural properties of α -helical peptides, was used to calculate the predicted hydrophobic moment and generate helix wheels²⁴. Collection of Anti-Microbial Peptides Release 3 (CAMPR3), a database of sequences, structures, and signatures of antimicrobial peptides, was used to computationally shorten the peptide sequence length applying the Random Forest algorithm; and CAMPR4, a subsequent release which expands and accelerates AMP based studies by providing curated information on natural and synthetic AMPs, was used to predict the novel peptide’s activity also applying the Random Forest algorithm. ToxinPred3.0, a web server that predicts peptide toxicity based on primary sequence using machine learning models, was used to obtain predicted toxicity scores and improve selectivity by replacing one amino

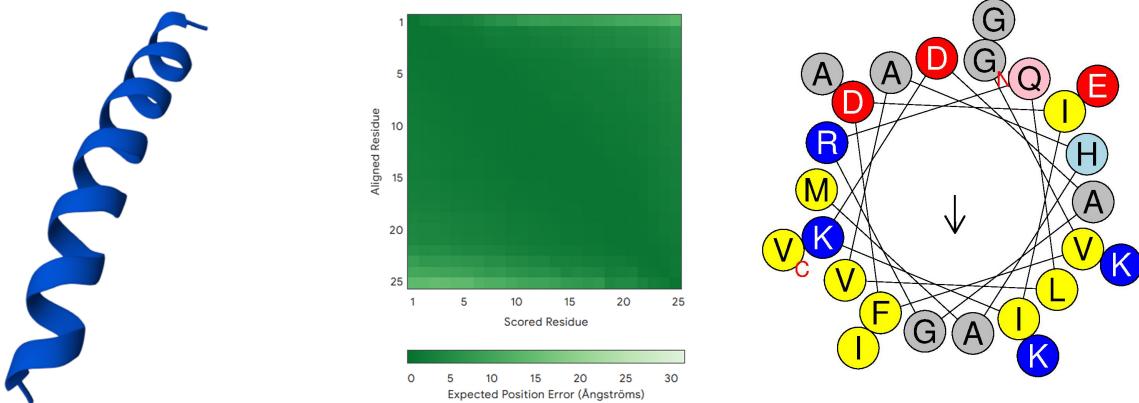


Fig. 2 The 3D structure and helix wheel of the initial peptide

Note: For the 3D structure, dark blue stands for very high ($p\text{LDDT} > 90$); light blue stands for confident ($90 > p\text{LDDT} > 70$); yellow stands for low ($70 > p\text{LDDT} > 50$); and orange stands for very low ($p\text{LDDT} < 50$).

Table 1 Predicted characteristics of the initial peptide “GVFDIIKDAGRQLVAHAMGKIAEKV”

Strain	MIC ($\mu\text{g/mL}$)	Net Charge	Molecular Weight (Da)	Isoelectric point	Type of Structure	Hydrophobic moment	AMP Probability	Toxicity Hybrid Score
ATCC 700603	300	1	2667.166	9.81	α -helical	0.268	0.531	0.5

Table 2 Prediction Results for Antimicrobial Regions in the Shortened Peptide Sequence

Position	Sequence	Class	AMP Probability
1–20	GVFDIIKDAGRQLVAHAMGK	AMP	0.676
2–21	VFDIIKDAGRQLVAHAMGKI	AMP	0.689
3–22	FDIIKDAGRQLVAHAMGKIA	AMP	0.691
4–23	DIIKDAGRQLVAHAMGKIAE	AMP	0.547
5–24	IIKDAGRQLVAHAMGKIAEK	AMP	0.859
6–25	IKDAGRQLVAHAMGKIAEKV	AMP	0.811

acid through targeted substitution computationally. In Toxin-Pred3.0, toxicity scores range from 0 to 1, with higher values indicating greater toxicity; this study used a 0.38 cutoff, and scores are probabilistic and require experimental validation. AlphaFold is a web service powered by AlphaFold 3 that generates highly accurate 3D structure predictions for proteins, DNA, RNA, ligands, ions, and chemical modifications in a single platform (<https://alphafoldserver.com/>).

A dataset of peptide reported against the bacterium *K. pneumoniae* and meeting the CAMPR3 sequence length requirement was used for this study. A peptide active against *K. pneumoniae* but with a relatively weak predicted antimicrobial performances, characterized by a high predicted MIC and

a relatively low predicted AMP score, was selected for improvement. Its existing activity against *K. pneumoniae*, combined with suboptimal predicted efficacy, made it a suitable candidate for optimization. Three steps were taken to increase its predicted antimicrobial potential and reduce predicted toxicity. First, the peptide sequence was shortened from 25 to 20 amino acids computationally. Second, polar uncharged and negatively charged amino acids were computationally substituted with positively charged or hydrophobic residues to enhance predicted antimicrobial activity and AMP scores. Third, predicted peptide toxicity was further minimized through targeted substitutions computationally.

Table 3 The top 10 mutant peptide variants ranked by predicted toxicity

Mutant ID	Sequence	Class	Predicted AMP Probability	Predicted MERCI (+ve) score	Predicted MERCI (-ve) score	Toxicity prediction
Mutant_K6D	IICKADRHLVAHAMKKIAKK	AMP	0.91	0	-0.5	0
Mutant_I17E	IICKAKRHLVAHAMKKEAKK	AMP	0.91	0	-0.5	0
Mutant_H12D	IICKAKRHLVADAMKKIAKK	AMP	0.92	0	-0.5	0
Mutant_I2D	IDKKAKRHLVAHAMKKIAKK	AMP	0.93	0	-0.5	0
Mutant_M14D	IICKAKRHLVAHADKKIAKK	AMP	0.97	0	-0.5	0
Mutant_I2Y	IYKKAKRHLVAHAMKKIAKK	AMP	0.99	0	-0.5	0
Mutant_K3Y	IIYKAKRHLVAHAMKKIAKK	AMP	0.99	0	-0.5	0
Mutant_H12Y	IICKAKRHLVAYAMKKIAKK	AMP	0.99	0	-0.5	0
Mutant_K19Y	IICKAKRHLVAHAMKKIAYK	AMP	0.99	0	-0.5	0
Mutant_M14Y	IICKAKRHLVAHAYKKIAKK	AMP	1	0	-0.5	0

A 20-residue peptide was chosen because many AMPs are around this length and shorter peptides are generally easier and cheaper to synthesize. Other lengths were not explored in this project which is a limitation and could be examined in future work. Amino acid substitutions were conducted because of three reasons: first, introducing positively charged residues (K, R, H) can enhance binding to negatively charged bacterial membranes; second, increasing hydrophobic residues can help the peptide insert into and disrupt membranes; and third, replacing very flexible residues like glycine with residues that favor helices can help stabilize an alpha-helical structure.

Results

From the GRAMPA database, 200 AMPs targeting *K. pneumoniae* were obtained. The AMP probability of each peptide was predicted computationally using CAMPR3 with the Random Forest classifier. Four peptides were excluded from the dataset due to sequence lengths shorter than four amino acids as indicated by CAMPR3, leaving 196 peptides for further analysis.

K. pneumoniae K6 (ATCC 700603, “GVFDIIKDA-GRQLVAHAMGKIAEKV”) is a clinical isolate resistant to oxyimino-beta-lactams, used as a quality control strain for detecting extended-spectrum beta-lactamases (ESBLs) due to clavulanic acid-mediated MIC reduction²⁵. The sequence exhibited one of the highest predicted MIC values (300 µg/mL) and one of the lowest predicted AMP probability scores (0.531), and was selected for mutational optimization. Table 1 summarizes its key characteristics, including a low predicted antimicrobial probability, high predicted MIC, low predicted net positive charge (+1), and high predicted toxicity (0.5).

Figure 2 shows the predicted 3D structure and helix wheel of the initial peptide, as generated by AlphaFold and HeliQuest, respectively (default settings applied).

The initial peptide sequence (“GVFDIIKDA-GRQLVAHAMGKIAEKV”), consisting of 25 residues, was computationally shortened to 20 residues using CAMPR3 2 with the Random Forest algorithm. Table 2 shows the results from the ‘Predict Antimicrobial Region within Peptides’ tool. The shortened sequence ‘IICKDAGRQLVAHAMGKIAEK’ (positions 5–24) showed the highest AMP probability and was selected for further mutation and optimization.

In the shortened sequence “IICKDAGRQLVAHAMGKIAEK”, amino acids with polar uncharged side chains (Q) and the negatively charged amino acids (D and E) and the structurally flexible glycine (G) were computationally substituted with amino acids bearing positively charged side chains (R; H; and K) to enhance the predicted AMP probability. An exhaustive set of 243 combinations (3^5 , corresponding to substitutions at five positions 4-D, 6-G, 8-Q, 15-G, and 19-E) was generated. The sequence ‘IICKAKRHLVAHAMKKIAKK’, achieved the highest predicted AMP probability score (0.968) according to CAMPR3.

Although the newly created sequence ‘IICKAKRHLVAHAMKKIAKK’ was predicted to be non-toxic (toxicity score: 0.155) by ToxinPred3.0, the tool was further utilized to guide single amino acid substitutions aimed at reducing predicted toxicity and optimizing the peptide’s predicted antimicrobial profile—potent against the target pathogen while remaining safe for host cells. A total of 381 mutants were generated computationally, among which ten were predicted to have zero toxicity and AMP probability above 0.90. Here, a toxicity score of 0 indicates a non-toxic prediction by the ToxinPred3.0

Table 4 Predicted characteristics of the final optimized peptide “IICKAKRHLVAHAYKKIAKK”

Net Charge	Molecular Weight (Da)	Type of Structure	Hydrophobic moment	AMP Probability	Toxicity Hybrid Score
+8.5	2344.94	α -helical	0.201	1	0

Table 5 Predicted characteristics of the LL-37 “LLGDFFRKSKKEKIGKEFKRIVQRIKDFLRNLVPRTES”

Strain	Net Charge	Molecular Weight (Da)	Type of Structure	Hydrophobic moment	AMP Probability	Toxicity Hybrid Score
LL-37	+6.0	4493.312	α -helical	0.521	0.749	0.0

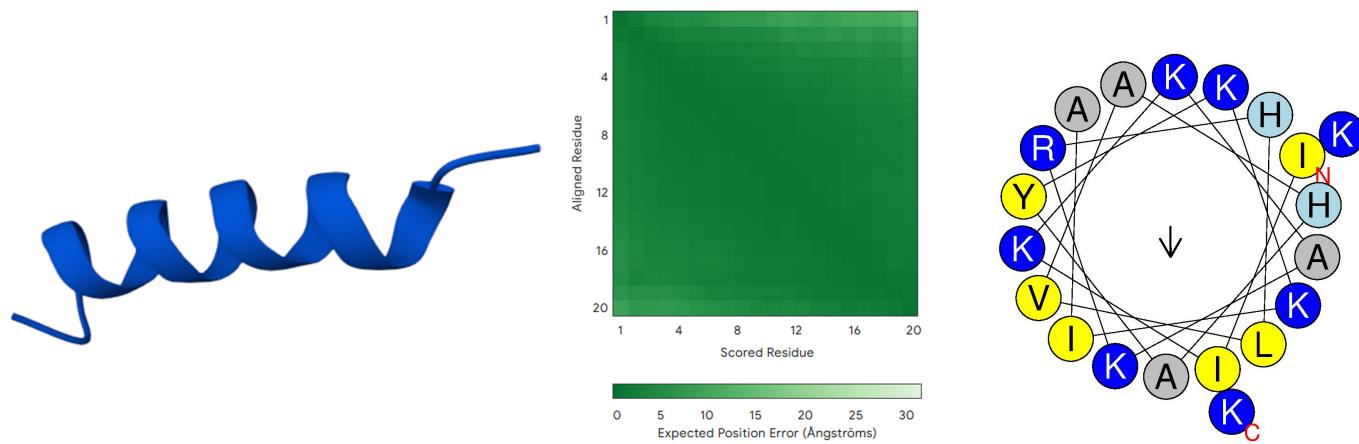


Fig. 3 The 3D structure and helix wheel of the optimized peptide

Note: For the 3D structure, dark blue stands for very high (pLDDT > 90); light blue stands for confident (90 > pLDDT > 70); yellow stands for low (70 > pLDDT > 50); and orange stands for very low (pLDDT < 50).

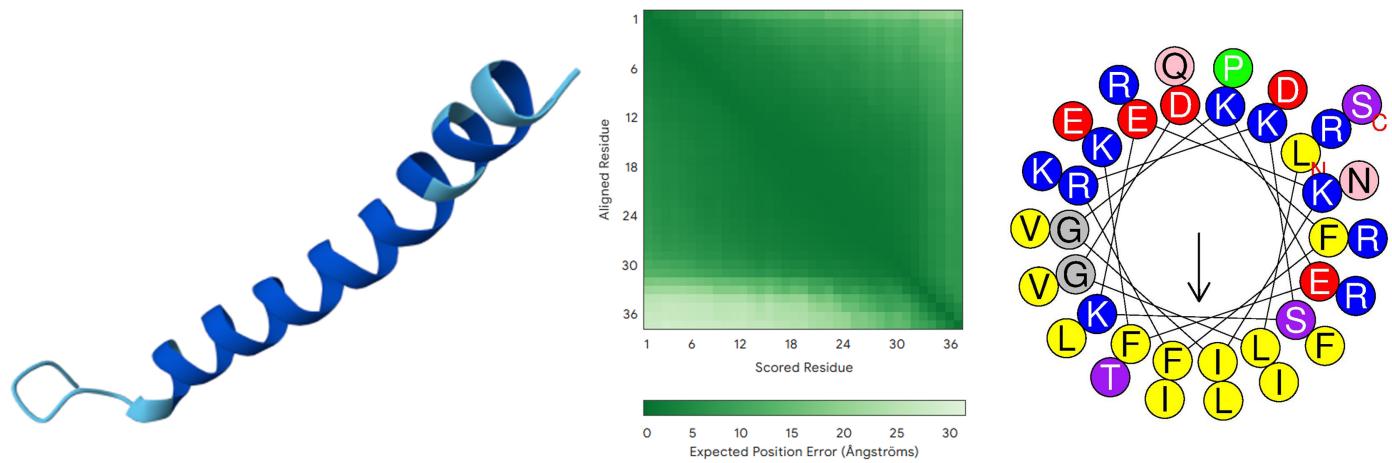


Fig. 4 The 3D structure and helix wheel of LL-37

Note: For the 3D structure, dark blue stands for very high (pLDDT > 90); light blue stands for confident (90 > pLDDT > 70); yellow stands for low (70 > pLDDT > 50); and orange stands for very low (pLDDT < 50).

model rather than experimentally confirmed absence of toxicity. The ten sequences with the lowest predicted toxicity scores (from ToxinPred3.0) and their corresponding predicted AMP probabilities (from CAMPR4) are presented in Table 3. The remaining 371 sequences were excluded due to high predicted toxicity or low AMP probability, and therefore not presented for brevity.

The sequence ‘IICKAKRHLVAHAYKKIAKK’ (last sequence in Table 3) combines high predicted AMP probability and minimal predicted toxicity. MERCI (+ve) and (−ve) scores are motif-based predictors of antimicrobial potential, with (+ve) indicating AMP-like motifs (0 to +1) and (−ve) indicating non-AMP motifs (−1 to +1); the peptide showed a (−ve) score of −0.5 and a (+ve) score of 0, suggesting minimal non-AMP motifs and moderate AMP-associated motifs, consistent with its predicted antimicrobial activity. Compared with the peptide before single amino acid substitutions (‘IICKAKRHLVAHAMKKIAKK’), Methionine (M) is replaced by Tyrosine (Y) at position 14. Table 4 summarizes the key characteristics of this optimized peptide, including its AMP probability score (1), predicted toxicity score (0), net positive charge (+8.5), and predicted α -helical structure. Figure 3 presents the predicted 3D structure and Helix wheel of this final candidate.

For reference, the optimized peptide was compared with the widely studied human cathelicidin antimicrobial peptide LL-37 (“LLGDFFRKSKEKIGKEFKRIVQRIKDFLRN-LVPRTES”). Table 5 summarizes the key characteristics of this optimized peptide, including its AMP probability score (0.749), predicted toxicity score (0), net positive charge (+6.0), and predicted α -helical structure. Figure 4 presents the predicted 3D structure and Helix wheel of this final candidate.

Discussion

AMP probability describes how similar the sequence is to known AMPs according to the model, while potency / effectiveness refers to how strongly a peptide kills bacteria in experiments (exp. MIC values). The higher predicted AMP probability suggests a greater likelihood of antimicrobial activity, but the true potency remains unknown without experimental MIC measurements. Because short peptides can be flexible in aqueous solution, the AlphaFold structure should be viewed as a possible model of how the peptide might look when interacting with a membrane, rather than a confirmed conformation in solution.

Predicted Local Distance Difference Test (pLDDT) estimates how well a predicted structure agrees with an experimental structure based on backbone carbon coordinates²⁶. It provides a per-atom confidence score from 0–100, where higher values indicate greater confidence: scores

above 90 are highly reliable, while scores below 50 suggest likely inaccuracies (AlphaFold Server, 2025, (<https://alphafoldserver.com/guides#section-3:interpreting-results-from-alphafold-server>)). Predicted Aligned Error (PAE) measures AlphaFold 3’s confidence in the relative positions of two items within the predicted structure²⁷. Higher PAE values indicate lower confidence in the predicted spatial relationship between residues.

AlphaFold3 predicts the initial peptide (Figure 2) to adopt a stable α -helical conformation with high pLDDT scores (> 90; dark blue) and low predicted aligned error (mostly dark green), indicating a well-defined structure with modest terminal flexibility typical of α -helical antimicrobial peptides. The final optimized peptide (Figure 3) shows a similar α -helical fold with high pLDDT (mostly dark blue) and predominantly dark green PAE, reflecting very high local confidence and modest terminal flexibility. Comparison with the AlphaFold3-predicted LL-37 structure (Figure 4) reveals high pLDDT across the α -helical core (dark blue) with slightly reduced confidence at both termini (light blue), and predominantly low PAE with modest terminal flexibility (mostly dark green, with lighter green regions in the lower and right parts).

Our findings suggest that computational amino acid length shortening and substitutions can enhance predicted AMP scores and reduce predicted toxicity. Shorter AMPs often retain strong antibacterial activity while reducing hemolysis and improving selectivity, stability, and resistance profiles²⁸. Mechanistically, their activity—driven by binding to negatively charged membranes and peptide aggregation—tends to plateau with increasing length, showing diminishing per-residue effectiveness in longer sequences²⁹. Substituting polar uncharged residues and glycine—a structurally flexible, non-polar special case—with positively charged or hydrophobic amino acids increases the predicted likelihood that a peptide will disrupt bacterial membranes or cross into the cytoplasm to interfere with intracellular processes, while limiting toxicity to host cells³⁰.

As an ESKAPE pathogen, *K. pneumoniae* poses a serious threat to hospitalized patients, especially those in intensive care units. Its multi-drug resistance contributes to high mortality and substantial healthcare costs. Developing safe and effective antimicrobial peptides represents a promising strategy to reduce deaths and protect vulnerable populations. A meta-analysis of 157 studies including 37,915 patients showed *K. pneumoniae* is associated with high, progressively increasing mortality: 17% at 7 days, 24% at 14 days, 29% at 30 days, and 34% at 90 days, with significantly higher rates in ICU patients⁴. As a leading cause of carbapenem-resistant Enterobacteriaceae (CRE) infections, *K. pneumoniae* contributes to major healthcare expenditures, with direct hospital costs per CRE case ranging from \$22,000 to \$66,000 and total annual

U.S. hospital costs estimated between \$275 million and \$1.4 billion, alongside significant societal costs and loss of quality-adjusted life years³¹. Given the rapid rise in drug resistance, there is an urgent need to invest in the development of novel antimicrobials and alternative therapeutic strategies to effectively manage and contain *K. pneumoniae* infections.

This work contributes to the broader effort to combat the growing threat of antibiotic resistance by applying a systematic and exhaustive computational approach to optimizing antimicrobial peptides. We computationally shortened the sequence and exhaustively tested 243 substitutions replacing polar uncharged, negatively charged residues, and glycine with positively charged amino acids (arginine, histidine, lysine), followed by targeted single-residue hydrophobic substitutions to reduce predicted toxicity. We first discarded variants with high predicted toxicity. Among the remaining sequences, we prioritized those with high predicted AMP probability, zero predicted toxicity, and suitable net charge and hydrophobic moment, and from these we selected one peptide as the final design.

This strategy computationally identified a novel peptide with greatly enhanced antimicrobial potential and reduced toxicity, achieving a predicted AMP probability of 1.0 and predicted toxicity of 0—improving from the original peptide’s predicted 0.531 and 0.5 scores, respectively. The predicted hydrophobic moment—a measure of the dipole moment created by the separation of charged and uncharged residues—decreased from 0.268 to 0.201, indicating membrane-inserting potential remaining within a relatively low to moderate range and the reduction suggesting a trade-off in peptide efficacy.

To place the novel design in a broader AMP context, the designed peptide was compared with LL-37, the only human cathelicidin. LL-37 is a 37-residue antimicrobial peptide produced by immune and epithelial cells that exerts broad antimicrobial activity, modulates immune responses, promotes tissue repair, and shows therapeutic potential in infection models³². The final optimized peptide showed a higher predicted AMP probability (1.0 vs. 0.749), equally low predicted toxicity (0), but a lower hydrophobic moment (0.201 vs. 0.521). The optimized peptide also had a higher predicted net charge (+8.5 vs. +6.0), suggesting stronger electrostatic interactions with negatively charged bacterial membranes. Both peptides had an α -helical structure.

The reduced hydrophobic moment may help decrease disruptive interactions with mammalian membranes, potentially lowering toxicity, but it could also weaken membrane-disrupting activity against bacteria. Future work could explore different hydrophobic moments to better balance predicted toxicity and antimicrobial effectiveness. The computationally generated helix wheels of the initial and final peptides exhibit some predicted amphipathic character that may support membrane activity, but their relatively weak predicted hydrophobic

moment compared with LL-37 may limit membrane selectivity and potency. In this context, selectivity means being more harmful to bacteria than to human cells. Experimentally, this is often quantified using a selectivity index, such as the ratio between a toxicity value (e.g., hemolysis or cell viability) and MIC. In this project, we did not measure this index, so selectivity is only predicted, not experimentally confirmed. Repositioning residues to enhance the hydrophobic–hydrophilic separation could improve amphipathicity promoting binding to the bilayer. These simulations were made possible by leveraging computational tools, enabling efficient and cost-effective design of novel antimicrobial peptides.

This study demonstrates that AI and computational approaches can effectively help design novel AMPs targeting *K. pneumoniae* by optimizing amino acid length and substitutions to enhance predicted antimicrobial activity while reducing predicted toxicity. Such computational sequence-optimization strategies are broadly applicable for developing potent AMPs against diverse bacterial pathogens beyond *K. pneumoniae*. AMP-based drug development is a long and costly process, typically taking 13–14 years and costing tens of millions of euros per candidate, with delays and attrition potentially adding €50–100 million to early-stage expenses³³. AI and computational methods hold great promises to substantially reduce both time and costs during the early discovery and lead optimization stages, offering a promising, cost-effective, and time-efficient alternative to traditional antibiotic development pathways.

This study, however, has several limitations:

- First, while CAMPR3/4 and ToxinPred are powerful tools, they are machine-learning models that use patterns in existing data to provide probabilistic predictions, not guarantees. Scores such as “AMP probability=1.0” or “toxicity=0” should be interpreted as “very likely” or “very unlikely” according to the model. The true behavior of the peptide must ultimately be confirmed by experiments.
- Second, this study did not specifically predict red blood cell hemolysis, which is a key concern for AMPs intended for systemic use. Future work should include *in vitro* hemolysis assays to access whether the peptide damages human red blood cells.
- Third, highly cationic peptides often increase hemolysis, reduced hydrophobic moment may reduce activity, strain-to-strain MIC variability.
- Fourth, while this study shows that computationally targeted single-site substitutions can improve predicted antimicrobial activity and reduce predicted toxicity, it is limited by focusing solely on shortening the peptide from

25 to 20 residues with an arbitrary cutoff, without exploring other lengths or sequence arrangements. Potentially valuable strategies such as multi-site substitutions and structural modifications to improve physicochemical properties and enhance the hydrophobic–hydrophilic separation were also not examined.

As all results are based solely on computational predictions, future research should build upon these findings by investigating a wider range of sequence alterations and performing wet lab experimental validation to confirm and further optimize the peptide design.

Molecular dynamics simulations incorporating an atomistic bilayer and surrounding solvent offer detailed insight into peptide–bilayer interactions³⁴. The energetics of AMP addition to a transmembrane pore, quantified through molecular dynamics simulations, are essential for evaluating pore formation and growth, providing critical insight into the pore-forming ability of peptides in cell membranes³⁵. However, conducting membrane simulations and energetics studies is computationally intensive and represents an important direction for future research.

Acknowledgements

I sincerely thank Dr. Christopher Lockhart, Dr. Elena Rivas, and Dr. Danielle Ross for their invaluable guidance throughout this research.

References

- 1 H. G. Braun, S. R. Perera, Y. D. N. Tremblay and J.-L. Thomassin, *Canadian Journal of Microbiology*, 2024, **70**, 507–528.
- 2 M. K. Paczosa and J. Mecsas, *Microbiology and Molecular Biology Reviews*, 2016, **80**, 629–661.
- 3 GBD 2021 Lower Respiratory Infections and Antimicrobial Resistance Collaborators, *The Lancet Infectious Diseases*, 2024.
- 4 D. Li, X. Huang, H. Rao, H. Yu, S. Long, Y. Li and J. Zhang, *Frontiers in Cellular and Infection Microbiology*, 2023, **13**, 1157010.
- 5 X. C. Lin, C. L. Li, S. Y. Zhang, X. F. Yang and M. Jiang, *Open Forum Infectious Diseases*, 2023, **11**, ofad649.
- 6 F. Svara and D. J. Rankin, *BMC Evolutionary Biology*, 2011, **11**, 130.
- 7 T. Islam, N. T. Tamanna, M. S. Sagor, R. M. Zaki, M. F. Rabbee and M. Lackner, *Pharmaceutics*, 2024, **16**, 1542.
- 8 P. Gagat, M. Ostrówka, A. Duda-Madej and P. Mackiewicz, *International Journal of Molecular Sciences*, 2024, **25**, 10821.
- 9 L. A. C. Branco, P. F. N. Souza, N. A. S. Neto, T. K. B. Aguiar, A. F. B. Silva, R. F. Carneiro, C. S. Nagano, F. P. Mesquita, L. B. Lima and C. D. T. Freitas, *Antibiotics*, 2022, **11**, 1753.
- 10 H. Zhou, X. Du, Y. Wang, J. Kong, X. Zhang, W. Wang, Y. Sun, C. Zhou, T. Zhou and J. Ye, *Microbiology Spectrum*, 2024, **12**, e03979–23.
- 11 J. Z. Lau, S. H. Kuo, Y. Belo, E. Malach, B. Maron, H. E. Caraway, M. W. Oh, Y. Zhang, N. Ismail, G. W. Lau and Z. Hayouka, *Antimicrobial Agents and Chemotherapy*, 2023, **67**, e00574–23.
- 12 M. Nedyalkova, A. S. Paluch, D. Potes Vecini and M. Lattuada, *Digital Discovery*, 2024, **3**, 9–22.
- 13 H. Meng, *Probiotics and Antimicrobial Proteins*, 2025.
- 14 S. Sur and H. Nimesh, *Advances in Pharmacology*, 2025, vol. 103, pp. 415–428.
- 15 D. P. Mihai and G. M. Nitulescu, *Pharmaceutics*, 2025, **18**, 436.
- 16 D. Bae, M. Kim, J. Seo and H. Nam, *Briefings in Bioinformatics*, 2025, **26**, bbaf343.
- 17 A. Mesa, A. Orrego, J. W. Branch-Bedoya *et al.*, *Current Microbiology*, 2025, **82**, 379.
- 18 J. Y. Kwon, M. K. Kim, L. Mereuta, C. H. Seo, T. Luchian and Y. Park, *AMB Express*, 2019, **9**, 122.
- 19 L. M. Yin, M. A. Edwards, J. Li, C. M. Yip and C. M. Deber, *Journal of Biological Chemistry*, 2012, **287**, 7738–7745.
- 20 J. Tan, J. Huang, Y. Huang and Y. Chen, *Molecules*, 2014, **19**, 10803–10817.
- 21 A. Al-Khdhairawi, D. Sanuri, R. Akbar, S. D. Lam, S. Sugumar, N. Ibrahim, S. Chieng and F. Sairi, *Computational Biology and Chemistry*, 2023, **102**, 107800.
- 22 T. Li, X. Ren, X. Luo *et al.*, *Nature Communications*, 2024, **15**, 7538.
- 23 P. Das, T. Sercu, K. Wadhawan, I. Padhi, S. Gehrmann, F. Cipcigan, V. Chenthamarakshan, H. Strobelt, C. dos Santos, P.-Y. Chen, Y. Y. Yang, J. Tan, J. Hedrick, J. Crain and A. Mojsilovic, *Nature Biomedical Engineering*, 2021.
- 24 M. Schiffer and A. B. Edmundson, *Biophysical Journal*, 1967, **7**, 121–135.
- 25 J. K. Rasheed, G. J. Anderson, H. Yigit, A. M. Queenan, A. Doménech-Sánchez, J. M. Swenson, J. W. Biddle, M. J. Ferraro, G. A. Jacoby and F. C. Tenover, *Antimicrobial Agents and Chemotherapy*, 2000, **44**, 2382–2388.
- 26 E. F. McDonald, T. Jones, L. Plate, J. Meiler and A. Gulsevin, *Structure*, 2023, **31**, 111–119.e2.
- 27 P. G. Magana Gomez and O. Kovalevskiy, *AlphaFold: a practical guide*, EMBL-EBI, Hinxton, Cambridgeshire, United Kingdom, 2024, <https://doi.org/10.6019/TOL.AlphaFold-w.2024.00001.1>.
- 28 A. Barreto-Santamaría, Z. J. Rivera, J. E. García, H. Curtidor, M. E. Patarroyo, M. A. Patarroyo and G. Arévalo-Pinzón, *Microorganisms*, 2020, **8**, 867.
- 29 S. Clark, T. A. Jowitt, L. K. Harris, C. G. Knight and C. B. Dobson, *Communications Biology*, 2021, **4**, 605.
- 30 R. M. Fleeman, L. A. Macias, J. S. Brodbelt and B. W. Davies, *Proceedings of the National Academy of Sciences of the United States of America*, 2020, **117**, 27620–27626.
- 31 S. M. Bartsch, J. A. McKinnell, L. E. Mueller, L. G. Miller, S. K. Gohil, S. S. Huang and B. Y. Lee, *Clinical Microbiology and Infection*, 2017, **23**, 48.e9–48.e16.
- 32 O. E. Voronko, V. A. Khotina, D. A. Kashirskikh, A. A. Lee and V. A. Gasanov, *International Journal of Molecular Sciences*, 2025, **26**, 8103.
- 33 L. Cresti, G. Cappello and A. Pini, *International Journal of Molecular Sciences*, 2024, **25**, 4870.
- 34 P. La Rocca, P. C. Biggin, D. P. Tielemans and M. S. P. Sansom, *Biochimica et Biophysica Acta*, 1999, **1462**, 185–200.
- 35 Y. Lyu, N. Xiang, X. Zhu and G. Narsimhan, *The Journal of Chemical Physics*, 2017, **146**, 155101.