Technical Note

Exploratory AI-Assisted In Silico Evidence That Bacterial Signaling Molecules May Occupy a Drug-Like Pharmacokinetic Space: A Preliminary Comparative Cheminformatics Analysis and Candidate Identification

Maxwel Adriano Abegg *

Posted Date: 11 February 2026

doi: 10.20944/preprints202602.0854.v1

Keywords: antimicrobial resistance; drug-likeness; bacterial signaling; quorum sensing; Lipinski rules; in silico screening; antibiotic discovery; signaling-first triage

 doi:10.20944/preprints202602.0854.v1

Technical Note

Exploratory AI-Assisted in Silico Evidence That Bacterial Signaling Molecules May Occupy a Drug-Like Pharmacokinetic Space: A Preliminary Comparative Cheminformatics Analysis and Candidate Identification

Maxwel Adriano Abegg

Institute of Exact Sciences and Technology; Graduate Program in Sciences, Technology and Health (PPGCTS), Federal University of Amazonas (UFAM), Itacoatiara, Amazonas, Brazil; maxabegg@gmail.com

Abstract

The hypothesis that antibiotics evolved from or co-opted bacterial intercellular signaling molecules predicts that signaling-active compounds should exhibit physicochemical properties favorable for drug development. We explored this prediction through a comparative cheminformatics analysis of 74 bacterial signaling molecules — comprising 48 small-molecule diffusible signals, 17 peptide autoinducers and intracellular messengers, and 9 antibiotics with documented signaling activity — against 71 randomly selected drug-like molecules. Small-molecule signaling molecules showed significantly higher Lipinski Rule-of-5 compliance than random drugs (95.8% vs. 73.2%; Odds Ratio = 8.40, 95% CI: 1.86–38.1, p = 0.001; Bonferroni-adjusted p = 0.009), while peptide autoinducers and intracellular messengers showed near-zero compliance (5.9%), confirming that the enrichment is specific to small-molecule diffusible signals. Predicted membrane permeability was over 2-fold higher for small-molecule signals (72.9% vs. 32.4% classified as "Good"), and enrichment was observed across six of seven drug-likeness filters (Enrichment Factors 1.12–1.31). We identified 10 signaling-active molecules with favorable profiles whose direct antibiotic activity has not been systematically tested; six are free of PAINS alerts. This analysis was conducted with substantial AI assistance, molecular properties were compiled from databases rather than computed de novo, and results require independent verification with validated cheminformatics tools. We present these findings as a hypothesis-generating study to stimulate experimental testing of the signaling-first triage concept for antibiotic discovery. For all abbreviations and acronyms used throughout this manuscript, please refer to the Glossary section.

Keywords: antimicrobial resistance; drug-likeness; bacterial signaling; quorum sensing; Lipinski rules; in silico screening; antibiotic discovery; signaling-first triage

1. Introduction

1.1. The Antibiotic Discovery Crisis

Antimicrobial resistance (AMR) represents a growing global health threat while discovery pipelines remain optimized for a single endpoint: minimum inhibitory concentration (MIC) in standardized growth assays (Brown & Wright, 2016). This framing may systematically discard molecules whose primary bioactivity manifests at sub-inhibitory concentrations (Andersson & Hughes, 2014; Fajardo & Martínez, 2008).

1.2. The Signaling-First Hypothesis

 doi:10.20944/preprints202602.0854.v1

A companion opinion article (Abegg, 2025, in preparation) proposes that molecules capable of modulating bacterial cell signaling at sub-inhibitory concentrations represent a systematically underexplored source for antibiotic discovery. This "signaling-first" triage framework rests on the empirical observation that many antibiotics act as intercellular signals at low concentrations (Linares et al., 2006; Davies & Davies, 2010; Romero et al., 2011), and that concentration-dependent transitions from signaling to growth inhibition have been documented for multiple compound classes (Beyersmann et al., 2017; Hoffman et al., 2005). Recent studies have strengthened this basis: Williams et al. (2024) showed that sub-inhibitory antibiotics elicit structured signaling responses in *Pseudomonas aeruginosa*, while Amer et al. (2025) demonstrated growth-phase-dependent modulation of quorum sensing (QS) by sub-MIC antibiotics.

1.3. An Untested Prediction

An implicit prediction of this hypothesis is that molecules with demonstrated bacterial signaling activity should exhibit favorable drug-likeness properties. The evolutionary logic is that the physicochemical constraints enabling effective intercellular signaling — small molecular size for passive diffusion, moderate lipophilicity for membrane transit — overlap with the properties that predict drug-likeness (Lipinski et al., 2001; Veber et al., 2002). Critically, this prediction applies specifically to small-molecule diffusible signals. Gram-positive bacteria frequently employ peptide-based quorum sensing systems, and intracellular second messengers operate under entirely different physicochemical logic. Including these non-drug-like signaling molecules in the analysis provides an essential internal control.

1.4. Objectives

This exploratory technical note addresses: **Part 1:** Do small-molecule bacterial signaling compounds exhibit systematically different drug-likeness profiles compared to both peptide/intracellular signals and randomly selected drug-like molecules? **Part 2:** Can specific signaling-active molecules with favorable profiles be identified that lack systematic antibiotic testing at elevated concentrations?

2. Methods

2.1. AI-assisted Workflow

This analysis was conducted with AI assistance using Claude (Anthropic; Claude Opus 4.6 Extended for all analysis, dataset construction, statistical verification, and manuscript revision, December 2024 – February 2025). The specific workflow was:

Step 1 — Dataset construction. The author provided Claude with the conceptual framework, inclusion/exclusion criteria, and iterative corrections over multiple sessions. Claude was prompted to identify bacterial signaling molecules with documented, reproducible signaling activity from the quorum sensing and bacterial communication literature, including: (a) all major classes of Gram-negative autoinducers (AHLs with varying acyl chain lengths and substitutions, quinolone signals, DSF family, AI-2); (b) Gram-positive peptide autoinducers (AIPs, CSPs, competence pheromones, lantibiotics); (c) intracellular second messengers (c-di-GMP, c-di-AMP, ppGpp, pppGpp, cAMP); (d) *Streptomyces* butyrolactone signals (A-factor, SCB1, VB-A); (e) interspecies signals (indole derivatives, CAI-1, photopyrones); (f) QS inhibitors and antivirulence compounds with documented signaling modulation; and (g) antibiotics with documented sub-MIC signaling effects. A comparison set of 71 drug-like molecules spanning diverse therapeutic areas was independently constructed. Selection was based on Claude's training knowledge of published literature, not a systematic database query with reproducible search terms — this is a key methodological limitation.

Step 2 — Property compilation. For each molecule, Claude retrieved MW, LogP, HBA, HBD, TPSA, and rotatable bonds from PubChem compound records (preferred source) or DrugBank. LogP

 doi:10.20944/preprints202602.0854.v1

values correspond to XLogP3 as reported in PubChem unless otherwise noted. For Gram-positive peptide autoinducers (MW > 500 Da), properties were estimated from peptide sequence-based calculations and published structural data, as PubChem records are incomplete or absent for many of these molecules. Properties were compiled from database records, not computed de novo from canonical SMILES using a single cheminformatics toolkit (e.g., RDKit, CDK, OpenBabel). This means that (a) values for the same property may derive from different computational algorithms across molecules, (b) results cannot be precisely reproduced by re-running a computational pipeline, and (c) inter-source inconsistencies may exist. **This is the most critical methodological weakness of this study.**

Step 3 — Filter application and statistical analysis. Claude applied seven drug-likeness filters to all datasets using the published threshold criteria for each filter (Section 2.3). Statistical tests were performed using Python (scipy.stats v1.11): Mann-Whitney U tests with rank-biserial correlation for continuous property comparisons, Fisher's exact test for categorical pass/fail comparisons, and Enrichment Factors (EF) calculated as EF = (% passing in test set) ÷ (% passing in reference set). Bonferroni correction was applied separately for property comparisons (6 tests, adjusted $\alpha = 0.0083$) and filter comparisons (7 tests, adjusted $\alpha = 0.0071$). The 95% confidence interval (CI) for the Lipinski Odds Ratio (OR) was computed using the Woolf logit method. All scripts were generated by Claude and executed in the same session.

Step 4 — ADMET prediction and PAINS screening. Simplified ADMET predictions were derived from threshold-based classifications (Section 2.4). All 26 candidate molecules were screened for PAINS alerts via manual structural assessment guided by AI (Section 2.5).

Step 5 — Testing gap assessment. For each candidate molecule, Claude assessed whether published MIC or dose-response data at growth-inhibitory concentrations could be identified. This was based on Claude's training knowledge of published literature (training data cutoff: approximately April 2024), not a formal systematic review with defined search protocols. Consequently, unpublished negative results ("file drawer problem") and literature published after the training cutoff are not captured.

Step 6 — Figure generation and manuscript drafting. Python scripts (matplotlib v3.8, scipy v1.11) were generated by Claude and executed to produce all figures. The author critically reviewed all outputs, challenged assumptions, identified errors, and directed iterative refinement across multiple revision cycles. All Python scripts and raw data are provided as supplementary materials.

2.2. Dataset Construction

Signaling molecule dataset (n = 74 total; three subsets).

Small-molecule diffusible signals (n = 48, "Primary" subset): AHL autoinducers including unsubstituted (C4- through C14-HSL), 3-oxo-substituted (3-oxo-C6- through 3-oxo-C12-HSL), and 3-hydroxy-substituted (3-OH-C4-HSL, 3-OH-C14:1-HSL) variants (n = 12); quinolone signals (PQS, HHQ, HQNO, NHQ; n = 4); universal/interspecies signals including AI-2 (DPD) and indole derivatives (indole, 5-hydroxyindole, isatin, indole-3-carboxaldehyde, indole-3-acetic acid; n = 6); DSF family (DSF, BDSF, CDSF; n = 3); *Streptomyces* butyrolactone signals (A-factor, SCB1, virginiae butanolide A; n = 3); *Vibrio* CAI-1 (n = 1); *Photorhabdus* photopyrone A (n = 1); dual signal/antibiotic tropodithietic acid (TDA; n = 1); bacterial signal metabolites (pyocyanin, phenazine-1-carboxylic acid, 2-aminoacetophenone, tyramine, hordenine; n = 5); QS inhibitors with documented signaling modulation (furanone C-30, 5-fluorouracil, cinnamaldehyde, eugenol, ajoene, curcumin, baicalin; n = 7); and antivirulence signal modulators (savirin/M64, mBTL, virstatin, LED209, LsrK inhibitor 11e; n = 5).

Peptide autoinducers and intracellular messengers (n = 17, "Negative" subset): Gram-positive peptide autoinducers and lantibiotics (AIP-I, AIP-II, AIP-III, AIP-IV from *S. aureus*; CSP-1 and CSP-2 from *S. pneumoniae*; GBAP from *E. faecalis*; ComX from *B. subtilis*; SHP from *Streptococcus*; iAM373 and cCF10 from *E. faecalis*; nisin A; n = 12) and intracellular second messengers (cyclic-di-GMP, cyclic-di-AMP, cAMP, ppGpp, pppGpp; n = 5). These serve as an internal negative control — they are bona

 doi:10.20944/preprints202602.0854.v1

fide signaling molecules but are predicted to fail drug-likeness filters due to fundamentally different physicochemical constraints (large peptide size for receptor specificity; high charge density for intracellular enzyme recognition).

Antibiotics with sub-MIC signaling effects (n = 9, "Excluded" subset): Azithromycin, tobramycin, ciprofloxacin, tetracycline, erythromycin, trimethoprim, rifampicin, colistin, and daptomycin. These are excluded from the primary enrichment analysis to avoid circular reasoning (antibiotics are already known drugs, so including them would artificially inflate drug-likeness metrics).

The complete dataset with all 74 molecules and their physicochemical properties is in Supplementary Table S1.

Random drug-like molecule dataset (n = 71). Molecules selected to span diverse therapeutic areas: cardiovascular (amlodipine, atorvastatin, clopidogrel, digoxin, lisinopril, losartan, metoprolol, warfarin; n = 8), CNS/psychiatric (carbamazepine, diazepam, donepezil, fluoxetine, gabapentin, levetiracetam, risperidone, sertraline; n = 8), oncology (cyclophosphamide, doxorubicin, erlotinib, imatinib, methotrexate, paclitaxel, sorafenib, tamoxifen; n = 8), metabolic/endocrine (glipizide, levothyroxine, metformin, pioglitazone, sitagliptin; n = 5), anti-inflammatory (celecoxib, colchicine, ibuprofen, indomethacin, naproxen, prednisone; n = 6), antiviral (acyclovir, lopinavir, oseltamivir, remdesivir, ribavirin, tenofovir; n = 6), antifungal (fluconazole, griseofulvin, ketoconazole, terbinafine; n = 4), miscellaneous FDA-approved drugs (allopurinol, chloroquine, finasteride, hydroxychloroquine, montelukast, omeprazole, ranitidine, sildenafil; n = 8), failed/withdrawn (BMS-986165/deucravacitinib, lorcaserin, rofecoxib, torcetrapib; n = 4), large/complex (cyclosporine, ivermectin, sirolimus, vincristine; n = 4), and additional diverse drugs (apixaban, aspirin, caffeine, canagliflozin, cetirizine, diphenhydramine, empagliflozin, loratadine, pantoprazole, rosuvastatin; n = 10). Complete dataset in Supplementary Table S2.

Candidate molecule dataset (n = 26). Subset curated for Part 2: all molecules from the Primary subset with documented signaling activity and without established clinical antibiotic use. Complete with testing gap annotations, PAINS flags, and priority classifications in Supplementary Table S3.

2.3. Drug-Likeness Filters

Seven independent filters were applied using published threshold criteria:

1. Lipinski Rule-of-5 — strict (0 violations: MW ≤ 500, LogP ≤ 5, HBA ≤ 10, HBD ≤ 5) (Lipinski et al., 2001)

2. Lipinski Rule-of-5 — relaxed (≤1 violation allowed) (Lipinski et al., 2001)

3. Veber Criteria — Rotatable bonds ≤ 10 AND TPSA ≤ 140 Å² (Veber et al., 2002)

4. O'Shea & Moser Antibiotic-Specific — MW < 600, −2 < LogP < 6, HBD ≤ 5, HBA ≤ 12 (O'Shea & Moser, 2008)

5. Ghose Filter — 160 ≤ MW ≤ 480, −0.4 ≤ LogP ≤ 5.6, 40 ≤ atoms ≤ 130, 20 ≤ MR ≤ 130 (Ghose et al., 1999). Note: atom count and molar refractivity (MR) were not available in the compiled dataset; Ghose pass rates are based on MW and LogP criteria only and may therefore be overestimated.

6. Egan Filter — LogP ≤ 5.88, TPSA ≤ 131.6 Å² (Egan et al., 2000)

7. Muegge Filter — 200 ≤ MW ≤ 600, −2 ≤ LogP ≤ 5, TPSA ≤ 150, HBA ≤ 10, HBD ≤ 5, RotBonds ≤ 15, rings ≤ 7 (Muegge et al., 2001)

A composite drug-likeness score (0–6) was calculated for candidate molecules as the number of filters passed (excluding the Lipinski relaxed variant to avoid double-counting).

2.4. ADMET Predictions

Simplified ADMET predictions were derived from threshold-based classifications applied to the physicochemical properties in the dataset: membrane permeability (Good: TPSA < 75 Å² AND 0 < LogP < 5; Poor: TPSA > 140 Å² OR LogP < −2 OR LogP > 6; Moderate: all intermediate values) and BBB penetration (High: TPSA < 60 Å² AND LogP > 1; Low: TPSA > 120 Å² OR LogP < −1; Moderate: intermediate). These threshold-based classifications are substantially less reliable than validated

 doi:10.20944/preprints202602.0854.v1

computational tools (SwissADME BOILED-Egg model, pkCSM, ADMETlab 3.0) and should be treated as rough estimates only.

2.5. PAINS Screening

All 26 candidate molecules were evaluated for known PAINS alerts based on published criteria (Baell & Holloway, 2010). Molecules containing Michael acceptors, reactive electrophiles, quinone/redox-cycling scaffolds, or known promiscuous motifs (catechols, thioallyl groups) were flagged. This was a manual structural assessment guided by AI, not a computational PAINS filter (e.g., RDKit FilterCatalog).

2.6. Statistical Analysis

Mann-Whitney U tests for continuous property comparisons; Fisher's exact test for categorical pass/fail comparisons; rank-biserial correlation (r) as effect size metric. Bonferroni correction applied separately for property comparisons (6 tests, adjusted α = 0.0083) and filter comparisons (7 tests, adjusted α = 0.0071). Both uncorrected and Bonferroni-corrected p-values are reported. Power: with n = 48 and n = 71, the Mann-Whitney U test has approximately 80% power to detect medium effects (r ≈ 0.27) at α = 0.05.

3. Results

3.1. Part 1: Three-Group Comparison Reveals Specificity of Drug-Likeness Enrichment

3.1.1. Property Distributions

Small-molecule signaling molecules showed significantly lower molecular weight, fewer hydrogen bond acceptors, and lower TPSA than random drug-like molecules — all surviving Bonferroni correction (Table 1, Figure 1). LogP, HBD, and rotatable bonds did not differ significantly after correction. Critically, peptide autoinducers and intracellular messengers showed property profiles dramatically different from both groups, confirming that drug-likeness is not a universal property of bacterial signaling molecules.

Table 1. Physicochemical properties: small-molecule signals (n = 48) vs. random drugs (n = 71).

Property	Signaling (median)	Random (median)	p-value	Bonferroni p	r
Molecular Weight (Da)	225.8	379.5	9.26×10^{-11}	5.56×10^{-10} ***	0.702
LogP	2.10	2.70	0.142	0.853 ns	0.159
H-Bond Acceptors	3.0	5.0	5.17×10^{-7}	3.10×10^{-6} ***	0.539
H-Bond Donors	1.0	2.0	0.015	0.090 ns	0.248
TPSA (Å²)	55.8	81.7	7.40×10^{-5}	4.44×10^{-4} ***	0.430
Rotatable Bonds	5.5	6.0	0.418	1.000 ns	0.088

*Mann-Whitney U tests. *** $p < 0.001$ after Bonferroni correction ($\alpha/6 = 0.0083$); ns = not significant.*

 doi:10.20944/preprints202602.0854.v1

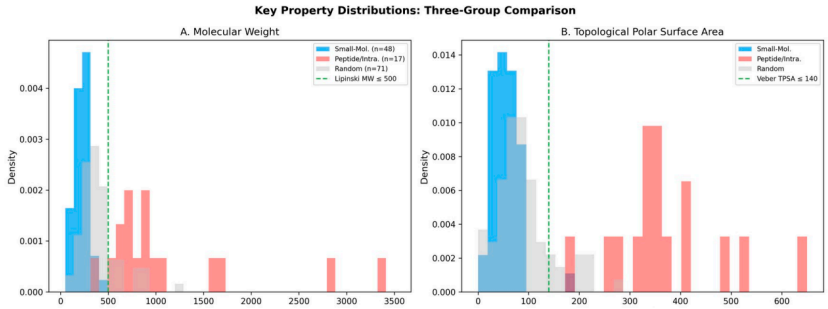

Figure 1. Key property distributions across three groups. (A) Molecular weight: small-molecule signals cluster below 500 Da, while peptides extend to >3000 Da. (B) TPSA: small-molecule signals are concentrated below 140 Å² (Veber threshold), while intracellular messengers show TPSA > 300 Å².

3.1.2. Drug-Likeness Filter Enrichment: Three-Group Comparison

The three-group comparison reveals a striking pattern: small-molecule signals showed high pass rates (56–98%), random drugs showed intermediate rates (63–83%), and peptide/intracellular signals showed near-zero rates on all filters (Table 2, Figure 2). After Bonferroni correction, three filters reach significance: Lipinski strict (adjusted p = 0.009), Egan (adjusted p = 0.004), and O'Shea antibiotic-specific (adjusted p = 0.016).

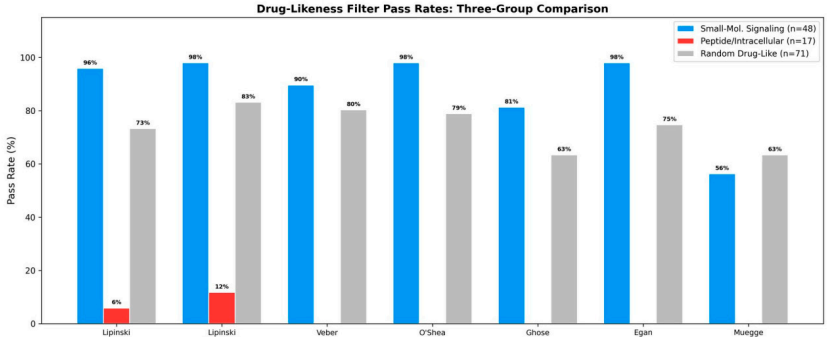

Figure 2. Drug-likeness filter pass rates across three groups. Small-molecule signaling molecules (blue) show consistently high pass rates, random drugs (gray) show intermediate rates, and peptide autoinducers/intracellular messengers (red) show near-zero rates on all filters.

Table 2. Drug-likeness filter pass rates: three-group comparison.

Filter	Small-Mol. (n=48)	Pept./Intra. (n=17)	Random (n=71)	EF	Fisher's p	Bonf. p
Lipinski Ro5 (strict)	95.8% (46/48)	5.9% (1/17)	73.2% (52/71)	1.31	0.001	0.009 **
Lipinski Ro5 (≤1 viol.)	97.9% (47/48)	11.8% (2/17)	83.1% (59/71)	1.18	0.014	0.099 ns
Veber Criteria	89.6% (43/48)	0.0% (0/17)	80.3% (57/71)	1.12	0.209	1.000 ns
O'Shea Abx	97.9% (47/48)	0.0% (0/17)	78.9% (56/71)	1.24	0.002	0.016 *
Ghose Filter	81.2% (39/48)	0.0% (0/17)	63.4% (45/71)	1.28	0.042	0.292 ns

 doi:10.20944/preprints202602.0854.v1

Egan Filter	97.9% (47/48)	0.0% (0/17)	74.6% (53/71)	1.31	< 0.001	0.004 **
Muegge Filter	56.2% (27/48)	0.0% (0/17)	63.4% (45/71)	0.89	0.451	1.000 ns

EF = Enrichment Factor (% pass small-mol. ÷ % pass random). Fisher's p compares small-mol. vs. *random only. The Muegge exception (EF = 0.89) reflects its MW ≥ 200 lower bound excluding small signals (AI-2: 132 Da; Indole: 117 Da).*

Key finding: Lipinski compliance. Small-molecule signaling: 46/48 (95.8%); Peptide/intracellular: 1/17 (5.9%); Random drugs: 52/71 (73.2%). Fisher's exact (small-mol. vs. random): OR = 8.40, 95% CI: 1.86–38.1, p = 0.001; Bonferroni-adjusted p = 0.009.

3.1.3. Chemical Space Mapping

The chemical space map (Figure 3) shows the three distinct populations. Small-molecule signals cluster tightly within both the Lipinski zone (MW ≤ 500, LogP ≤ 5) and the antibiotic-relevant O'Shea zone, while peptide autoinducers (MW 680–3354 Da) and intracellular messengers (LogP −5 to −7) fall far outside drug-like chemical space. Random drugs show the expected broad distribution.

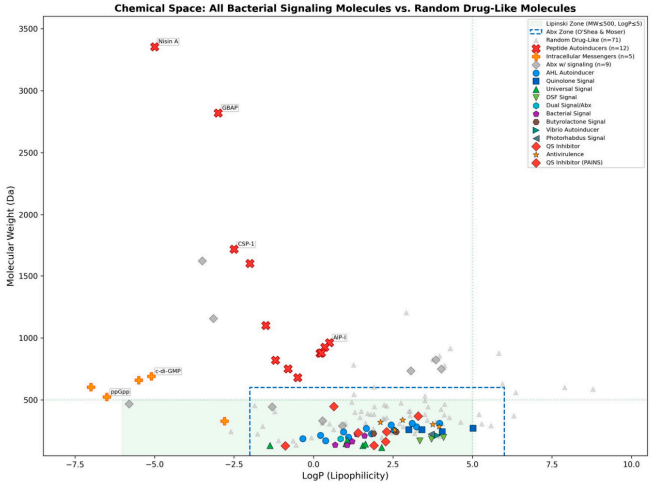

Figure 3. Chemical space (LogP vs. MW) showing all 74 bacterial signaling molecules and 71 random drug-like molecules. Small-molecule signals (colored by category) cluster within the Lipinski zone (green shading) and antibiotic-relevant zone (blue dashes). Peptide autoinducers (red ×) and intracellular messengers (orange +) fall far outside drug-like space. Key negative examples annotated.

3.1.4. Predicted ADMET Properties

72.9% of small-molecule signals were classified as "Good" predicted membrane permeability vs. 32.4% of random drugs and 0% of peptide/intracellular signals (Table 3).

Table 3. Predicted membrane permeability: three-group comparison.

	Good	Moderate	Poor
Small-Mol. Signaling (n = 48)	72.9% (35/48)	25.0% (12/48)	2.1% (1/48)
Peptide/Intracellular (n = 17)	0.0% (0/17)	0.0% (0/17)	100.0% (17/17)

 doi:10.20944/preprints202602.0854.v1

Random Drug-Like (n = 71)	32.4% (23/71)	46.5% (33/71)	21.1% (15/71)

Based on TPSA/LogP thresholds (Section 2.4); treat as rough estimates.

3.2. Part 2: Candidate Identification

Among 26 curated small-molecule signaling molecules (Supplementary Table S3), 10 were classified as HIGH priority based on composite drug-likeness score ≥ 5/6 and a documented testing gap. After PAINS assessment, three were flagged for structural alerts (Table 4). The six highest-confidence candidates — Savirin, LED209, LsrK inhibitor 11e, 3-oxo-C12-HSL, Hordenine, and Indole-3-acetic acid — are free of major PAINS alerts.

Table 4. High-priority signaling-active candidates.

Molecule	Category	MW	LogP	Score	Testing Gap	PAINS
Savirin (M64)	Antivirulence	302.4	3.75	6/6	MIC not reported	None
LED209	Antivirulence	338.4	2.80	6/6	Designed non-bactericidal	None
LsrK inhib. 11e	AI-2 Modulator	320.0	2.10	6/6	Untested (Milli et al., 2024)	None
3-oxo-C12-HSL	QS Autoinducer	297.4	2.44	5/6	Untested at MIC range	None
Hordenine	QS Modulator	165.2	1.21	5/6	Untested at high conc.	None
Indole-3-acetic acid	Interspecies	175.2	1.06	5/6	Untested as antibiotic	None
Virstatin	Antivirulence	286.3	3.95	6/6	Unexplored at high conc.	Acridone — monitor
mBTL	QS Inhibitor	257.1	2.55	6/6	MIC not published	Thiolactone — reactive
Furanone C-30	QS Modulator	243.5	2.30	6/6	Not tested	Michael acceptor
C4-HSL	QS Autoinducer	171.2	0.38	5/6	Untested as antibiotic	Lactone — labile

Score = filters passed out of 6 (Lipinski strict, Veber, O'Shea, Ghose, Egan, Muegge). All 10 show 0 Lipinski violations.

4. Discussion

4.1. Drug-Likeness Enrichment Is Specific to Small-Molecule Diffusible Signals

The three-group comparison provides stronger evidence than a simple two-group design. Small-molecule diffusible signals show ~96% Lipinski compliance (OR = 8.40, Bonferroni-adjusted p = 0.009), while peptide autoinducers and intracellular messengers show ~6% compliance. This demonstrates that drug-likeness enrichment is not a universal property of all bacterial signaling but is specific to the class of small molecules that function via passive diffusion and membrane transit.

 doi:10.20944/preprints202602.0854.v1

Three filters now survive Bonferroni correction (Lipinski, Egan, O'Shea), up from one in the smaller initial dataset.

This suggests a practical application: in silico filtering for signaling-like properties simultaneously enriches for drug-likeness (EF 1.12–1.31), membrane permeability (72.9% vs. 32.4% "Good"), and biological relevance to bacterial physiology. This convergence is further supported by findings that indole signaling modulates efflux pump expression (Salama et al., 2024) and that resistance machinery components can serve as signal transducers (Koberska et al., 2021).

4.2. The Testing Gap and Dose-Escalation Rationale

The antivirulence compounds (Savirin, LED209, Virstatin, mBTL) were explicitly designed NOT to kill bacteria — the question "what happens at 10–100× the anti-virulence EC_{50}?" has not been systematically published. 3-oxo-C12-HSL induces eukaryotic cell apoptosis at ~100 μM (Tateda et al., 2003), yet its effects on bacterial viability remain uncharacterized. Indole-3-acetic acid directly controls antibiotic production in bacteria (Matilla et al., 2018).

We acknowledge that unpublished negative data may exist, that most virulence-associated QS systems are not essential for growth under laboratory conditions, and that our testing gap assessment was AI-mediated rather than a systematic review.

4.3. Biological Considerations

Species-specificity. QS systems differ between Gram-negative (AHL-based) and Gram-positive (peptide-based) bacteria (Whiteley et al., 2017). Universal signals (AI-2, indole) may offer broader targets. **Toxicity.** 3-oxo-C12-HSL modulates NF-κB signaling (Kravchenko et al., 2008); Hordenine is sympathomimetic. **Resistance.** Efflux pumps export AHL and quinolone signals; AHL lactonases degrade AHL candidates; LasR loss-of-function mutants are common in chronic infections (Feltner et al., 2016). **PAINS.** Furanone C-30, curcumin (Nelson et al., 2017), and cinnamaldehyde contain reactive electrophilic motifs and were deprioritized (Baell & Holloway, 2010).

4.4. Limitations

(1) AI-assisted methodology. This work used Claude (Anthropic) as the primary computational tool. The author lacks formal cheminformatics training. AI analyses are stochastic and may differ across sessions. Results should be considered provisional until reproduced using deterministic workflows (RDKit, OpenBabel).

(2) Compiled rather than computed properties. Properties were sourced from PubChem/DrugBank, not computed from canonical SMILES with a single validated toolkit. Different algorithms may underlie values across molecules. This is the most critical methodological weakness.

(3) Peptide property estimation. Properties for Gram-positive peptide autoinducers (n = 12) were estimated from published data and sequence-based calculations. These values are sufficient to demonstrate peptides fall far outside drug-like space but are not experimentally determined.

(4) Sample sizes. n = 48 vs. n = 71 provides ~80% power for $r \approx 0.27$. Larger validation against ChEMBL/ZINC (n > 10,000) is needed.

(5) Reference set bias. The random set consists primarily of FDA-approved drugs, already biased toward drug-likeness.

(6) Simplified ADMET and PAINS. Threshold-based ADMET (Section 2.4) and manual PAINS assessment are less reliable than validated tools (SwissADME, pkCSM, RDKit FilterCatalog).

(7) No experimental validation. Drug-likeness does not predict antibacterial activity.

4.5. Recommended Validation Pathway

Computational (immediate): Recompute all properties from canonical SMILES using RDKit; validated ADMET via SwissADME/pkCSM; RDKit PAINS filters; expand to n > 100 from ChEMBL; compare against 10,000+ unbiased compounds.

 doi:10.20944/preprints202602.0854.v1

Literature: PRISMA-compliant systematic review for MIC data on all candidates.

Experimental: MIC at up to 256 μg/mL against ESKAPE pathogens; dose-response profiling; cytotoxicity (HepG2, HEK293); checkerboard synergy assays.

Definitive test: Prospective comparison of signaling-enriched vs. random library hit rates.

5. Conclusions

This AI-assisted exploratory analysis provides preliminary evidence that small-molecule bacterial signaling molecules exhibit drug-like physicochemical profiles, while peptide-based and intracellular signaling molecules do not. The strongest evidence is Lipinski compliance (OR = 8.40, Bonferroni-adjusted $p = 0.009$); three of seven drug-likeness filters now survive Bonferroni correction. Ten candidates with testing gaps were identified, six free of PAINS alerts.

The ultimate question is empirical: do signaling molecules yield antibiotic hits at rates higher than random libraries?

Supplementary Materials: The following supporting information can be downloaded at website of this paper posted on Preprints.org, **Supplementary Table S1:** Complete signaling molecule dataset (n = 74; 48 primary + 17 negative + 9 excluded) with all physicochemical properties, categories, signaling roles, and subset assignments. **Supplementary Table S2:** Complete random drug-like molecule dataset (n = 71) with all properties, therapeutic area classifications, and recalculated ADMET predictions. **Supplementary Table S3:** Candidate molecule dataset (n = 26) with testing gap annotations, PAINS flags, and priority classifications. All deposited as CSV files.

Author Contributions: M.A.A. conceived the framework, designed prompts, directed the AI-assisted analysis, critically evaluated outputs, and assumes full intellectual responsibility. Claude (Anthropic) was used as a computational tool for data curation, statistical analysis, figure generation, and text drafting.

Funding: No specific funding was received.

AI Declaration: This work was generated with AI assistance (Claude Opus 4.6 Extended, Anthropic; December 2024 – February 2025). The complete workflow is described in Section 2.1. The author assumes full intellectual responsibility and invites independent validation using standard cheminformatics tools.

Conflicts of Interest: The author declares no conflicts of interest.

Glossary of Abbreviations and Acronyms

Abbreviation	Full Term
ADMET	Absorption, Distribution, Metabolism, Excretion, and Toxicity
AHL	N-Acyl Homoserine Lactone
AI-2	Autoinducer-2
AIP	Autoinducing Peptide
AMR	Antimicrobial Resistance
BBB	Blood–Brain Barrier
BDSF	Burkholderia Diffusible Signal Factor (cis-2-Dodecenoic Acid)
CAI-1	Cholera Autoinducer-1
cAMP	Cyclic Adenosine Monophosphate
CDSF	cis,cis-11-Methyldodeca-2,5-dienoic Acid
c-di-AMP	Cyclic di-Adenosine Monophosphate
c-di-GMP	Cyclic di-Guanosine Monophosphate
CI	Confidence Interval
CNS	Central Nervous System

 doi:10.20944/preprints202602.0854.v1

CSP	Competence-Stimulating Peptide
Da	Dalton
DPD	(4S)-4,5-Dihydroxy-2,3-pentanedione
DSF	Diffusible Signal Factor
EC50	Half-Maximal Effective Concentration
EF	Enrichment Factor
ESKAPE	Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, Enterobacter species
FDA	U.S. Food and Drug Administration
GBAP	Gelatinase Biosynthesis-Activating Pheromone
HBA	Hydrogen Bond Acceptors
HBD	Hydrogen Bond Donors
HHQ	4-Hydroxy-2-Heptylquinoline
HQNO	4-Hydroxy-2-Heptylquinoline N-oxide
HSL	Homoserine Lactone
LogP	Logarithm of Octanol/Water Partition Coefficient
MIC	Minimum Inhibitory Concentration
MR	Molar Refractivity
MW	Molecular Weight
NF-κB	Nuclear Factor Kappa-light-chain-enhancer of Activated B Cells
NHQ	2-Nonyl-4-Hydroxyquinoline
ns	Not Significant
OR	Odds Ratio
PAINS	Pan Assay Interference Compounds
ppGpp	Guanosine Tetraphosphate
pppGpp	Guanosine Pentaphosphate
PQS	Pseudomonas Quinolone Signal
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QS	Quorum Sensing
RND	Resistance-Nodulation-Division
Ro5	Rule of Five
SHP	Short Hydrophobic Peptide
TDA	Tropodithietic Acid
TPSA	Topological Polar Surface Area
VB-A	Virginiae Butanolide A

References

1. 1. Amer, A.N. et al. (2025). Growth-phase-dependent modulation of quorum sensing and virulence factors in Pseudomonas aeruginosa ATCC 27853 by sub-MICs of antibiotics. Antibiotics 14, 4.
2. Andersson, D.I. & Hughes, D. (2014). Microbiological effects of sublethal levels of antibiotics. Nat. Rev. Microbiol. 12, 465–478.
3. Baell, J.B. & Holloway, G.A. (2010). New substructure filters for removal of pan assay interference compounds (PAINS). J. Med. Chem. 53, 2719–2740.

 doi:10.20944/preprints202602.0854.v1

4. Beyersmann, P.G. et al. (2017). Dual function of tropodithietic acid as antibiotic and signaling molecule in global gene regulation of the probiotic bacterium Phaeobacter inhibens. Sci. Rep. 7, 730.
5. Brown, E.D. & Wright, G.D. (2016). Antibacterial drug discovery in the resistance era. Nature 529, 336–343.
6. Davies, J. & Davies, D. (2010). Origins and evolution of antibiotic resistance. Microbiol. Mol. Biol. Rev. 74, 417–433.
7. Egan, W.J., Merz, K.M. & Baldwin, J.J. (2000). Prediction of drug absorption using multivariate statistics. J. Med. Chem. 43, 3867–3877.
8. Fajardo, A. & Martínez, J.L. (2008). Antibiotics as signals that trigger specific bacterial responses. Curr. Opin. Microbiol. 11, 161–167.
9. Feltner, J.B. et al. (2016). LasR variant cystic fibrosis isolates reveal an adaptable quorum-sensing hierarchy in Pseudomonas aeruginosa. mBio 7, e01513-16.
10. Ghose, A.K. et al. (1999). A knowledge-based approach in designing combinatorial or medicinal chemistry libraries. J. Comb. Chem. 1, 55–68.
11. Hoffman, L.R. et al. (2005). Aminoglycoside antibiotics induce bacterial biofilm formation. Nature 436, 1171–1175.
12. Koberska, M. et al. (2021). Beyond self-resistance: ABCF ATPase LmrC is a signal-transducing component of an antibiotic-driven signaling cascade accelerating the onset of lincomycin biosynthesis. mBio 12, e01731-21.
13. Kravchenko, V.V. et al. (2008). Modulation of gene expression via disruption of NF-κB signaling by a bacterial small molecule. Science 321, 259–263.
14. Linares, J.F. et al. (2006). Antibiotics as intermicrobial signaling agents instead of weapons. Proc. Natl. Acad. Sci. U.S.A. 103, 19484–19489.
15. Lipinski, C.A. et al. (2001). Experimental and computational approaches to estimate solubility and permeability. Adv. Drug Deliv. Rev. 46, 3–26.
16. Matilla, M.A. et al. (2018). An auxin controls bacterial antibiotics production. Nucleic Acids Res. 46, 11229–11238.
17. Milli, G. et al. (2024). New LsrK ligands as AI-2 quorum sensing interfering compounds against biofilm formation. J. Med. Chem. 67, 18139–18156.
18. Muegge, I. et al. (2001). Simple selection criteria for drug-like chemical matter. J. Med. Chem. 44, 1841–1846.
19. Nelson, K.M. et al. (2017). The essential medicinal chemistry of curcumin. J. Med. Chem. 60, 1620–1637.
20. O'Shea, R. & Moser, H.E. (2008). Physicochemical properties of antibacterial compounds. J. Med. Chem. 51, 2871–2878.
21. Romero, D. et al. (2011). Antibiotics as signal molecules. Chem. Rev. 111, 5492–5505.
22. Salama, G.G. et al. (2024). Downregulation of Klebsiella pneumoniae RND efflux pump genes following indole signal produced by Escherichia coli. BMC Microbiol. 24, 312.
23. Tateda, K. et al. (2003). The Pseudomonas aeruginosa autoinducer N-3-oxododecanoyl homoserine lactone accelerates apoptosis in macrophages and neutrophils. Infect. Immun. 71, 5785–5793.
24. Veber, D.F. et al. (2002). Molecular properties that influence the oral bioavailability of drug candidates. J. Med. Chem. 45, 2615–2623.
25. Whiteley, M. et al. (2017). Progress in and promise of bacterial quorum sensing research. Nature 551, 313–320.
26. Williams, C.C. et al. (2024). Signaling molecules in Pseudomonas aeruginosa response to antibiotics at sub-inhibitory concentrations. Asian J. Biotechnol. Bioresour. Technol. 10, 19–33.