Research Article

BIOMEDICAL Journal of Scientific & Technical Research

ISSN: 2574 -1241

DOI: 10.26717/BJSTR.2026.67.010494

Identification of Natural Ligands Targeting LivK of Porphyromonas gingivalis Through Virtual Screening and Its Potential as a Molecular Target in Periodontitis

Salvador Pérez-Mora[1,2,3]*

[1]*Undergraduate Student in Dentistry, Universidad UCAD, Mexico City 07340, Mexico*

[2]*Doctoral Student in Educational Sciences, Centro de Estudios Profesionales y de Investigación (ESPI), Ciudad López Mateos 52970, State of Mexico, Mexico*

[3]*Postdoctoral Researcher, Molecular Biomedicine Laboratory I, Escuela Nacional de Medicina y Homeopatía (ENMyH), Instituto Politécnico Nacional (IPN), Mexico City 07320, Mexico*

***Corresponding author:** Salvador Pérez-Mora, Undergraduate Student in Dentistry, Universidad UCAD, Mexico City 07340, Mexico, Doctoral Student in Educational Sciences, Centro de Estudios Profesionales y de Investigación (ESPI), Ciudad López Mateos 52970, State of Mexico, Mexico, Postdoctoral Researcher, Molecular Biomedicine Laboratory I, Escuela Nacional de Medicina y Homeopatía (ENMyH), Instituto Politécnico Nacional (IPN), Mexico City 07320, Mexico

ARTICLE INFO

Received: November 18, 2026

Published: November 25, 2026

Citation: Salvador Pérez-Mora. Identification of Natural Ligands Targeting LivK of Porphyromonas gingivalis Through Virtual Screening and Its Potential as a Molecular Target in Periodontitis. Biomed J Sci & Tech Res 67(3)-2026. BJSTR. MS.ID.010494.

ABSTRACT

Porphyromonas gingivalis is a periodontal pathogen whose nutritional adaptation and invasive capacity contribute to its persistence and potential extraoral dissemination. LivK, a periplasmic substrate-binding protein associated with branched-chain amino acid (BCAA) transport, represents a molecular target of interest due to its association with infection-related phenotypes. This study aimed to structurally characterize LivK and prioritize natural compounds through virtual screening. The AlphaFold model AF-B2RKF2-F1 was evaluated using structural quality assessments and comparisons with experimentally determined homologs. A putative binding site was identified using P2Rank and geometric cavity detection, while an independent assessment using the 1USK homolog recovered 83.3% of the residues comprising its crystallographic leucine-binding site. A total of 738,794 compounds from the COCONUT database were processed through physicochemical filtering, AutoDock Vina molecular docking, ligand efficiency analysis, pharmacophore-based selection, and refinement using five independent random seeds. Ten candidates with more favorable docking scores than L-leucine were prioritized. These ten compounds and the three reference substrates shared contacts with seven residues: Thr138, Arg151, Ser153, Pro154, Glu159, Tyr183, and Gly257.

Integration of physicochemical properties, ADME predictions, and toxicity alerts highlighted four candidates: pentenomycin I, (2R,3R,4S)-4-hydroxyisoleucine, (R)-piperazic acid, and 2-oxoimidazolidine-4-carboxylic acid. These findings support the computational prioritization of potentially competitive LivK ligands but do not establish direct binding or functional inhibition. Experimental binding, competition, and transport assays are required to determine their relevance to periodontal colonization.

Keywords: *Porphyromonas gingivalis*; Periodontitis; LivK; Bacterial Nutrient Acquisition; Branched-Chain Amino Acid Transport; Structure-Based Drug Discovery; Natural Product Screening; Molecular Docking; Protein–Ligand Interactions; Therapeutic Target Identification

Abbreviations: LIV: Leucine–Isoleucine–Valine; ABC: ATP-Binding Cassette; PAE: Predicted Aligned Error; RMSD: Root-Mean-Square Deviation; TPSA: Topological Polar Surface Area; QED: Quantitative Estimate of Drug-Likeness; ADME: Absorption, Distribution, Metabolism, and Excretion; PAE: Predicted Aligned Error; BCAA: Branched-Chain Amino Acid; GO: Gene Ontology; COG: Clusters of Orthologous Groups; DPP7: Dipeptidyl Peptidase 7

Introduction

Periodontitis is a chronic inflammatory disease driven by a dysbiotic microbial community that progressively destroys the periodontal ligament and alveolar bone. In its severe form, it affects approximately 11% of the global adult population and ranks as the sixth most prevalent human condition [1,2]. In dental practice, periodontitis is the leading cause of tooth loss in adults, with direct consequences for masticatory function, aesthetics, and quality of life, as well as substantial treatment costs that place it among the most economically burdensome oral diseases [2]. Unlike dental caries, its management cannot always be achieved through mechanical biofilm removal alone, as a proportion of patients continue to experience attachment loss despite appropriate periodontal therapy. This has sustained interest in identifying molecular targets that could complement periodontal debridement [3,4]. *Porphyromonas gingivalis (P. gingivalis)* occupies a distinctive position within this microbial community. Although it is not the most abundant microorganism in the periodontal pocket, it acts as a keystone pathogen capable of reshaping the subgingival biofilm even at low abundance. By manipulating the complement system and Toll-like receptor signaling, it subverts host immune responses and promotes the transition from a commensal community to a dysbiotic, tissue-destructive microbiota [3,4].

Its virulence repertoire includes gingipains, fimbriae, atypical lipopolysaccharide, and peptidylarginine deiminase. Furthermore, its ability to invade gingival epithelial and endothelial cells enables persistence within protected niches, where it may evade antimicrobial agents and host immune defenses [5,6]. This invasive capacity also extends the relevance of periodontitis beyond the oral cavity. Transient bacteremia associated with chewing, toothbrushing, or periodontal scaling may facilitate the dissemination of *P. gingivalis* and its products to distant organs and tissues. The bacterium has been detected in atherosclerotic plaques, synovial fluid, liver, placenta, and brain tissue [5,7-9]. Epidemiological studies and clinical consensus reports have linked periodontitis to atherosclerotic cardiovascular disease, type 2 diabetes mellitus, rheumatoid arthritis, adverse pregnancy outcomes, nonalcoholic fatty liver disease, and Alzheimer's disease [5,7-10]. These systemic associations have broadened the clinical significance of periodontitis, emphasizing that periodontal treatment addresses not only the loss of tooth-supporting tissues but also a potential source of chronic inflammation with systemic implications. Consequently, understanding the mechanisms that support *P. gingivalis* persistence within the host has become increasingly relevant. One such mechanism involves nutritional adaptation. *P. gingivalis* is an asaccharolytic bacterium that does not ferment carbohydrates and instead relies primarily on peptides and amino acids as sources of carbon and energy.

This metabolic characteristic is particularly important in deep periodontal pockets, where sugars are scarce and gingival crevicular fluid provides host-derived proteins [11,12]. To exploit these nutritional resources, the bacterium employs an integrated system of proteolytic degradation and nutrient uptake involving gingipains, dipeptidyl and tripeptidyl peptidases, and specialized transporters such as RagAB [6,11-14]. Because these pathways are essential to its nutritional metabolism and cannot be replaced by carbohydrate utilization, nutrient recognition and transport represent attractive yet relatively underexplored targets for molecular intervention. Within this nutritional network, the leucine–isoleucine–valine (LIV) transport system is of particular interest. LivK is the periplasmic substrate-binding component of an ATP-binding cassette (ABC) transporter involved in branched-chain amino acid (BCAA) uptake. Using dual RNA sequencing (GSE184085), Wu and colleagues observed significantly increased expression of livh and livk in intracellular *P. gingivalis* recovered from infected human umbilical vein endothelial cells compared with bacteria grown in culture medium. Deletion of *livk* reduced cellular invasion and hepatic colonization in mice fed a high-fat diet, whereas deletion of either *livh* or *livk* attenuated increases in circulating BCAA levels and markers of liver injury [15].

These findings are particularly relevant because they demonstrate induction of the LIV system during intracellular endothelial infection, a biological context that may contribute to the relationship between periodontal infection and systemic disease. Independent studies have associated alterations in BCAA metabolism in *P. gingivalis* with systemic metabolic effects in mice [16]. Furthermore, substrate-binding proteins of the LIV system in *Campylobacter jejuni* have been implicated in colonization-related functions extending beyond amino acid transport [17]. Together, these observations suggest that the biological relevance of LivK may extend beyond nutrient acquisition, making the characterization of its ligand-recognition properties a promising starting point for investigating its potential modulation. An additional practical consideration is its periplasmic localization. Because nutrients reach the periplasm through outer membrane porins, compounds targeting LivK would not need to cross the bacterial inner membrane to enter the cytoplasm [18,19]. Moreover, ABC import systems and their associated substrate-binding proteins are found in bacteria, archaea, and certain plants, but are absent in humans, potentially reducing the likelihood of direct interactions with homologous host proteins [18,20]. Virtual screening of natural compounds offers an opportunity to explore molecules with structural and physicochemical characteristics resembling those of the metabolites recognized by these proteins [21].

Nevertheless, docking scores must be interpreted cautiously because scoring functions tend to favor larger molecules and should therefore be assessed alongside ligand efficiency and the geometry of protein–ligand interactions [22]. The objective of this study was to characterize the substrate-recognition site of LivK from *P. gingivalis* using independent structural approaches and, based on this characterization, to identify and prioritize a focused set of natural compounds from the COCONUT database that reproduce key interaction patterns of the natural substrates and may be suitable for subsequent experimental evaluation.

Materials and Methods

Target Selection

LivK from *P. gingivalis* ATCC 33277 (UniProt B2RKF2; PGN_RS06345 / PGN_1328) was selected based on previously published experimental evidence [15]. The study associated with dataset GSE184085 reported 7.54- and 4.40-fold increases in the expression of *livh* and *livk*, respectively, in bacteria recovered from infected endothelial cells. Accordingly, LivK was selected as the structural target. Its selection over LivH, despite the greater induction of the latter, was based on three criteria: LivK is the substrate-binding component directly involved in substrate recognition and system selectivity; its periplasmic localization makes it accessible without requiring compounds to cross the inner membrane; and, unlike *livh* deletion, *livk* deletion reduces cellular invasion and hepatic colonization [15].

Three-Dimensional Structural Model of LivK and Structural Validation

The predicted structure was retrieved from the AlphaFold Protein Structure Database (https://alphafold.ebi.ac.uk/; entry AF-B2RKF2-F1, model version 6; accessed October 20, 2026) [23,24]. The N-terminal signal peptide and low-confidence C-terminal segment were excluded, retaining residues 38–373 as the mature protein model for structural calculations and molecular docking. The protein sequence and locus identifiers were verified using UniProt (https://www.uniprot.org/; entry B2RKF2; accessed October 2026) [25]. Local and interdomain structural confidence was assessed using the predicted Local Distance Difference Test (pLDDT), which estimates residue-level confidence in the predicted model, and the Predicted Aligned Error (PAE) matrix, which evaluates uncertainty in the relative positioning and orientation of different structural regions.

Structural geometry was examined using Ramachandran analysis through PDBsum (https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/; accessed October 2026) [26] to evaluate the distribution of protein backbone dihedral angles and identify residues in unfavorable conformational regions; ProSA-web (https://prosa.services.came.sbg.ac.at/prosa.php; accessed October 2026) [27] to assess overall model quality through its Z-score and identify potential energetic anomalies; ERRAT, implemented through the SAVES server (https://saves.mbi.ucla.edu/; accessed October 2026) [28], to examine the quality of nonbonded atomic interactions; and VADAR (http://vadar.wishartlab.com/; accessed October 25, 2026) [29] to characterize geometric and structural properties, including solvent accessibility, secondary structure, and residue conformational parameters.

Binding-Site Definition and Assessment Using Experimental Homologs

The three-dimensional model of LivK was superimposed onto the experimentally determined structures 1USK from *Escherichia coli (E. coli)*, containing crystallographic L-leucine, and 3TD9 from *Thermotoga maritima (T. maritima)*, using CEalign in PyMOL 3.0. The root-mean-square deviation (RMSD) and number of aligned residues were recorded [30,31]. Potential binding cavities were identified using P2Rank 2.5.1, implemented through the PrankWeb server [32,33], and a LIGSITE-like geometric cavity detection algorithm developed in Python using a 1.0 Å grid [34]. The geometric centers and residue sets identified by both approaches were compared to assess the spatial agreement between the predicted regions. P2Rank identified 13 residues contained within the main cavity of 38 residues delineated by the geometric procedure. Accordingly, the intersection of both sets was defined as a structural consensus core comprising 13 residues: Leu54, His136, Thr138, Ser153, Pro154, Glu159, Tyr183, Leu256, Gly257, Val260, Phe278, Ile304, and Ser305. For subsequent virtual screening analyses, an expanded operational binding pocket comprising 16 residues was used, consisting of the 13 overlapping positions and three additional residues: Thr155, Val156, and Asn232. This expansion was considered an exploratory methodological definition of the binding environment and established the residue set used to evaluate protein–ligand contacts.

The three additional residues were distinguished from those identified by consensus, without attributing independent validation to them or assuming that their inclusion demonstrated a functional role in substrate recognition. The center of the final docking search box was determined independently of the number of residues included in the operational binding pocket. Exploratory molecular docking of L-leucine, L-isoleucine, and L-valine was performed using a 24 × 33 × 20 Å box centered at coordinates (−4.34, −0.16, −7.27) Å. Heavy-atom centroids were calculated from the highest-ranked poses of these amino acids to establish the center of the final docking box at (2.50, 2.41, −5.58) Å. This procedure was based exclusively on the natural reference substrates, without incorporating information from the COCONUT compounds subsequently subjected to virtual screening. As an independent control for geometric cavity detection, the experimental L-leucine-binding site of homolog 1USK was used. This site was defined by chain A residues located within 4.0 Å of any heavy atom of the crystallographic ligand LEU A1347. After removing the ligand, fpocket 4.2.3 was executed using its default parameters [35]. Agreement between the experimental binding site and the predicted cavity was quantified using coverage, $C = |E \cap P| / |E|$, and the Jaccard index, $J = |E \cap P| / |E \cup P|$, where E represents the set of experimental binding-site residues and P represents the set of predicted residues.

These metrics were used to assess the geometric recovery of a known binding pocket, without interpreting this assessment as direct validation of AutoDock Vina's ability to reproduce the crystallographic ligand orientation. Additionally, fpocket was applied to the LivK model as a complementary analysis of geometric sensitivity. Positive spatial correspondence was defined as at least 50% of the heavy atoms of docked L-leucine being located within 3.0 Å of any alpha sphere identified by the algorithm. This criterion was used exclusively to examine spatial agreement between the geometric cavity and the reference docking pose, without considering it experimental evidence of molecular recognition.

Structural Comparison and Alignment of LivK with Experimental Homologs

The structural similarity of LivK to the experimental homologs 1USK from *E. coli* and 3TD9 from *T. maritima* was assessed using two complementary procedures. In addition, structural alignments were performed with TM-align using the alignment tool of the RCSB Protein Data Bank (https://www.rcsb.org/alignment; accessed November 5, 2026). The TM-score, RMSD, number of aligned residues, and sequence identity were recorded for the LivK–1USK, LivK–3TD9, and 1USK–3TD9 comparisons in order to assess the conservation of the three-dimensional fold among the structures [30,31]. Structural superpositions and image generation were performed in PyMOL 3.0.

Natural Compound Library and Ligand Preparation

The October 2026 version of the COCONUT database [21,36] was processed in batches of 100,000 records using Python 3.11.17 and pandas 3.1.0, excluding entries lacking canonical SMILES or molecular weight information. The following criteria were then applied sequentially: molecular weight between 100 and 400 Da; no more than six rotatable bonds; a maximum of five hydrogen-bond donors and eight hydrogen-bond acceptors; topological polar surface area (TPSA) ≤140 $Å^2$; ALogP between −3 and 4; no more than one violation of Lipinski's rule of five; and 7–28 heavy atoms. Salts and mixtures represented by multicomponent SMILES, structures containing metals or metalloids, glycosides identified through available annotations, and duplicates based on canonical SMILES were excluded. To further match molecular size to the LivK binding cavity, additional restrictions of molecular weight ≤200 Da and ≤15 heavy atoms were applied. Ligands underwent chemical consistency verification using RDKit 2026.03.1 [37], including assessment of atomic valences, aromaticity, and structural properties. Three-dimensional conformers were generated using ETKDG (random seed 42) and optimized with the MMFF94 force field whenever possible (maximum of 1,000 iterations). The resulting structures were subsequently converted to PDBQT format using Open Babel 3.1.1, with assignment of Gasteiger partial charges [38]. Unspecified stereocenters were enumerated, and each configuration was processed independently.

Structures that failed to complete molecular preparation or docking successfully were excluded, and the corresponding counts were reported at each stage. Ligand preparation did not include pH-dependent protonation or tautomer enumeration. Hydrogen atoms were added according to the implicit valence of each deposited SMILES, and partial charges were assigned to the resulting structures. Consequently, ligands were docked in their neutral forms rather than the zwitterionic species predominantly present at physiological pH. The three natural substrates underwent the same preparation procedure, ensuring a consistent basis for internal comparisons among the compounds in the evaluated panel.

Molecular Docking, Preselection, and Pharmacophore-Based Filtering

Molecular docking was performed using AutoDock Vina 1.2.7 with its default scoring function [39,40]. Compounds passing the physicochemical prefiltering stage were docked using a 15 × 15 × 15 Å search box centered at (2.50, 2.41, −5.58) Å, with an exhaustiveness of 16 and up to five poses generated per ligand. The highest-ranked pose was retained for subsequent analyses. L-leucine, L-isoleucine, and L-valine were used as internal reference substrates. Poses establishing contacts with at least nine binding-site residues at distances below 4.0 Å were retained. The retained compounds were ranked by ligand efficiency, and 600 structures were preselected for redocking with an exhaustiveness of 32, after excluding compounds containing neither nitrogen nor oxygen atoms. The pharmacophore-based filter included four criteria: N/O atom proximities within 3.5 Å reaching at least 60% of the minimum value observed among the reference substrates, rounded down; distances below 4.0 Å to the carboxyl oxygen atoms of Glu159, Asp181, Asp182, Asp277, or Glu306, included as geometric references for evaluating polar contacts without expanding the consensus binding site; contacts with at least nine binding-site residues; and the presence of at least two nitrogen or oxygen atoms in the ligand.

Compounds satisfying these criteria were ranked using a combined index incorporating polar atom proximities and ligand efficiency.

The 25 highest-ranked candidates and the three reference substrates were refined using an exhaustiveness of 64 and five independent random seeds (42, 137, 2024, 777, and 31,415), from which mean docking scores and standard deviations were calculated. Finally, the ten candidates with mean docking scores more favorable than that of L-leucine were selected, yielding a final panel of 13 ligands, including the three natural substrates. Ligand efficiency was calculated as LE = −E_Vina / N_heavy, where E_Vina represents the mean docking score (kcal/mol) and N_heavy denotes the number of non-hydrogen atoms. This metric normalizes the docking score according to molecular size [22].

Evaluation of the Discriminatory Ability of Molecular Docking

To evaluate the ability of molecular docking to distinguish the natural substrates of LivK from other structurally related compounds, 42 molecules were analyzed: L-leucine, L-isoleucine, and L-valine (natural substrates), together with 39 amino acids and analogs encompassing stereoisomers, structural modifications, and amino acids with different chemical properties. This analysis was designed to examine the specificity of molecular recognition, considering that a favorable docking score does not necessarily imply binding affinity or selective recognition by LivK. Accordingly, docking scores, ligand efficiency,

and binding-pocket occupancy were evaluated collectively to identify potential limitations of the method in discriminating between natural substrates and alternative compounds. The analysis was performed using AutoDock Vina with an 18 × 18 × 18 Å search box centered on the binding site and an exhaustiveness of 64.

Verification of Chemical Identity

The chemical identities of the ten final candidates were verified using canonical SMILES, molecular formulas, InChIKeys, and stereochemical configurations, recalculated with RDKit 2026.03.1 [37] from the structures actually used for docking, independently of the compound names recorded in COCONUT. Absolute configurations were assigned according to the Cahn–Ingold–Prelog rules, and compound identities were cross-checked against PubChem (https://pubchem.ncbi.nlm.nih.gov/; accessed October 2026) and the Human Metabolome Database (https://hmdb.ca/; accessed October 2026).

Molecular Interaction Analysis

The interactions of the ten final candidates and the three reference substrates with LivK were evaluated using custom Python 3.11.17 scripts and NumPy 2.4.6. Contacting residues were identified using distances below 4.0 Å between heavy atoms, while proximities between nitrogen and oxygen atoms within 3.5 Å were analyzed separately as indicators of potential polar interactions. The analysis included the 16 binding-site residues and Arg151, which was subsequently incorporated for contact characterization. Two-dimensional protein–ligand interaction diagrams were generated using BIOVIA Discovery Studio Visualizer 2024.

Physicochemical Profiling, ADME, and Predictive Toxicity Assessment

Physicochemical descriptors were calculated using RDKit 2026.03.1 [37], including molecular weight, Crippen logP, topological polar surface area (TPSA), hydrogen-bond donors and acceptors, fraction of sp^3-hybridized carbon atoms, and quantitative estimate of drug-likeness (QED) [41]. Lipinski, Veber, Ghose, and Egan criteria were applied descriptively, without excluding compounds, considering the periplasmic localization of LivK. Absorption, distribution, metabolism, and excretion (ADME) properties and medicinal chemistry indicators were obtained from the SwissADME server (http://www.swissadme.ch/; accessed October 26, 2026) [42], including gastrointestinal absorption, blood–brain barrier permeability, P-glycoprotein substrate status, inhibition of cytochrome P450 isoforms, Lipinski violations, PAINS and Brenk alerts, and consensus logP. Toxicity was independently assessed using ProTox 3.0 (https://tox.charite.de/protox3/; accessed October 26, 2026) [43]. Only predictions classified as Active with probabilities ≥0.70 were retained in the integrated table.

An exploratory three-tier classification was established to prioritize the candidates. The high-priority category comprised compounds without Brenk alerts or ProTox Active predictions reaching the established threshold. The intermediate-priority category included compounds presenting Brenk alerts without non-cardiotoxicity predictions above the threshold, or compounds showing only a cardiotoxicity signal.

The low-priority category included compounds with at least one active non-cardiotoxicity prediction exceeding the threshold. Cardiotoxicity was excluded as a discriminatory criterion because all three reference substrates also yielded positive predictions. This classification was intended exclusively for internal comparison and does not constitute a validated measure of safety. RDKit and SwissADME descriptors were recorded separately, distinguishing Crippen logP from consensus logP and retaining the hydrogen-bond acceptor counts generated by each tool without assuming equivalence between methods.

Genomic Context and Functional Characterization

The genomic context of *P. gingivalis* ATCC 33277 was analyzed using GenBank (AP009380.1; assembly GCA_000010505.1; https://www.ncbi.nlm.nih.gov/genbank/) and KEGG (organism code: pgn; https://www.kegg.jp/), accessed in October 2026 [44,45]. The correspondence between LivK identifiers in both databases was verified, and the PGN_1324–PGN_1330 genomic region was examined to characterize the location, orientation, and proximity of genes associated with the LIV system, without assuming that their genomic proximity demonstrates cotranscription. To characterize the functional context of LivK, 39 proteins, including LivK, annotated in KEGG pathways pgn02010 (ABC transporters) or pgn02024 (quorum sensing) were selected. Proteins were classified according to the KEGG BRITE hierarchy (ko00001) and represented in a network constructed using NetworkX 3.6.1 and visualized with Cytoscape 3.10.5 [46]. Connections were established between proteins sharing at least one KEGG pathway, with edge weights corresponding to the number of shared pathways, without interpreting these connections as physical or regulatory interactions.

The analysis was considered descriptive. Overrepresentation tests were performed using a one-sided hypergeometric test with Benjamini–Hochberg correction [47]. The overrepresentation analyses were not used to support the study conclusions because the proteins were selected based on their membership in specific KEGG pathways. Consequently, their enrichment does not constitute independent functional evidence.

Reproducibility and Documentation Limitations

All calculations were performed in Ubuntu 24.04 under the Windows Subsystem for Linux, using Python 3.11.17 in dedicated conda environments. The documented software versions included pandas 3.1.0, NumPy 2.4.6, SciPy 1.17.1, and NetworkX 3.6.1. Input and processed structures, COCONUT identifiers, docking poses, fpocket outputs, analysis scripts, intermediate tables, network files, and SHA-256 checksums were preserved and organized by workflow stage.

 DOI: 10.26717/BJSTR.2026.67.010494

Results

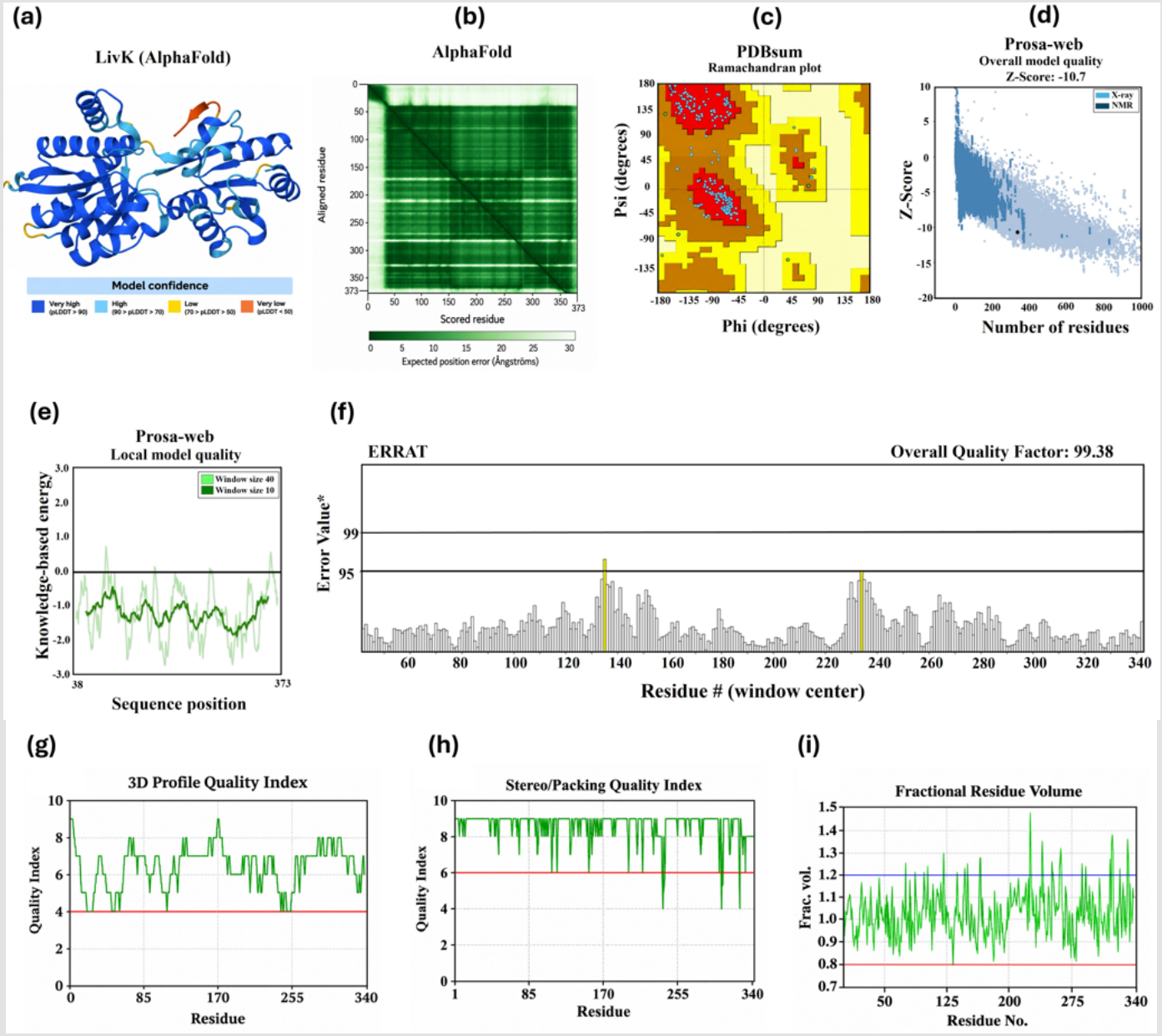


Figure 1: Structural model and quality assessment of LivK.

a) Three-dimensional structure predicted by AlphaFold, colored according to local model confidence (pLDDT).
b) Predicted Aligned Error (PAE) matrix.
c) Ramachandran plot generated using PDBsum.
d) Global and local structural quality assessments performed using ProSA-web.
e) Global and local structural quality assessments performed using ProSA-web.
f) ERRAT profile of nonbonded atomic interactions.
g) 3D Profile Quality Index; values above 4 indicate the acceptable range (red line).
h) Stereo/Packing Quality Index; values above 6 indicate the acceptable range (red line).
i) Fractional Residue Volume, with the acceptable range defined between 0.8 (red line) and 1.2 (blue line).

The Three-Dimensional Structural Model of LivK Is Suitable for Molecular Docking

The three-dimensional model of LivK obtained using AlphaFold exhibited a protein fold predominantly characterized by regions of high and very high structural confidence within the segment spanning residues 38–373 (Figure 1a). Analysis of the Predicted Aligned Error (PAE) matrix revealed low error values in regions corresponding to intradomain organization, supporting the consistency of the predicted structural architecture (Figure 1b). Notably, the residues comprising the consensus binding site exhibited a mean pLDDT value of 88.6, indicating high confidence in their local structural conformation. Ramachandran plot analysis showed that 98.3% of the evaluated residues (236/240) were located in favored conformational regions, while the remaining 1.7% (4/240) occupied allowed regions, with no outliers identified (Figure 1c). This distribution indicates favorable protein backbone geometry, with φ and ψ dihedral angles consistent with stereochemically acceptable conformations. ProSA-web analysis yielded an overall Z-score of −10.7, within the reference range of experimentally determined protein structures of comparable size. Furthermore, the local energy profile displayed predominantly negative values throughout the sequence, consistent with a favorable structural conformation (Figure 1d & Figure 1e). Similarly, ERRAT yielded an overall quality factor of 99.38, indicating a high degree of consistency in the model's nonbonded atomic interactions (Figure 1f).

Complementary structural quality assessments revealed predominantly favorable sequence–structure compatibility (3D Profile Quality Index) and stereochemical and packing quality (Stereo/Packing Quality Index), while fractional residue volume values (Fractional Residue Volume) remained largely within the reference limits (Figure 1g-Figure 1i). Although localized deviations were identified in certain regions, they did not indicate widespread structural abnormalities. Collectively, the agreement between AlphaFold structural confidence metrics and independent assessments of stereochemical quality, energy profiles, and molecular packing supports the global and local consistency of the three-dimensional LivK model and its suitability for binding-cavity characterization and comparative molecular docking analyses.

Prediction and Characterization of the LivK Binding Site

Binding-site analysis of LivK using P2Rank identified a candidate region comprising 13 residues, with a score of 3.17 and a predicted probability of 0.111 (11.1%). In parallel, geometric analysis using a LIGSITE-based cavity detection algorithm (1.0 Å grid) identified a primary cavity comprising 38 residues, which contained all positions identified by P2Rank (Table 1). This cavity had an approximate volume of 288 $Å^3$, larger than that of the second identified cavity (103 $Å^3$), while the distance between the centers predicted by the two methods was 4.58 Å. The overlap of 13 residues enabled the definition of a structural consensus core comprising Leu54, His136, Thr138, Ser153, Pro154, Glu159, Tyr183, Leu256, Gly257, Val260, Phe278, Ile304, and Ser305. For subsequent analyses, an expanded operational binding pocket comprising 16 residues was used by incorporating Thr155, Val156, and Asn232 as additional positions within the binding environment. This residue set exhibited a mean pLDDT of 88.8 and served as the reference for evaluating protein–ligand contacts during virtual screening. The expansion was considered an exploratory methodological definition rather than evidence of greater agreement between the prediction methods. As a complementary assessment, the experimental L-leucine-binding site of the *E. coli* homolog (PDB: 1USK) was examined. This site was defined by 12 residues located within 4.0 Å of the crystallographic ligand.

Table 1: Prediction of the substrate-binding pocket of LivK from P. gingivalis and comparison with the L-leucine-bound experimental homolog from E. coli (PDB: 1USK).

Protein	Method	Probability	No. of Residues	Pocket Center (x, y, z)	Residues Forming the Pocket
LivK (B2RKF2)	P2Rank 2.5.1 (PrankWeb)	0.111 (11.1%)	13	−1.06, −1.53, −4.38	Leu54, His136, Thr138, Ser153, Pro154, Glu159, Tyr183, Leu256, Gly257, Val260, Phe278, Ile304, Ser305
LivK (B2RKF2)	Geometric cavity detection (LIGSITE-like grid, 1.0 Å)	—	38	−4.34, −0.16, −7.27	Contains the 13 P2Rank residues (100% overlap)
LivK (B2RKF2)	Expanded operational binding pocket used for docking	—	16	2.50, 2.41, −5.58	Leu54, His136, Thr138, Ser153, Pro154, Thr155, Val156, Glu159, Tyr183, Asn232, Leu256, Gly257, Val260, Phe278, Ile304, Ser305
Experimental homolog (PDB: 1USK)	Crystallographic L-leucine-binding site (X-ray, 2.00 Å)	—	12	Defined by the ligand	Trp18, Leu77, Cys78, Ser79, Gly100, Ala101, Thr102, Tyr150, Tyr202, Glu226, Gly227, Tyr276

Note: Binding-pocket centers are expressed as Cartesian coordinates (x, y, z) in Å. The expanded operational binding pocket comprises 13 consensus residues and three additional exploratory residues selected for molecular docking, whereas the docking box center was determined independently from the substrate poses, as described in the Materials and Methods section. The experimental reference binding site (PDB: 1USK) was defined by residues located within 4.0 Å of the crystallographic L-leucine, using the residue numbering of the *E. coli* structure. —, not applicable.

After ligand removal, independent analysis using fpocket recovered 10 of the 12 experimental binding-site residues, corresponding to 83.3% coverage and a Jaccard index of 0.769 (Table 1). These findings supported the ability of the geometric procedure to identify an experimentally characterized amino acid-binding cavity, although they do not directly validate the docking poses or the binding affinities of ligands toward LivK. In contrast, applying fpocket to the LivK model identified 26 cavities, none of which met the predefined criterion for spatial correspondence with the docked L-leucine pose. The closest cavity encompassed only 1 of the ligand's 9 heavy atoms within ≤3.0 Å of an alpha sphere (11.1%) and 2 of 9 within ≤4.0 Å (22.2%), with its centroid located 6.53 Å from the docking center. These findings indicate limited agreement between fpocket and the proposed operational binding site of LivK. Accordingly, this site was retained as a candidate binding region supported by the integration of P2Rank, geometric cavity detection, and exploratory docking of the natural substrates, rather than as an experimentally validated functional binding site.

Structural Comparison and Superposition of LivK with Experimental Homologs

TM-align analysis revealed substantial similarity between the three-dimensional fold of LivK and the experimentally determined structures 1USK from *E. coli* and 3TD9 from *T. maritima*. Despite low sequence identity (15–16%), LivK exhibited TM-scores of 0.73 and 0.72, respectively, with structural coverage ranging from 74.7% to 78.0%. Furthermore, comparison of the experimental structures 1USK and 3TD9 revealed greater structural similarity, with a TM-score of 0.92 (Table 2). Three-dimensional superposition illustrated the overall conservation of the protein fold, along with local structural differences among the proteins (Figure 2). Collectively, these findings support the similarity of the three-dimensional architecture of LivK to that of both experimental homologs, despite their limited sequence identity, and provide a structural reference for investigating interactions with potential ligands and exploring possible functional modulation.

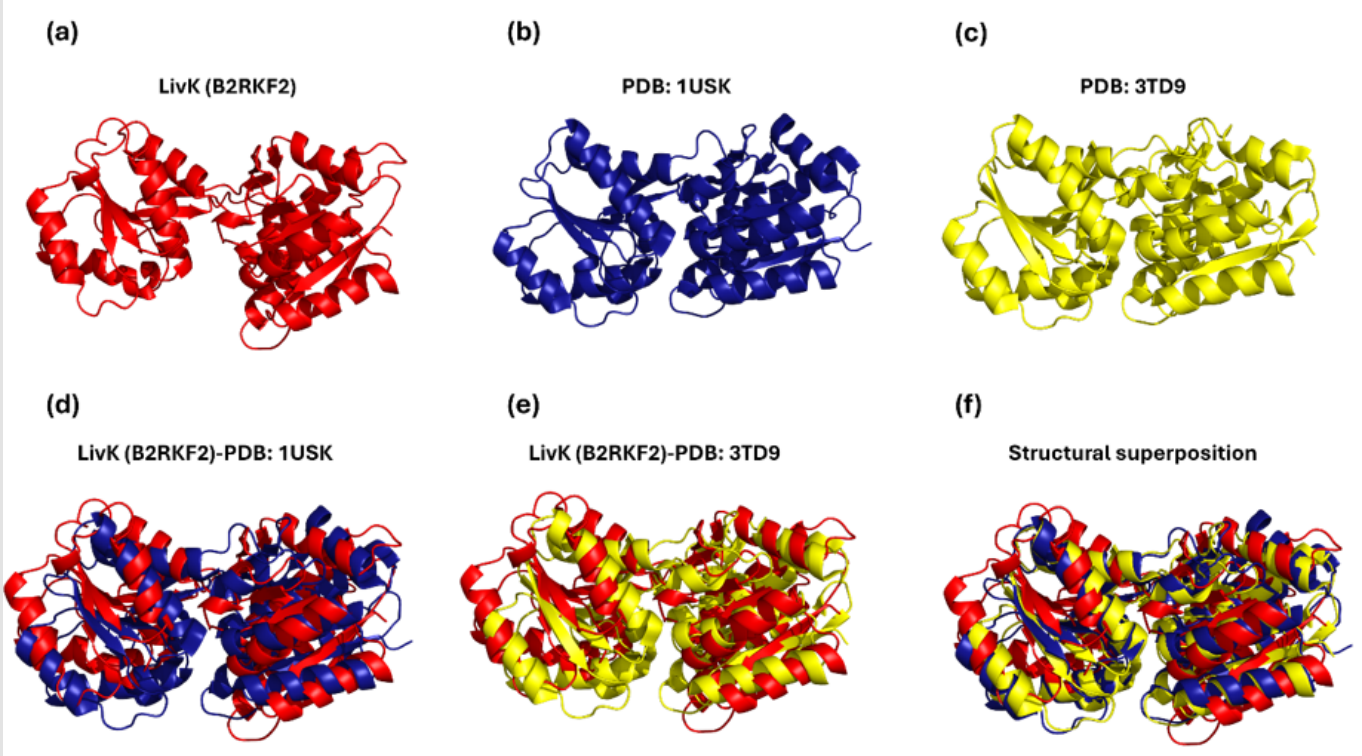


Figure 2: Structural comparison and superposition of LivK with experimentally determined homologs.

a) AlphaFold-predicted structure of LivK from *P. gingivalis* (red).

b) Crystal structure of the leucine-binding protein from E. coli (PDB: 1USK; blue).

c) Crystal structure of a branched-chain amino acid-binding protein from *T. maritima* (PDB: 3TD9; yellow).

d) Structural superposition of LivK and 1USK.

e) Structural superposition of LivK and 3TD9.

f) Combined superposition of the three structures, highlighting conservation of the overall fold and local structural differences. Structural similarity was quantitatively assessed using TM-align, whereas three-dimensional visualization and figure preparation were performed using PyMOL.

Table 2: Structural comparison of LivK with experimental proteins using TM-align.

Comparison	TM-Score	RMSD (Å)	Aligned Residues	Coverage (%)	Identity (%)
LivK–1USK	0.73	4.06	262	78	16
LivK–3TD9	0.72	4.16	251	74.7	15
1USK–3TD9	0.92	1.82	334	—	25

Note: Structural coverage for the comparisons with LivK was calculated relative to the 336 residues of its model (38–373). The TM-score represents a normalized measure of structural similarity. —, not applicable.

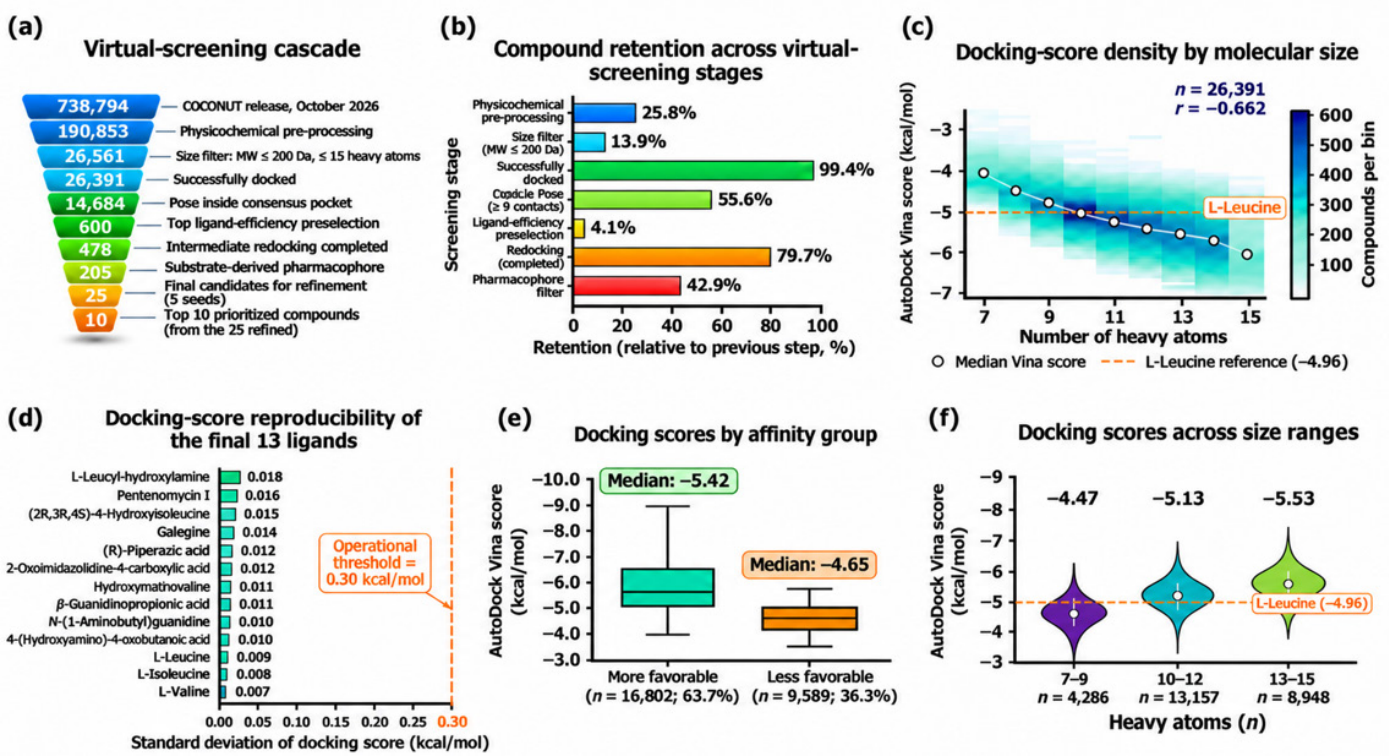


Figure 3: Virtual screening cascade and global analysis of docking scores for natural compounds against LivK.

a) Progressive reduction of the COCONUT library through physicochemical, geometric, ligand efficiency, and pharmacophore-based filters, followed by refinement and candidate prioritization.

b) Retention percentage relative to the previous stage, highlighting the most selective filters.

c) Density distribution of AutoDock Vina scores as a function of heavy-atom count, including the corresponding medians and the L-leucine reference.

d) Reproducibility of docking scores for the ten prioritized candidates and the three natural substrates, expressed as the standard deviation calculated from five independent random seeds; the dashed line indicates the operational threshold of 0.30 kcal/mol.

e) Box plots descriptively comparing the scores of compounds classified as more favorable than or equal to L-leucine and those classified as less favorable, without implying statistical significance.

f) Violin plots showing the distribution of docking scores across molecular size intervals, expressed as the number of heavy atoms. More negative scores indicate more favorable AutoDock Vina values and do not represent experimental measurements of binding affinity.

The Virtual Screening Cascade Integrated Geometric, Ligand Efficiency, and Pharmacophore-Based Criteria

Virtual screening of the COCONUT library, initially comprising 738,794 compounds, progressively reduced the chemical space through the sequential application of physicochemical, geometric, ligand efficiency, and pharmacophore-based criteria. Physicochemical preprocessing retained 190,853 structures, while molecular size filtering reduced the dataset to 26,561 compounds, of which 26,391 successfully completed molecular docking with AutoDock Vina. Subsequently, 14,684 compounds satisfied the criterion of establishing contacts with at least nine residues of the defined LivK binding site. Ligand efficiency-based preselection retained 600 candidates, of which 478 completed the intermediate redocking stage and 205

satisfied the pharmacophore-based filter derived from the natural substrates. From this subset, 25 candidates were selected for refinement using five independent random seeds, and the ten compounds with the most favorable mean AutoDock Vina scores were prioritized for subsequent comparison with the three natural substrates (Figure 3a). For the complementary analysis of contacts with acidic residues, 29 molecular poses were examined, corresponding to the 25 refined candidates, the three natural substrates, and one additional pose of the (2S,3R,4S)-4-hydroxy-L-isoleucine stereoisomer, independently evaluated for comparison with the (2R,3R,4S) isomer recorded in COCONUT.

This additional structure was not included in the primary set of 28 refined ligands. Analysis of these 29 poses revealed distances below 4.0 Å to Glu159 for all evaluated ligands (2.688–3.961 Å), whereas Asp181, Asp182, Asp277, and Glu306 showed no proximities within this threshold. These findings identified Glu159 as the only acidic residue satisfying the geometric criterion across the analyzed poses. Retention analysis identified the most selective stages of the screening workflow. Physicochemical preprocessing retained 25.8% of the initial library, whereas molecular size filtering retained 13.9% of the structures from the preceding stage. Docking was successfully completed for 99.4% of the prepared compounds, while the contact-based filter retained 55.6% of the docked structures. Ligand efficiency-based preselection represented the most restrictive stage, with a retention rate of 4.1%, followed by molecular size filtering (13.9%) and pharmacophore-based filtering, which retained 42.9% of the redocked candidates. Intermediate redocking was successfully completed for 79.7% of the selected compounds (Figure 3b). These results indicate that the progressive reduction of the library was primarily driven by the sequential application of molecular selection criteria, whereas losses associated with docking execution were comparatively smaller. Analysis of the 26,391 docked compounds revealed a negative correlation between the number of heavy atoms and AutoDock Vina scores ($r = -0.662$), with a tendency toward more negative scores as molecular size increased (Figure 3c).

Consistently, stratification by molecular size intervals showed a progressive shift toward more favorable score distributions among compounds containing greater numbers of heavy atoms (Figure 3f). This association highlights the influence of molecular size on the scoring function and the importance of incorporating additional parameters during candidate prioritization. Docking scores across the global screening dataset ranged from −7.37 to −2.49 kcal/mol, with a median of −5.16 kcal/mol. Using the L-leucine score (−4.96 kcal/mol) as a reference, 16,802 compounds (63.7%) exhibited more favorable scores, whereas 9,589 (36.3%) yielded scores equal to or less favorable than that of L-leucine. The median scores of these two groups were −5.42 and −4.65 kcal/mol, respectively (Figure 3e). Furthermore, refinement of the final panel of 13 ligands, comprising the ten prioritized candidates and three natural substrates, revealed low variability in docking scores across five independent random seeds, with standard deviations below the predefined operational threshold of 0.30 kcal/mol (Figure 3d). This finding reflects the energetic reproducibility of the search procedure but does not, by itself, demonstrate geometric convergence of docking poses or predictive accuracy, as discussed in greater detail in Section 3.6.

Collectively, these results demonstrate that, although a substantial proportion of the evaluated compounds exhibited more favorable docking scores than L-leucine, the observed association with molecular size prevents these differences from being interpreted as sufficient evidence of selective recognition by LivK. Accordingly, candidate prioritization integrated binding-site contact criteria, ligand efficiency, pharmacophore compatibility, and docking score reproducibility during refinement, providing a more comprehensive assessment than one based exclusively on AutoDock Vina scores.

Discriminatory Ability of Molecular Docking

Evaluation of 42 amino acids and analogs revealed limited discriminatory ability of AutoDock Vina scores. The natural substrates L-leucine, L-isoleucine, and L-valine exhibited docking scores of −4.96, −4.84, and −4.50 kcal/mol, respectively, whereas 10 of the 39 comparator compounds achieved more favorable scores than L-leucine. Among these, L-tryptophan (−6.44 kcal/mol) and L-phenylalanine (−6.08 kcal/mol), both classified as negative controls, exhibited particularly favorable scores. These findings demonstrate that docking scores alone cannot conclusively discriminate natural substrates from other amino acids and structural analogs, highlighting the need to consider geometric and pharmacophore-based criteria when interpreting molecular recognition (Supplementary Table S1). The scores obtained for this comparative series were generated using an independent molecular docking protocol with an 18 × 18 × 18 Å search box and an exhaustiveness of 64, whereas the final ligand panel was evaluated using a 15 × 15 × 15 Å search box and five independent random seeds. Accordingly, L-valine exhibited a docking score of −4.50 kcal/mol in the discriminatory analysis and a mean score of −4.49 kcal/mol during final refinement. Although this difference of 0.01 kcal/mol is small, both results should be interpreted within the context of their respective protocols rather than as directly equivalent measurements.

Prioritized Ligands Occupied a Common LivK Subsite and Exhibited Stable Docking Scores Across Independent Runs

Structural visualization of LivK revealed a defined binding cavity on the protein surface, within which the docking poses of the ten prioritized candidates and three natural substrates were concentrated (Figure 4a). An enlarged view of the predicted binding pocket illustrated the arrangement of the ligands within this region (Figure 4b), while their superposition revealed a compact spatial cluster consistent with occupancy of a common binding subsite (Figure 4c). The mean distance of individual ligand centroids from the common

centroid was 0.51 ± 0.24 Å, with a maximum deviation of 1.05 Å and a maximum pairwise centroid separation of 1.53 Å. All individual distances remained below the predefined operational threshold of 2.0 Å, indicating limited spatial dispersion of the final docking poses within the evaluated binding pocket. Comparison of the 13 ligands showed that the ten prioritized candidates exhibited mean AutoDock Vina scores ranging from −5.71 to −5.06 kcal/mol, which were more negative than those obtained for L-leucine (−4.96 kcal/mol), used as the internal operational threshold, as well as L-isoleucine (−4.84 kcal/mol) and L-valine (−4.49 kcal/mol). The most negative scores were obtained for L-leucyl hydroxylamine (−5.71 kcal/mol), pentenomycin I (−5.66 kcal/mol), and (2R,3R,4S)-4-hydroxyisoleucine (−5.65 kcal/mol), followed by galegine (−5.51 kcal/mol) and (R)-piperazic acid (−5.32 kcal/mol) (Figure 4d).

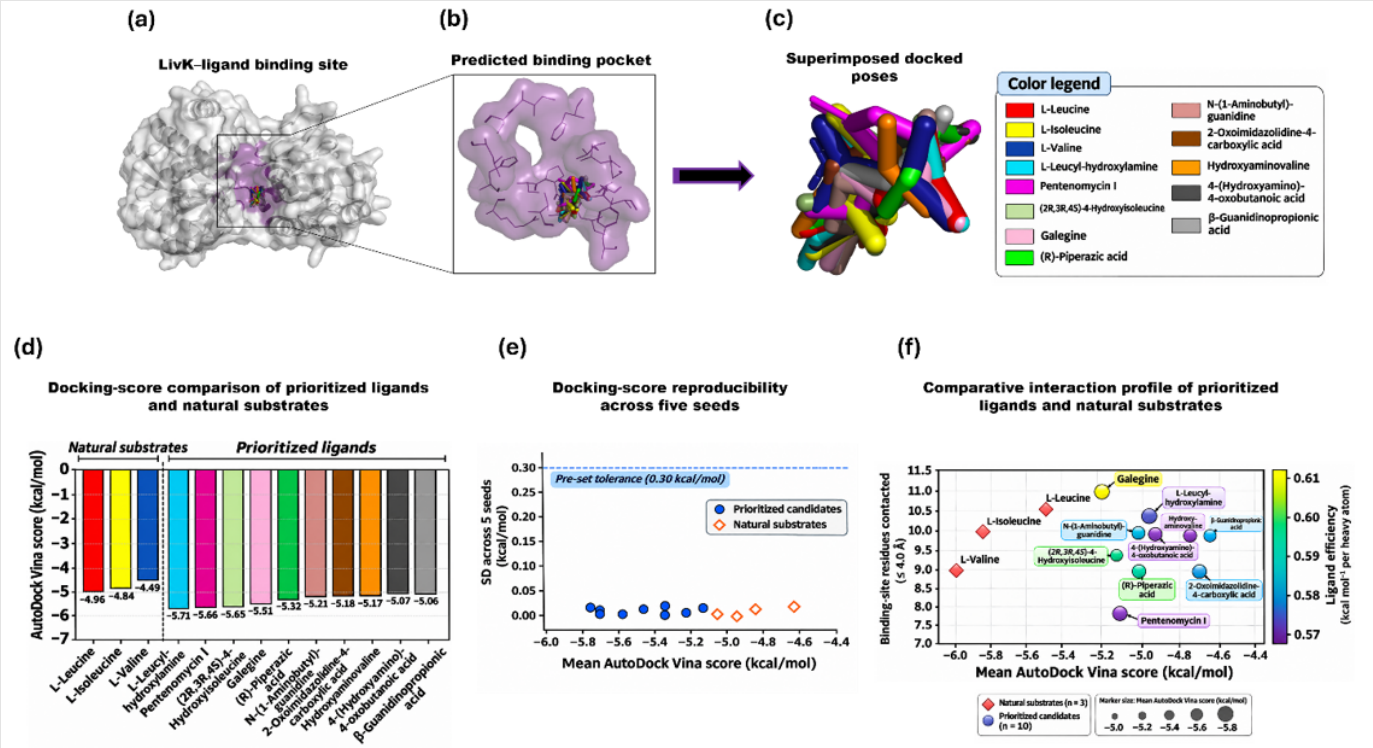


Figure 4: Structural characterization and comparative assessment of the ten prioritized ligands and three natural substrates of LivK.

a) Surface representation of LivK showing the location of the binding region.

b) Enlarged view of the predicted binding pocket and the spatial arrangement of the ligands.

c) Superposition of the docked poses of the ten prioritized candidates and three natural substrates, identified by a color code.

d) Comparison of the mean AutoDock Vina scores of the 13 ligands.

e) Energetic reproducibility of refinement using five independent random seeds, expressed as the standard deviation of docking scores around their mean; the dashed line indicates the operational threshold of 0.30 kcal/mol.

f) Comparative profile of docking scores and binding-site contact metrics, with ligand efficiency represented by a color scale. More negative scores indicate more favorable values according to AutoDock Vina and do not represent experimental measurements of binding affinity.

Stereochemical verification confirmed that CNP0207254.7 corresponds to the (2R,3R,4S) isomer, whose configuration at the α-carbon differs from that of (2S,3R,4S)-4-hydroxy-L-isoleucine described as a natural product. Additional docking of the latter stereoisomer under equivalent computational conditions yielded a score of −5.41 kcal/mol, representing a difference of 0.24 kcal/mol relative to the compound recorded in COCONUT. Refinement using five independent random seeds revealed low energetic variability across runs, with a maximum standard deviation of 0.018 kcal/mol for the 13 ligands, substantially below the predefined operational threshold of 0.30 kcal/mol (Figure 4e). This variability reflects convergence of the search algorithm under a common scoring function and provides an indicator of procedural stability across different initializations rather than predictive accuracy. In this regard, the differences observed between the prioritized candidates and L-leucine (0.10–0.75 kcal/mol) are smaller than the characteristic error of the AutoDock Vina scoring function relative to experimental binding affinities [39,40]. Therefore, docking scores were used for internal ranking and comparison against the natural substrate threshold, without interpreting them as binding affinity estimates or evidence of stronger binding than L-leucine. This consideration should be interpreted alongside the limited discriminatory ability documented in Section 3.5. Normalization of docking scores by heavy-atom count altered the relative ranking of the compounds (Figure 4f).

The ten prioritized candidates exhibited ligand efficiencies ranging from 0.56 to 0.61 kcal/mol per heavy atom, whereas the three natural substrates showed values between 0.54 and 0.56 kcal/mol per heavy atom. The ranges were closely aligned at their boundaries, indicating that these differences do not establish a conclusive energetic ranking. Galegine exhibited the highest ligand efficiency (0.61), despite ranking fourth among the prioritized candidates based on its AutoDock Vina score (−5.51 kcal/mol). In contrast, L-leucyl hydroxylamine, which exhibited the most negative docking score (−5.71 kcal/mol), had a ligand efficiency of 0.57. These differences demonstrate that ligand efficiency normalized by molecular size provides complementary information to absolute docking scores. Collectively, occupancy of a common binding subsite, stable docking scores across independent runs, and agreement among geometric, pharmacophore-based, and structural diversity criteria support the computational prioritization of the ten candidates for further characterization. This prioritization is not based on the magnitude of the docking score differences, which fall within the uncertainty of the scoring method, nor does it demonstrate inhibitory activity or selective recognition by LivK. Accordingly, its interpretation requires integration with detailed molecular interaction analyses and subsequent experimental validation.

Evaluated Ligands Shared a Recurrent Core of Contacts within the LivK Binding Pocket

Geometric proximity analysis, restricted to a set of 17 binding-pocket residues and using a distance threshold of less than 4.0 Å between heavy atoms, identified seven residues consistently present in the contact profiles of all 13 ligands: Thr138, Arg151, Ser153, Pro154, Glu159, Tyr183, and Gly257. Thr155, Val156, Leu256, and Val260 were identified in 11 of the 13 compounds, whereas Asn232 was observed in four. Two-dimensional interaction profiles generated using BIOVIA Discovery Studio Visualizer, which applies interaction classification criteria distinct from the geometric distance threshold described above, allowed differentiation of the predicted interaction types in each complex (Figure 5). In the three natural substrates, conventional hydrogen bonds involved the α-amino and carboxylate groups, with recurrent participation of Thr138, Arg151, Ser153, and Pro154, as well as Glu159 in the case of L-valine. Candidates retaining an α-amino acid backbone—(2R,3R,4S)-4-hydroxyisoleucine, hydroxyaminovaline, and L-leucyl hydroxylamine—exhibited polar interactions with a partially overlapping set of residues, although their geometric equivalence was not evaluated. Among guanidino-containing compounds—galegine, N-(1-aminobutyl) guanidine, and β-guanidinopropionic acid—the predicted hydrogen bonds primarily involved Thr138, Pro154, and Glu159. Interpretation of these interactions depends on the protonation states assigned during molecular preparation.

DOI: 10.26717/BJSTR.2026.67.010494

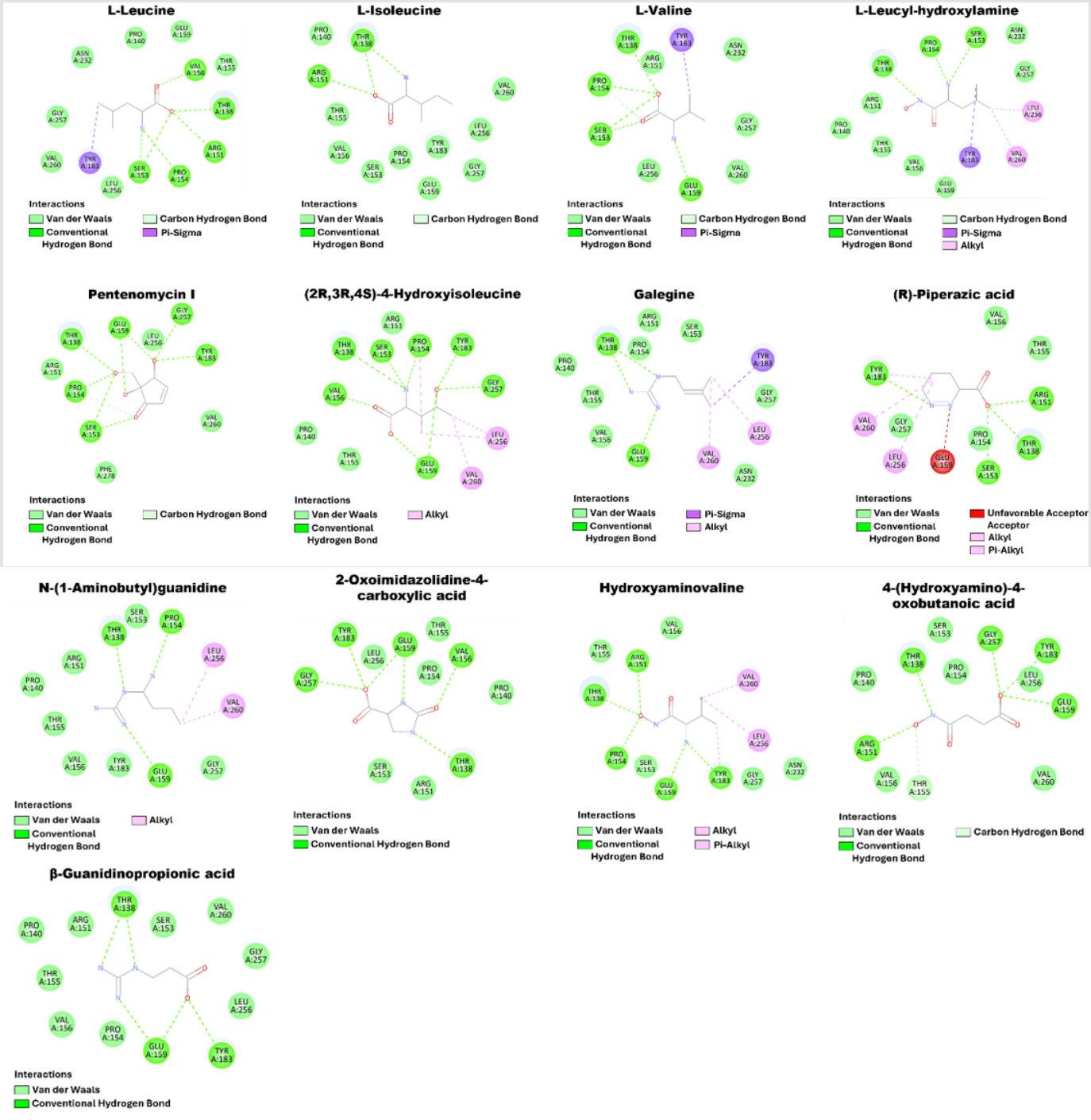


Figure 5: Two-dimensional interaction profiles of the natural substrates and ten prioritized candidates within the LivK binding pocket. Diagrams were generated using BIOVIA Discovery Studio Visualizer from the first docking pose (pose 1), corresponding to the conformation with the most favorable AutoDock Vina score for each ligand. The diagrams illustrate residues in proximity to each ligand and the predicted interactions, including hydrogen bonds, hydrophobic contacts, van der Waals interactions, and unfavorable contacts, according to the software's conventions.

Hydrophobic interactions classified as alkyl or π-alkyl primarily involved Leu256 and Val260 and were identified in the interaction profiles of L-leucyl hydroxylamine, (2R,3R,4S)-4-hydroxyisoleucine, galegine, hydroxyaminovaline, N-(1-aminobutyl)guanidine, and (R)-piperazic acid. Tyr183 participated in π–sigma interactions with L-leucine, L-valine, L-leucyl hydroxylamine, and galegine, whereas van der Waals contacts were identified in other profiles. Additionally, the interaction diagram of pentenomycin I indicated the involvement of Phe278, a residue that did not satisfy the geometric proximity criterion of less than 4.0 Å applied in the independent analysis. In the interaction profile of (R)-piperazic acid, Discovery Studio classified a contact between the ligand's carboxylate group and Glu159 as an unfavorable acceptor–acceptor interaction. This classification provides information complementary to the docking score because the AutoDock Vina scoring function incorporates steric, hydrophobic, and hydrogen-bonding terms but does not include a specific acceptor–acceptor repulsion term equivalent to the category used by Discovery Studio. Therefore, this contact represents a local interaction feature that warrants further examination during subsequent characterization of the candidate, without independently demonstrating complex instability or lack of binding. Overall, the two-dimensional interaction profiles revealed a shared core of binding-site residues alongside differences in the nature of the predicted interactions.

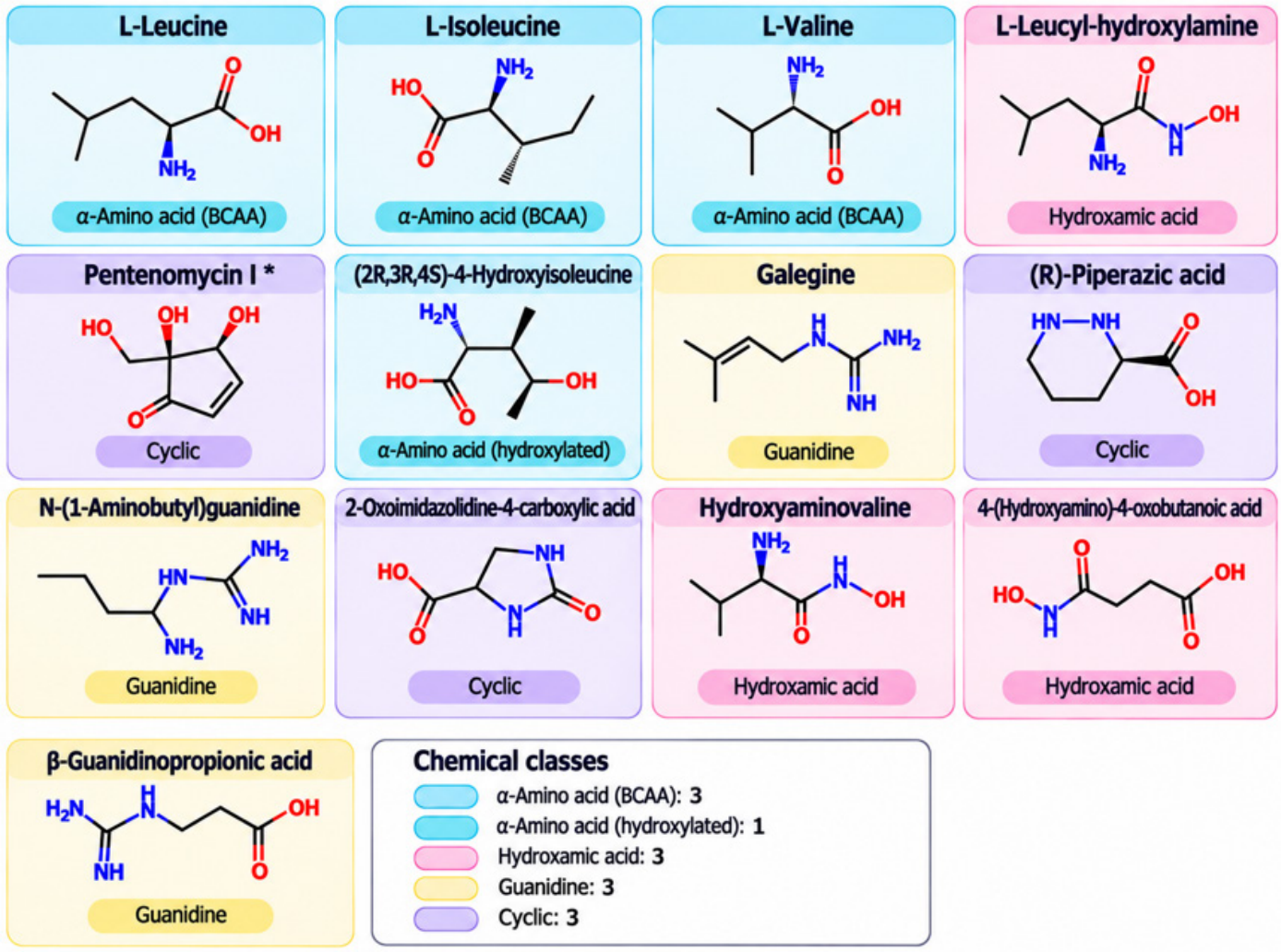


Figure 6: Structural classification of the natural substrates and prioritized candidates for LivK. Two-dimensional representations of the chemical structures of the three natural substrates and ten candidates selected through virtual screening. The compounds were organized into five visual categories corresponding to four major structural groups: branched-chain α-amino acids (BCAAs, n = 3), hydroxylated α-amino acids (n = 1), hydroxamic acid derivatives (n = 3), guanidino-containing compounds (n = 3), and cyclic structures (n = 3). Background colors indicate the structural category assigned to each compound. Pentenomycin I (*) was the only compound in the set lacking nitrogen atoms. The (2R,3R,4S)-4-hydroxyisoleucine configuration corresponds to the COCONUT stereoisomer evaluated in this study (CNP0207254.7), which differs from (2S,3R,4S)-4-hydroxy-L-isoleucine described as a natural product.

Selected Ligands Represent Four Chemotypes and Required Structural Identity Verification

Analysis of the chemical structures of the three natural substrates and ten prioritized candidates revealed four major structural groups (Figure 6). The first group comprised branched-chain α-amino acids, represented by L-leucine, L-isoleucine, and L-valine, together with (2R,3R,4S)-4-hydroxyisoleucine, which retains the amino acid backbone with an additional hydroxyl group at C4. The second group included three compounds containing hydroxamic acid functionalities: L-leucyl hydroxylamine, hydroxyaminovaline, and 4-(hydroxyamino)-4-oxobutanoic acid. The third group consisted of guanidino-containing compounds: galegine, N-(1-aminobutyl) guanidine, and β-guanidinopropionic acid. Finally, the fourth group comprised three chemically distinct cyclic structures: pentenomycin I, a polyhydroxylated cyclopentenone; (R)-piperazic acid, a heterocyclic compound containing an N–N bond; and 2-oxoimidazolidine-4-carboxylic acid, a cyclic urea derivative. Two characteristics of the ligand set were particularly relevant for subsequent characterization. Pentenomycin I was the only ligand lacking nitrogen atoms, unlike the natural substrates and the remaining prioritized candidates. Additionally, the ten candidates exhibited consensus partition coefficients (cLogP) ranging from −1.33 to 0.54, with galegine being the only compound with a positive value. The predominance of negative values reflects the polar nature of the compound set and provides a physicochemical context for interpreting the ADME predictions.

Structural verification of the COCONUT records identified two cases requiring clarification of chemical identity. Record CNP0207254.7 corresponded to (2R,3R,4S)-4-hydroxyisoleucine, whose configuration was confirmed using its InChIKey. This compound is a diastereomer of (2S,3R,4S)-4-hydroxy-L-isoleucine described in *Trigonella foenum-graecum*, differing in the configuration of the α-carbon. Accordingly, both structures were treated as distinct stereochemical entities, and the explicit nomenclature of the evaluated compound was retained throughout the manuscript figures and tables. Similarly, record CNP0547045.0 corresponded to N-(1-aminobutyl)guanidine rather than agmatine, although both molecules share the molecular formula $C_5H_{14}N_4$. Structural comparison confirmed that these compounds are constitutional isomers with different InChIKeys. In CNP0547045.0, the amino group and the nitrogen attached to the guanidino group are connected to the same carbon atom of the chain, forming a geminal aminal moiety, whereas in agmatine, these functional groups are located at opposite ends of the carbon chain. Additionally, two independent agmatine records were identified during primary screening (CNP0307539.0 and CNP0572375.0), with docking scores of −4.54 and −4.52 kcal/mol, respectively, which were less favorable than the L-leucine reference (−4.96 kcal/mol). Consequently, these compounds did not advance to subsequent screening stages.

The carbon atom bearing the two nitrogen-containing substituents in CNP0547045.0 constitutes a stereogenic center whose configuration was not specified in the original record. This characteristic was considered during molecular preparation and evaluation of the stereochemical configurations, according to the procedure described in the Materials and Methods section.

Integration of ADME Properties and Toxicity Predictions Enabled Differentiation of Candidate Priorities

Integration of physicochemical properties, Brenk structural alerts, and toxicity predictions enabled the establishment of an exploratory classification of the ten candidates selected for LivK (Table 3). This assessment did not alter the selection obtained through virtual screening but instead helped establish priorities for subsequent characterization. Three categories were defined: high priority, corresponding to compounds without Brenk alerts or active ProTox 3.0 predictions other than cardiotoxicity; medium priority, comprising compounds with one or more Brenk alerts but no other active toxicity predictions, or only a cardiotoxicity signal; and low priority, corresponding to compounds with at least one active toxicity prediction other than cardiotoxicity. In all cases, predictions classified as active with probabilities ≥0.70 were considered. Application of these criteria identified four high-priority, five medium-priority, and one low-priority candidate. The high-priority candidates were pentenomycin I, (2R,3R,4S)-4-hydroxyisoleucine, (R)-piperazic acid, and 2-oxoimidazolidine-4-carboxylic acid, none of which exhibited Brenk alerts or active toxicity predictions other than cardiotoxicity above the established threshold. In contrast, L-leucyl hydroxylamine, which achieved the most favorable docking score, exhibited two Brenk alerts and was assigned medium priority. Similarly, galegine, which showed the highest ligand efficiency, exhibited three alerts and was also classified as medium priority. These differences demonstrate that incorporating medicinal chemistry criteria altered the interpretation of the candidates relative to ranking based exclusively on docking scores.

Table 3: Integrated docking profile, physicochemical properties, medicinal chemistry alerts, and exploratory prioritization of LivK ligands.

Compound	COCONUT ID	Mean VINA SCORE (kcal/mol)	LE	Consensus cLogP	Brenk ALERTs (n)	ProTox 3.0 Active (P ≥ 0.70)	Exploratory Priority
L-Leucine	—	−4.96 ± 0.00	0.55	−0.30	0	Cardiotoxicity (0.95)	Reference
L-Isoleucine	—	−4.84 ± 0.01	0.54	−0.31	0	Cardiotoxicity (0.82)	Reference
L-Valine	—	−4.49 ± 0.02	0.56	−0.77	0	Cardiotoxicity (0.94)	Reference
L-Leucyl hydroxylamine	CNP0527266.1	−5.71 ± 0.01	0.57	−0.04	2	—	Medium
Pentenomycin I	CNP0256144.1	−5.66 ± 0.01	0.57	−0.92	0	—	High
(2R,3R,4S)-4-Hydroxyisoleucine	CNP0207254.7	−5.65 ± 0.00	0.57	−1.33	0	—	High
Galegine	CNP0549417.0	−5.51 ± 0.00	0.61	0.54	3	—	Medium
(R)-Piperazic acid	CNP0122625.1	−5.32 ± 0.01	0.59	−1.30	0	—	High
N-(1-Aminobutyl)guanidine	CNP0547045.0	−5.21 ± 0.01	0.58	−0.43	3	—	Medium
2-Oxoimidazolidine-4-carboxylic acid	CNP0324165.0	−5.18 ± 0.01	0.58	−1.09	0	—	High
Hydroxyaminovaline	CNP0510615.1	−5.17 ± 0.01	0.57	−0.26	2	—	Medium
4-(Hydroxyamino)-4-oxobutanoic acid	CNP0136196.0	−5.07 ± 0.00	0.56	−0.77	2	Mutagenicity (0.71)	Low
β-Guanidinopropionic acid	CNP0480617.0	−5.06 ± 0.02	0.56	−0.98	2	Cardiotoxicity (0.78)	Medium

Note: Vina score, mean AutoDock Vina score ± standard deviation of five independent runs (kcal/mol); LE, ligand efficiency (kcal mol^{-1} per heavy atom); cLogP, consensus octanol–water partition coefficient predicted by SwissADME; Brenk alerts, number of structural alerts; ProTox 3.0, toxicity prediction tool; P, predicted probability. Only Active predictions with P ≥ 0.70 are reported; (—) indicates that none reached this threshold, not the absence of toxicity.

The distribution of Brenk alerts was associated with chemotype: all six candidates exhibiting alerts belonged to the hydroxamic acid or guanidino-containing compound classes, whereas none of the four high-priority candidates presented such alerts. This association indicates that the classification was influenced by the structural characteristics of the compounds. However, Brenk alerts identify potentially problematic structural motifs for chemical screening and do not constitute experimental evidence of toxicity, particularly when the flagged functional groups are also present in endogenous metabolites and pharmacologically active compounds. 4-(Hydroxyamino)-4-oxobutanoic acid was the only low-priority candidate because of an active mutagenicity prediction (probability = 0.71). β-Guanidinopropionic acid retained medium-priority status despite exhibiting a cardiotoxicity signal (probability = 0.78). Cardiotoxicity was not considered a discriminatory criterion because ProTox 3.0 also classified all three natural substrates as cardiotoxic, with probabilities of 0.95, 0.82, and 0.94 for L-leucine, L-isoleucine, and L-valine, respectively, limiting its comparative utility within this compound series. Structural verification introduced two considerations independent of the ADME and toxicity predictions. (2R,3R,4S)-4-Hydroxyisoleucine retained high-priority status despite corresponding to a stereoisomer different from naturally occurring 4-hydroxy-L-isoleucine.

Similarly, N-(1-aminobutyl)guanidine remained in the medium-priority category; however, its geminal aminal moiety and unspecified stereogenic center require verification of chemical identity and stability before experimental evaluation. These considerations did not alter the exploratory classification (Table 3). SwissADME predictions were relatively homogeneous across the 13 ligands, indicating high gastrointestinal absorption, no predicted blood–brain barrier permeation, no predicted P-glycoprotein substrate activity, no predicted inhibition of the five evaluated CYP isoenzymes, and no Lipinski violations or PAINS alerts. However, these parameters cannot establish whether the compounds can reach the periplasmic space of *P. gingivalis* and showed limited discriminatory ability within the evaluated set. Consequently, they were not included in the main table or used as selection or exclusion criteria. Collectively, integration of these predictions enabled differentiation of candidate priorities for further characterization and identification of chemical limitations not directly reflected in docking scores. Nevertheless, the chemotype dependence of Brenk alerts and the limitations of the predictive models prevent this classification from being considered a validated assessment of safety or efficacy. These findings provide a preliminary basis for guiding the experimental selection of candidates, subject to confirmation of their chemical identity, stability, and biological activity.3.10.

LivK Is Integrated into a Functional Network Primarily Associated with Bacterial Transport Integration of GenBank and KEGG annotations for *P. gingivalis* ATCC 33277 identified LivK as the product of gene PGN_1328 (BAG33847.1; K01999), located within a genomic neighborhood associated with components of the LIV transport system (PGN_1324–PGN_1330). Based on shared annotations, a set of 39 proteins functionally associated with LivK was identified. Network visualization using Cytoscape revealed modules associated with

branched-chain amino acid (BCAA) transport, membrane transporters and ABC transport systems, Sec/SRP-mediated protein export, and other cellular processes (Figure 7a). Network connections represent functional associations derived from shared annotations rather than direct physical protein–protein interactions. The functional composition of this set was dominated by proteins involved in bacterial transport: 33 of the 39 proteins were associated with membrane transport categories, and 24 were annotated as ABC transporters (Figure 7b). Additionally, 20 proteins shared annotations with the KEGG quorum sensing pathway, primarily because of the presence of bacterial transport components, without demonstrating direct involvement of LivK in bacterial communication. Categories related to protein export, bacterial secretion, and protein folding, sorting, and degradation were also identified.

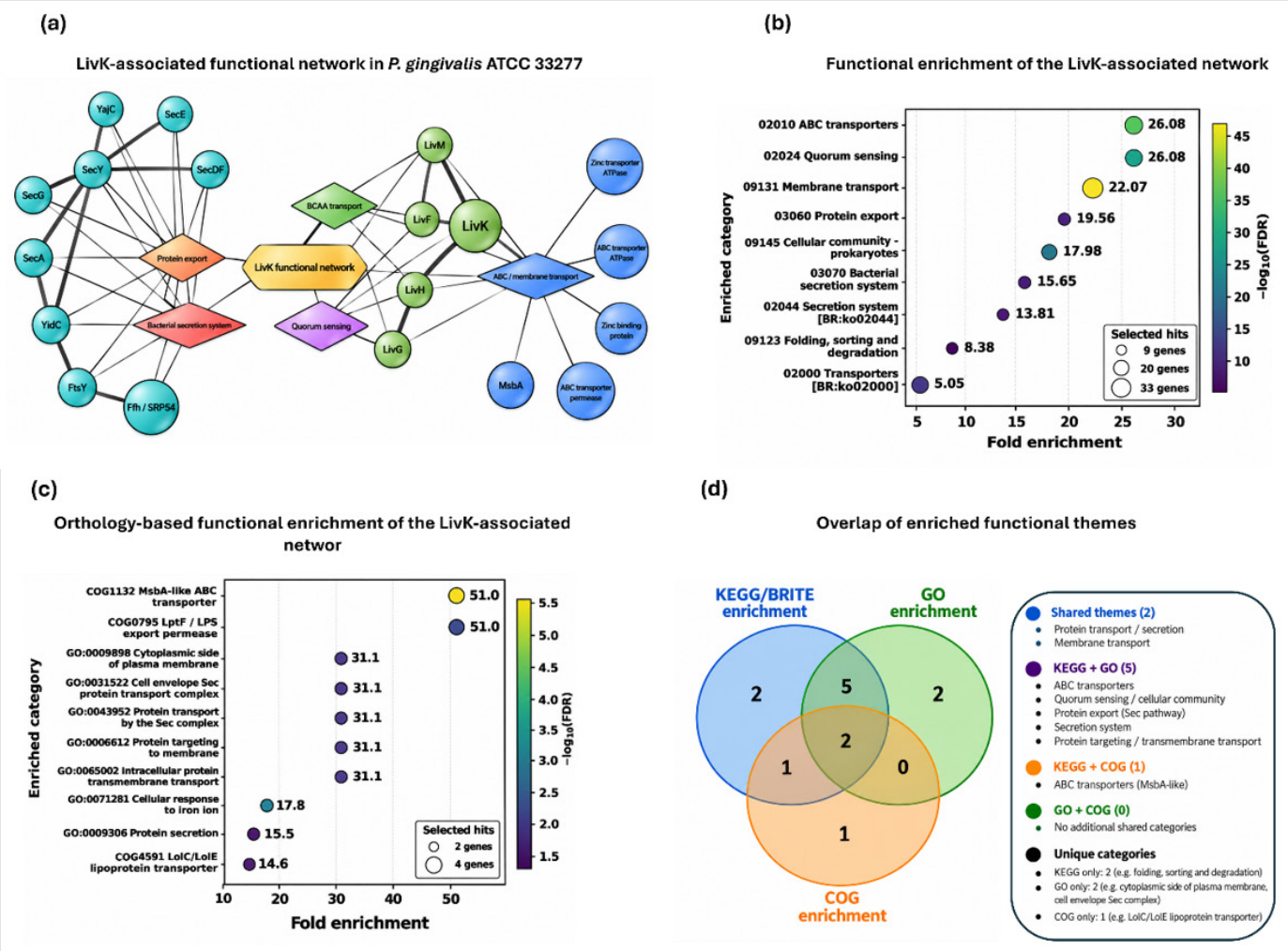


Figure 7: Genomic and functional context of LivK in *P. gingivalis* ATCC 33277.

a) Functional network of 39 proteins visualized using Cytoscape, organized into modules associated with LIV/BCAA transport, membrane/ABC transporters, Sec/SRP-mediated protein export, and related processes. Node shapes distinguish proteins, functional categories, and the central node; connections represent associations based on shared annotations rather than direct physical interactions.

b) KEGG/BRITE functional enrichment of the selected protein set.

c) Enrichment of Gene Ontology (GO) and Clusters of Orthologous Groups (COG) categories, including functions related to transport and protein export. In panels (b) and (c), bubble size represents the number of proteins, whereas color indicates $-\log_{10}$(FDR), with independent scales.

d) Venn diagram illustrating the overlap of functional themes among KEGG/BRITE, GO, and COG. Overrepresentation analyses were performed using a one-sided hypergeometric test with Benjamini–Hochberg correction. KEGG/BRITE results should be interpreted considering that the proteins were initially selected based on KEGG annotations shared with LivK.

Because the proteins were selected based on KEGG annotations shared with LivK, these findings describe the functional composition of the selected set rather than providing independent validation of biological associations. Complementary analysis using Gene Ontology (GO) and Clusters of Orthologous Groups (COG) classifications identified categories associated with membrane transport, protein export, and bacterial cell envelope organization (Figure 7c). Among the COG categories, COG1132, corresponding to MsbA-type ABC transporters, and COG0795, associated with lipopolysaccharide export permeases, were particularly prominent, both exhibiting enrichment factors of 51.0. COG4591, associated with LolC/LolE lipoprotein transporters, was also identified. GO annotations included transport processes, protein targeting to membranes, components of the Sec system, and cellular response to iron ions (GO:0071281), providing additional functional characterization of the selected protein set. Comparison among KEGG/BRITE, GO, and COG revealed partial convergence of functional categories (Figure 7d). Two functional themes were shared across all three classification systems, corresponding to membrane transport and protein secretion or transport. Additionally, five themes were shared between KEGG and GO, one between KEGG and COG, and none exclusively between GO and COG.

Despite differences in classification frameworks, these findings showed that the annotations converged primarily on functions associated with transport and the bacterial cell envelope. Nevertheless, this comparison should be interpreted in light of differences in coverage and functional organization among the annotation systems. Collectively, these findings place LivK within a functional context predominantly associated with membrane and amino acid transport, consistent with its annotation as a periplasmic substrate-binding protein of the LIV system. The convergence of transport-related categories across different annotation frameworks provides complementary functional context for interpreting the biological relevance of LivK, without demonstrating that the prioritized compounds alter its transport activity or the physiology of *P. gingivalis*.

Discussion

Periodontitis is a chronic inflammatory disease associated with subgingival biofilm dysbiosis, in which *P. gingivalis* plays a significant role through its ability to reshape microbial communities and alter host immune responses, even at low abundance [48]. The persistence of this bacterium depends, in part, on its adaptation to the ecological and nutritional conditions of the periodontal pocket. Owing to its asaccharolytic metabolism, *P. gingivalis* relies primarily on peptides and amino acids as carbon and energy sources [11,12]. Accordingly, nutrient acquisition and utilization mechanisms represent processes of particular interest for understanding its physiology and survival. This metabolic dependence is reflected in its specialized proteolytic system, comprising gingipains and dipeptidyl peptidases involved in protein and peptide degradation and processing [6,11,13,49]. Madej et al. [14] demonstrated that RagAB functions as a selective olig op eptide uptake system across the outer membrane, providing structural and functional evidence of nutrient acquisition in *P. gingivalis*. However, whereas proteolytic processing and olig op eptide uptake have been extensively characterized experimentally, the specific recognition of free amino acids by periplasmic proteins remains less well understood. This distinction highlights a complementary aspect of bacterial physiology that warrants further investigation.

The importance of these processes should also be considered in relation to the ability of *P. gingivalis* to interact with extraoral tissues. Farrugia, et al. [50] demonstrated that gingipains facilitate endothelial cell invasion and compromise vascular barrier integrity by reducing PECAM-1 and VE-cadherin levels. Furthermore, Wu, et al. [15] identified increased expression of the *livh* and *livk* genes during endothelial cell infection and observed that *livk* deletion reduced cellular invasion efficiency and hepatic colonization. Similarly, strains deficient in *livh* or *livk* exhibited a reduced ability to increase circulating BCAA levels and exacerbate liver injury in mice fed a high-fat diet. These findings suggest an association between the LIV system and specific infection-related phenotypes, although they do not establish that LivK directly participates in BCAA biosynthesis or export, nor do they demonstrate its specific role during periodontal colonization. Within this context, our study investigated the molecular characteristics of LivK, annotated as a periplasmic substrate-binding protein of the LivKHGMF ABC transport system. Structural comparison with the experimentally determined homologs 1USK and 3TD9 revealed conservation of the three-dimensional protein fold, while the integration of P2Rank and geometric approaches enabled the identification of a candidate binding region compatible with branched-chain amino acid recognition.

Although this agreement provides structural support for the proposed binding-site location, the absence of an experimentally determined LivK structure and the potential involvement of conformational changes during substrate binding limit the precision with which its functional binding pocket can be defined. Virtual screening enabled the prioritization of natural compounds by integrating docking scores, ligand efficiency, geometric criteria, and pharmacophore features. The observed association between molecular size and AutoDock Vina scores further emphasized the importance of avoiding reliance on absolute docking scores alone. This limitation was also evident in the assessment of the discriminatory ability of molecular docking, in which L-tryptophan (−6.44 kcal/mol) and L-phenylalanine (−6.08 kcal/mol), included as negative controls, exhibited more favorable scores than the ten prioritized candidates and three natural substrates. These findings indicate that AutoDock Vina scores, when considered independently, cannot establish selective recognition by LivK. Consequently, the L-leucine docking score was used exclusively as an operational threshold for internal ranking rather than as evidence of binding affinity or specificity.

Candidate prioritization was therefore based on the integration of geometric criteria, ligand efficiency, and pharmacophore compatibil-

ity, although these parameters alone cannot confirm selective recognition of the compounds. The selected candidates exhibited recurrent contacts with residues including Thr138, Arg151, Ser153, Pro154, Glu159, Tyr183, and Gly257, identifying potential molecular recognition sites for future investigation. Nevertheless, predicted occupancy of a binding cavity cannot distinguish between a transportable substrate, a nontransported ligand, and a potential competitor. Similarly, the stereochemical differences identified during chemical verification and the ADME/toxicity predictions require careful interpretation, particularly because human pharmacokinetic models cannot directly establish whether the compounds can reach the bacterial periplasmic space. Pharmacological targeting of functions associated with bacterial metabolism has relevant experimental precedents. Aoki-Nonaka, et al. [51] demonstrated that KGDI-109, an inhibitor of dipeptidyl peptidase 7 (DPP7), exhibits antibacterial activity against *P. gingivalis* and reduces both its oral bacterial burden and alveolar bone loss in a murine model of periodontitis. This finding provides preclinical evidence that certain functions associated with nutrient processing may represent targets for intervention.

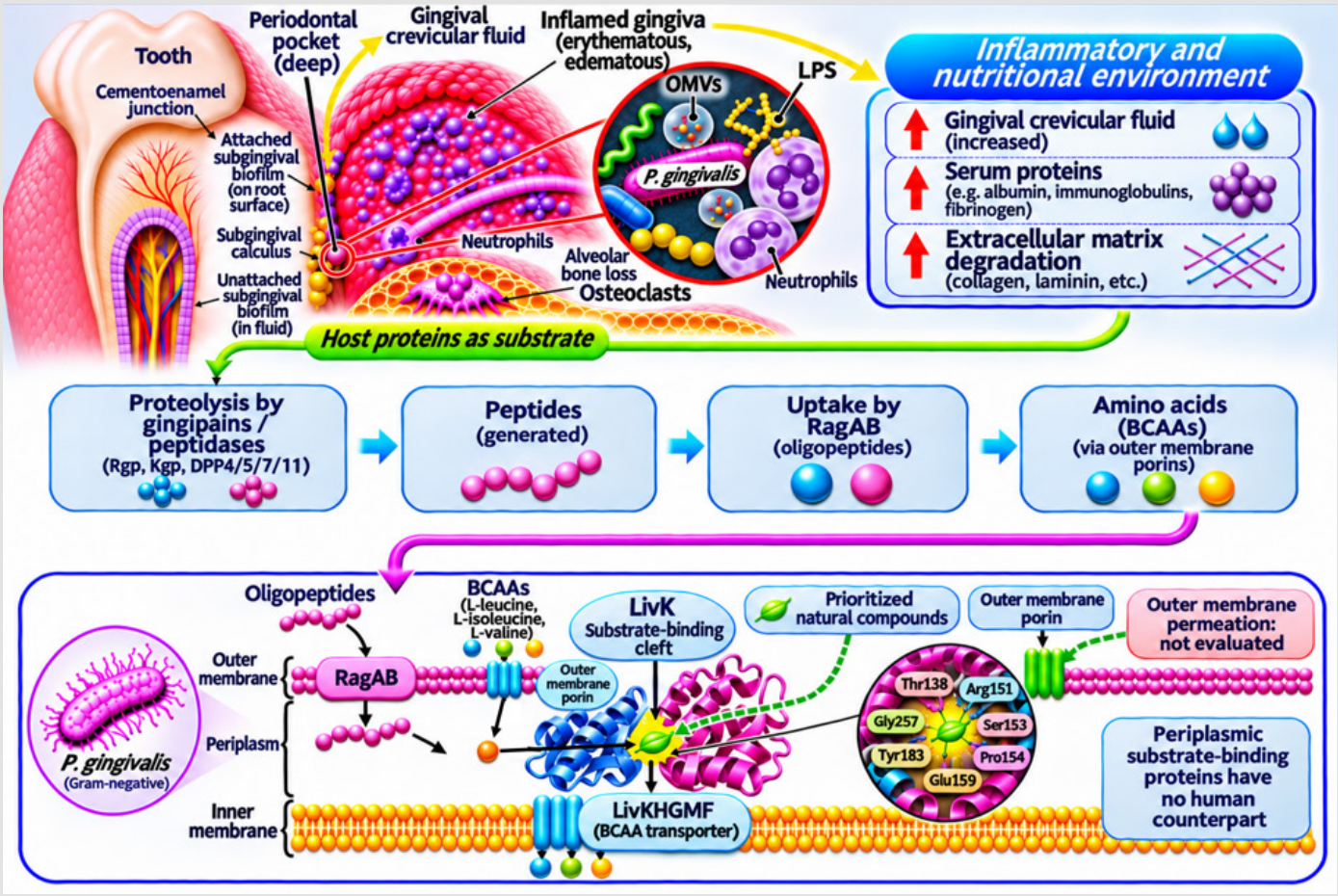


Figure 8: Proposed molecular mechanism of nutrient recognition by LivK and its potential modulation by natural compounds in *P. gingivalis*. Schematic representation of the inflammatory microenvironment of the periodontal pocket, the processing of host-derived proteins by gingipains and peptidases, and oligopeptide uptake through RagAB. The proposed periplasmic recognition of branched-chain amino acids (BCAAs: leucine, isoleucine, and valine) by LivK, associated with the LivKHGMF ABC transporter, is illustrated, along with the proposed occupancy of its binding pocket by natural compounds prioritized through virtual screening. Molecular competition and its potential functional consequences remain hypotheses requiring experimental validation.

However, these effects cannot be directly extrapolated to LivK, particularly because Wu, et al. [15] observed that *livk* deletion did not significantly affect bacterial growth under the conditions evaluated. Therefore, potential modulation of LivK may produce effects distinct from conventional bactericidal or bacteriostatic activity. Initial investigations should establish direct binding of the candidates and their ability to compete with natural BCAA substrates, followed by functional transport assays and studies determining whether modulation of this protein affects processes relevant to infection. Collectively, this study extends the molecular characterization of *P. gingivalis* beyond

its traditional virulence factors by investigating nutrient recognition through LivK, a protein experimentally associated with specific infection-related phenotypes [15]. Figure 8 integrates the available evidence into a proposed molecular mechanism linking the inflammatory microenvironment of the periodontal pocket to bacterial nutrient utilization. In this scenario, increased gingival crevicular fluid, the influx of serum proteins, and degradation of extracellular matrix components enhance the availability of protein substrates [3,4]. Their subsequent processing by gingipains and peptidases generates peptides that can be utilized by *P. gingivalis* [6,11,13]. Oligopeptide uptake through RagAB, as demonstrated by Madej et al. [14], represents one stage of this process, whereas LivK may participate in periplasmic BCAA recognition associated with the LivKHGMF transport system [15,52].

On this basis, our findings enabled the identification of a candidate binding region in LivK and the prioritization of ten natural compounds with more favorable docking scores than the natural substrate L-leucine. Among them, L-leucyl hydroxylamine (−5.71 kcal/mol), pentenomycin I (−5.66 kcal/mol), and (2R,3R,4S)-4-hydroxyisoleucine (−5.65 kcal/mol) exhibited the most favorable scores. However, integration of structural, physicochemical, and predictive toxicity criteria highlighted four higher-priority candidates: pentenomycin I, (2R,3R,4S)-4-hydroxyisoleucine, (R)-piperazic acid, and 2-oxoimidazolidine-4-carboxylic acid. These findings support the hypothesis that certain natural compounds may compete with LivK substrates, although their binding and functional modulatory potential require experimental validation. The relevance of this proposal is supported by experimental evidence implicating livk in cellular invasion and hepatic colonization [15], together with preclinical findings showing that targeting certain nutrient acquisition-related functions can alter outcomes of periodontal infection [51]. Nevertheless, a functional relationship between these mechanisms and the proposed molecular competition remains to be demonstrated. Experimental validation of ligand binding, transport, and their consequences for periodontal persistence and colonization will be essential to determine the potential of LivK as a molecular target for developing complementary strategies against *P. gingivalis*.

Conclusion

This study enabled the structural characterization of a putative LivK binding site in *P. gingivalis* and the prioritization of ten natural compounds through virtual screening. Integration of energetic, structural, and predictive toxicity criteria highlighted four candidates: pentenomycin I, (2R,3R,4S)-4-hydroxyisoleucine, (R)-piperazic acid, and 2-oxoimidazolidine-4-carboxylic acid. These findings provide a molecular basis for investigating potential competition with natural substrates and exploring LivK as a target for intervention against *P. gingivalis*. However, the pharmacological potential of these compounds and their relevance to periodontitis require experimental validation.

Acknowledgment

The author gratefully acknowledges the Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) for the financial support provided through the Estancias Posdoctorales por México program.

Data Availability Statement

The computational data supporting the findings of this study have been deposited in Zenodo under the DOI: https://doi.org/10.5281/zenodo.24254067.

The repository contains structural models, molecular docking inputs and outputs, compound identifiers, screening results, analytical scripts, and supporting documentation associated with the virtual screening of natural compounds targeting LivK of *P. gingivalis* ATCC 33277.

The COCONUT database is not redistributed in its entirety. Computational predictions should be interpreted within the methodological limitations described in the manuscript.

Declaration of Competing Interest

The author declares that there is no conflict of interest.

References

1. Kassebaum NJ, Bernabé E, Dahiya M, Bhandari B, Murray CJL, et al. (2014) Global burden of severe periodontitis in 1990-2010: a systematic review and meta- regression. J Dent Res 93(11): 1045-1053.
2. Tonetti MS, Jepsen S, Jin L, Otomo Corgel J (2017) Impact of the global burden of periodontal diseases on health, nutrition and wellbeing of mankind: A call for global action. J Clin Periodontol 44(5): 456-462.
3. Hajishengallis G, Lamont RJ (2021) Polymicrobial communities in periodontal disease: Their quasi-organismal nature and dialogue with the host. Periodontol 2000 86(1): 210-230.
4. Lamont RJ, Koo H, Hajishengallis G (2018) The oral microbiota: Dynamic communities and host interactions. Nat Rev Microbiol 16(12): 745-759.
5. Potempa J, Mydel P, Koziel J (2017) The case for periodontitis in the pathogenesis of rheumatoid arthritis. Nat Rev Rheumatol 13(10): 606-620.
6. Potempa J, Sroka A, Imamura T, Travis J (2003) Gingipains, the major cysteine proteinases and virulence factors of Porphyromonas gingivalis: Structure, function and assembly of multidomain protein complexes. Curr Protein Pept Sci 4(6): 397-407.
7. Mei F, Xie M, Huang X, Long Y, Lu X, et al. (2020) Porphyromonas gingivalis and its systemic impact: Current status. Pathogens 9(11): 944.
8. Hatasa M, Yoshida S, Takahashi H, Kenichi Tanaka, Yoshihito Kubotsu, et al. (2021) Relationship between NAFLD and periodontal disease from the view of clinical and basic research, and immunological response. Int J Mol Sci 22(7): 3728.
9. Hajishengallis G (2015) Periodontitis: from microbial immune subversion to systemic inflammation. Nat Rev Immunol 15(1): 30-44.
10. Sanz M, Marco del Castillo A, Jepsen S, et al. (2020) Periodontitis and cardiovascular diseases: Consensus report. J Clin Periodontol 47(3): 268-288.

11. Nemoto TK, Ohara Nemoto Y (2016) Exopeptidases and gingipains in Porphyromonas gingivalis as prerequisites for its amino acid metabolism. Jpn Dent Sci Rev 52(1): 22-29.
12. DP Miller, DA Scott (2021) Inherently and conditionally essential protein catabolism genes of Porphyromonas gingivalis. Trends Microbiol 29(1): 54-64.
13. Ohara Nemoto Y, Shimoyama Y, Kimura S, Kon A, Haraga H, et al. (2011) Asp- and Glu-specific novel dipeptidyl peptidase 11 of Porphyromonas gingivalis ensures utilization of proteinaceous energy sources. J Biol Chem 286(44): 38115-38127.
14. Madej M, White JBR, Nowakowska Z, Shaun Rawson, Carsten Scavenius, et al. (2020) Structural and functional insights into oligopeptide acquisition by the RagAB transporter from Porphyromonas gingivalis. Nat Microbiol 5(8): 1016-1025.
15. Wu L, Shi R, Bai H, Xingtong Wang, Jian Wei, et al. (2022) Porphyromonas gingivalis induces increases in branched- chain amino acid levels and exacerbates liver injury through livh/livk. Front Cell Infect Microbiol 12: 776996.
16. Tian J, Liu C, Zheng X, X Jia, X Peng, et al. (2020) Porphyromonas gingivalis induces insulin resistance by increasing BCAA levels in mice. J Dent Res 99(7): 839-846.
17. Ribardo DA, Hendrixson DR (2011) Analysis of the LIV system of Campylobacter jejuni reveals alternative roles for LivJ and LivK in commensalism beyond branched-chain amino acid transport. J Bacteriol 193(22): 6233-6243.
18. Davidson AL, Dassa E, Orelle C, Chen J (2008) Structure, function, and evolution of bacterial ATP-binding cassette systems. Microbiol Mol Biol Rev 72(2): 317-364.
19. Nikaido H (2003) Molecular basis of bacterial outer membrane permeability revisited. Microbiol Mol Biol Rev 67(4): 593-656.
20. Berntsson RP A, Smits SHJ, Schmitt L, Slotboom D J, Poolman B (2010) A structural classification of substrate-binding proteins. FEBS Lett 584(12): 2606-2617.
21. Sorokina M, Merseburger P, Rajan K, Yirik MA, Steinbeck C (2021) COCONUT online: Collection of Open Natural Products database. J Cheminform 13(1): 2.
22. Hopkins AL, Keserü GM, Leeson PD, Rees DC, Reynolds CH (2014) The role of ligand efficiency metrics in drug discovery. Nat Rev Drug Discov 13(2): 105-121.
23. Jumper J, Evans R, Pritzel A, Tim Green, Michael Figurnov, et al. (2021) Highly accurate protein structure prediction with AlphaFold. Nature 596(7873): 583-589.
24. Varadi M, Bertoni D, Magana P, Urmila Paramval, Ivanna Pidruchna, et al. (2024) AlphaFold Protein Structure Database in 2024: Providing structure coverage for over 214 million protein sequences. Nucleic Acids Res 52(D1): D368-D375.
25. (2025) UniProt: the Universal Protein Knowledgebase in 2025. Nucleic Acids Res. The UniProt Consortium 53(D1): D609-D617.
26. Laskowski RA, Jabłońska J, Pravda L, Vařeková RS, Thornton JM (2018) PDBsum: Structural summaries of PDB entries. Protein Sci 27(1): 129-134.
27. Wiederstein M, Sippl MJ (2007) ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35: W407-W410.
28. Colovos C, Yeates TO (1993) Verification of protein structures: Patterns of nonbonded atomic interactions. Protein Sci 2(9): 1511-1519.
29. Willard L, Ranjan A, Zhang H, Hassan Monzavi, Robert F Boyko, et al. (2003) VADAR: A web server for quantitative evaluation of protein structure quality. Nucleic Acids Res 31(13): 3316-3319.
30. Magnusson U, Salopek Sondi B, Luck LA, Mowbray SL (2004) X-ray structures of the leucine- binding protein illustrate conformational changes and the basis of ligand specificity. J Biol Chem 279(10): 8747-8752.
31. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The Protein Data Bank. Nucleic Acids Res 28(1): 235-242.
32. Krivák R, Hoksza D (2018) P2Rank: Machine learning based tool for rapid and accurate prediction of ligand binding sites from protein structure. J Cheminform 10(1): 39.
33. Jendele L, Krivák R, Skoda P, Novotny M, Hoksza D (2019) PrankWeb: A web server for ligand binding site prediction and visualization. Nucleic Acids Res 47(W1): W345-W349.
34. Hendlich M, Rippmann F, Barnickel G (1997) LIGSITE: Automatic and efficient detection of potential small molecule-binding sites in proteins. J Mol Graph Model 15(6): 359-363.
35. Le Guilloux V, Schmidtke P, Tuffery P (2009) Fpocket: An open source platform for ligand pocket detection. BMC Bioinformatics 10: 168.
36. Chandrasekhar V, Rajan K, Kanakam SRS, Sharma N, Weißenborn V, et al. (2025) COCONUT 2.0: a comprehensive overhaul and curation of the collection of open natural products database. Nucleic Acids Res 53(D1): D634-D643.
37. Landrum G, Tosco P, Kelley B, Ricardo Rodriguez, David Cosgrove, et al. (2026) RDKit: Open-source cheminformatics, release 2026_03_1 [software]. Zenodo.
38. O Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, et al. (2011) Open Babel: An open chemical toolbox. J Cheminform 3: 33.
39. Trott O, Olson AJ (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31(2): 455-461.
40. Eberhardt J, Santos Martins D, Tillack AF, Forli S (2021) AutoDock Vina 1.2.0: New docking methods, expanded force field, and Python bindings. J Chem Inf Model 61(8): 3891-3898.
41. Bickerton GR, Paolini GV, Besnard J, Muresan S, Hopkins AL (2012) Quantifying the chemical beauty of drugs. Nat Chem 4(2): 90-98.
42. Daina A, Michielin O, Zoete V (2017) SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep 7: 42717.
43. Banerjee P, Kemmler E, Dunkel M, Preissner R (2024) ProTox 3.0: A webserver for the prediction of toxicity of chemicals. Nucleic Acids Res 52(W1): W513-W520.
44. Kanehisa M, Furumichi M, Sato Y, Matsuura Y, Ishiguro Watanabe M (2025) KEGG: Biological systems database as a model of the real world. Nucleic Acids Res 53(D1): D672-D677.
45. Sayers EW, Bolton EE, Fine AM, Kelly C, Kim S, et al. (2026) Database resources of the National Center for Biotechnology Information in 2026. Nucleic Acids Res 54(D1): D20-D27.
46. Shannon P, Markiel A, Ozier O, et al. (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res 13(11): 2498-2504.

47. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc Series B 57(1): 289-300.

48. Hajishengallis G, Darveau RP, Curtis MA (2012) The keystone-pathogen hypothesis. Nat Rev Microbiol 10(10): 717-725.

49. Ohara Nemoto Y, Rouf SMA, Naito M, Amie Yanase, Fumi Tetsuo, et al. (2014) Identification and characterization of prokaryotic dipeptidyl-peptidase 5 from Porphyromonas gingivalis. J Biol Chem 289(9): 5436-5448.

50. Farrugia C, Stafford GP, Potempa J, Wilkinson RN, Chen Y, et al. (2021) Mechanisms of vascular damage by systemic dissemination of the oral pathogen Porphyromonas gingivalis. FEBS J 288(5): 1479-1495.

51. Aoki Nonaka Y, Minato Y, Hidaka K, Yuko Warita, Daiki Ando, et al. (2026) A DPP inhibitor suppresses periodontitis via antibacterial effect targeting Porphyromonas gingivalis. Sci Rep 16(1): 21180.

52. Adams MD, Wagner LM, Graddis TJ, Landick R, Antonucci TK, et al. (1990) Nucleotide sequence and genetic characterization reveal six essential genes for the LIV-I and LS transport systems of Escherichia coli. J Biol Chem 265(20): 11436-11443.

ISSN: 2574-1241

DOI: 10.26717/BJSTR.2026.67.010494

Salvador Pérez-Mora . Biomed J Sci & Tech Res

This work is licensed under Creative *Commons* Attribution 4.0 License

Submission Link: https://biomedres.us/submit-manuscript.php



Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles

https://biomedres.us/