Contents lists available at Curevita Journals

CUREVITA INNOVATION OF BIODATA INTELLIGENCE

Biodiversity Mining of Novel Small Molecule Inhibitors Targeting the N-terminal Domain of *Enterococcus faecalis* Esp Protein

Shrey Jain[1], Usha Chouhan[2], Deedhiti Mistry[3], Sonu Kurmi[4]

Article info

Article history: Received 2nd Sept 2025, Revised 18 Nov 2025, Accepted 4th Dec 2025, Published Dec 2025

Keywords: Enterococcus faecalis, Esp protein, N-terminal domain biofilm inhibition, molecular docking, virtual screening, ADMET profiling.

Author: Dr Usha Chouhan, Prof., Department of Bioinformatics, Maulana Azad National Institute of Technology, Bhopal, Madhya Pradesh, India.

Email ID: chouhanu@manit.ac.in

[1]Student, Amity Institute of Biotechnology, Amity University Chhattisgarh, Raipur, India.

[3]Research Scholar, Department of Mathematics, Bioinformatics and Computer Applications, MANIT Bhopal, Madhya Pradesh, India.

[4]Research Scholar, Department of Mathematics, Bioinformatics and Computer Applications, MANIT Bhopal, Madhya Pradesh, India.

Citation: Shrey Jain, Chouhan Usha, Mistry Deedhiti. 2025.Biodiversity Mining of Novel Small Molecule Inhibitors Targeting the N-terminal Domain of *Enterococcus faecalis* Esp Protein.Curevita Innovation of BioData Intelligence 1,2,98-119.

Publisher: Curevita Research Pvt Ltd

Abstract

Enterococcus Faecalis poses a significant clinical challenge as it can form persistent biofilms and exhibits multidrug resistance, with the Enterococcal surface protein (Esp) acting as a vital virulence factor for stabilizing biofilms. A computational approach was used to identify small-molecule inhibitors that selectively target the Esp N-terminal domain to disrupt biofilm formation and decrease virulence. Using the Esp crystal structure, molecular docking simulations with AutoDock 4.2 and AutoDock Vina identified lead compounds such as Myricetin and Quercetin dihydrate, which demonstrated high binding affinities. Extensive virtual screening and ADMET profiling with SwissADME confirmed that these lead compounds possess favourable drug-like qualities and pharmacokinetic properties, including high gastrointestinal absorption and adherence to Lipinski's Rule of Five. This *in silico* study successfully highlights a set of potential lead compounds as novel anti-virulence agents.

Article highlights: -

What is already known

- *Enterococcus faecalis* is a significant nosocomial pathogen known for its extensive drug resistance and ability to form persistent biofilms that protect it from antibiotics and host immunity.
- The Enterococcal surface protein (Esp) is a critical virulence factor that mediates biofilm formation, with its N-terminal domain forming stabilising amyloid fibrils in a pH- pH-dependent manner.
- The crystal structure of the Esp N-terminal domain provides a suitable molecular target for structure-based drug discovery aimed at inhibiting biofilm.

What this paper adds

- This study identifies several promising small-molecule lead compounds, including Myricetin and Quercetin derivatives, as potent inhibitors of the Esp protein's N-terminal domain.
- The findings provide a foundational basis for a novel therapeutic strategy targeting the Esp protein to combat *E. faecalis* infections without promoting conventional antibiotic resistance.

Introduction

E. faecalis is a Gram-positive, facultative anaerobe coccus that is a natural inhabitant within the gastrointestinal system of humans. It is a commensal organism in people who have healthy immune systems but has gained widespread attention as a serious opportunistic pathogen in clinical cases involving impaired health. Over the past decades, *E. faecalis* has become a common causative agent in nosocomial infections that include urinary tract infection, bacteraemia, endocarditis, and implant-associated infections such as catheters, prosthetic heart valves, and orthopaedic implants (Singh et al., 2021; Abdeltawab et al., 2022). Such infections become characteristically chronic, recurrent, and refractory to therapy primarily due to the extraordinary ability of the bacterium to be resistant to drugs and to generaten persistent biofilm structures.

The increasing prevalence of

multidrug-resistant strains of *E. faecalis*, especially vancomycin-resistant enterococci (VRE) strains, is a significant concern for the current healthcare systems across the globe (Singh et al, 2021). Inherent and acquired resistance of *E. faecalis* is attributed to varying degrees of low-affinity penicillin-binding proteins coupled with horizontal gene transfer and other associated traits of vancomycin inhibition operons. Alarming predispositions to multidrug resistance profoundly narrow the treatment options available and highlight the urgent importance of new treatment paradigms that can bypass the conventional bactericidal methods. A highly disturbing feature associated with infection by *E. faecalis* is a recurrent link to biofilm production. Biofilms are highly ordered microbial communities attached to surfaces and encapsulated within a self-produced extracellular polymeric matrix made up of polysaccharides, proteins, and extracellular DNA. Bacteria within these biofilms reside in an environment insulated against host immune response while protected against drug invasion, such that conventional therapy stands largely ineffective (Abdeltawab et al., 2022). The biofilm lifestyle has a tolerance to antibiotics that is up to 1000 times higher compared to that of planktonic cells and makes eradicating infection related to medical device infections The protective extracellular polymeric matrix both shields against biofilm attack and supports its development, resulting from a combination of environmental factors and genetic traits of the bacteria, This matrix aids in cell communication and is crucial for sustained cell vitality. In clinical practice, biofilms developed by *E. faecalis* have been commonly reported in connection with urinary catheter use, endocardial tissues, recurrent endodontic infection in root canals, and orthopaedic implant infections. It is associated to a

significant extent with treatment failures, chronic infection, and high relapse rates (Singh et al., 2021). In the structure of a biofilm, bacterial metabolic activity is reduced such that active cellular function-targeting antibiotics become less effective. Additionally, penetration by antimicrobial agents and immune responders is prevented by the matrix structure. Overall, these traits highlight biofilms as a key survival mechanism for

E. faecalis while making it even more important to understand distinct molecular pathways controlling development in a biofilm. Among all the different virulence factors associated with *E. faecalis*, the Enterococcal surface protein has been highly studied due to its role in the process of biofilm development and stabilisation. Esp is a large protein that is anchored to the cell surface and has an estimated molecular weight of approximately 202 KDa (Spiegelman et al., 2022). It belongs to a family of periscope proteins that can be characterized by an elongated structure encompassing variable repeat domains extending out from the bacterial surface. These proteins function as "molecular periscopes," enabling a capacity for bacteria to modify interactions between them and the environment.

Esp is stably attached to peptidoglycan in the outer bacterial cell wall via a conserved LPXTG motif recognised by sortase enzymes and hence stably localised on the bacterial surface (Mohamed et al., 2005; Toledo-Arana et al., 2001). Its structural organization is modular and comprises four principal parts: (1) a signal peptide sequence approximately 49 residues in length; (2) an N-terminal non-repeat domain comprising a sequence of 694 amino acids; (3) a central component harbouring numerous copies of tandem repeats also referred to as A, B, and C repeats which differ between strains due to recombination; and (4) a C-terminal anchor region enabling attachment via

sortase.

Among these structures, the N-terminal non-repeat region of Esp is particularly critical in its role in forming biofilms. Through crystallographic work (PDB ID: 6ORI), this region has been determined to possess a β-sheet-rich structure made up of two DE-variant immunoglobulin-like folds in tandem (Spiegelman et al., 2022; Al-Ahmad et al., 2024). Through this structure, protein-protein interactions and higher-order structural assemblies on the bacterial surface become plausible.

Functionally, this N-terminal domain operates to drive biofilm development via a variety of mechanisms. Early work showed that this was sufficient to enable glucose-based biofilm development (Tendolkar et al., 2005). Subsequent studies have elucidated a more complex function: the domain possesses conformational plasticity dependent upon environmental changes in pH. Neutral pH provides a globular structure and functionally inactive amyloid fibrillation. However, it undergoes a significant conformational shift and unwinds into amyloid-like fibrils when exposed to acidic pH levels (≤4.3), which are characteristic of mature biofilms and inflammatory host tissues. By integrating into the extracellular biofilm matrix, these fibrils provide mechanical strength, resistance to enzymatic digestion, and decreased susceptibility to host defence. (Spiegelman et al., 2022). The conformation shift is tightly regulated by an autoinhibitory system comprising the Esp452 core plus residues 453–743 such that the globular conformation is preserved in neutral pH media. Removal of this autoinhibition due to proteolysis or a shift in pH permits fibril growth and maturation of biofilm.

Amyloidogenic traits inherent in the

N-terminal region of Esp represent a sophisticated mode of bacterial adaptation. Amyloid fibrils possess high structural rigidity, resist proteolytic degradation, and can act as structural reinforcement within the biofilm.

These characteristics enable *E. faecalis* biofilms to withstand a variety of environmental stressors, thwart the actions of host immune effectors, and withstand antimicrobial treatment. Esp-derived bioactive amyloid fibrils incorporated into the biofilm matrix offer a survival-evolutionarily beneficial persistence strategy, hence increasing a pathogen's capacity to colonize and withstand elimination in clinical settings (Spiegelman et al., 2022).

As a principal factor in biofilm stability, the Esp N-terminal domain emerges as a promising target for therapy. Notably, inhibition of Esp-mediated fibrillation does not kill *E. faecalis* directly but dampens its virulence, causing less selective pressure against resistance development. This anti-virulence approach is consistent with new paradigms of therapy that aim to disarm pathogens rather than eradicate them, thus conserving host microbiota while fighting infection. Esp as a putative pharmacological target. Structure-based virtual screening allows rapid evaluation of large compound collections against the Esp N-terminal domain. Molecular docking programs such as AutoDock4.2 and AutoDock Vina, complemented by ligand preparation using MGL Tools, improve the accuracy of predictions about ligand binding affinities and modes of engagement(Verma & Chouhan, 2018a).

MATERIALS & METHODS

All the steps completed in this project are performed in HP Pavilion laptop 15- eg2035tu, which has 12th Gen Intel(R) Core

(TM) i5-1240P (1.70 GHz) and 16.0 GB (15.7 GB usable).

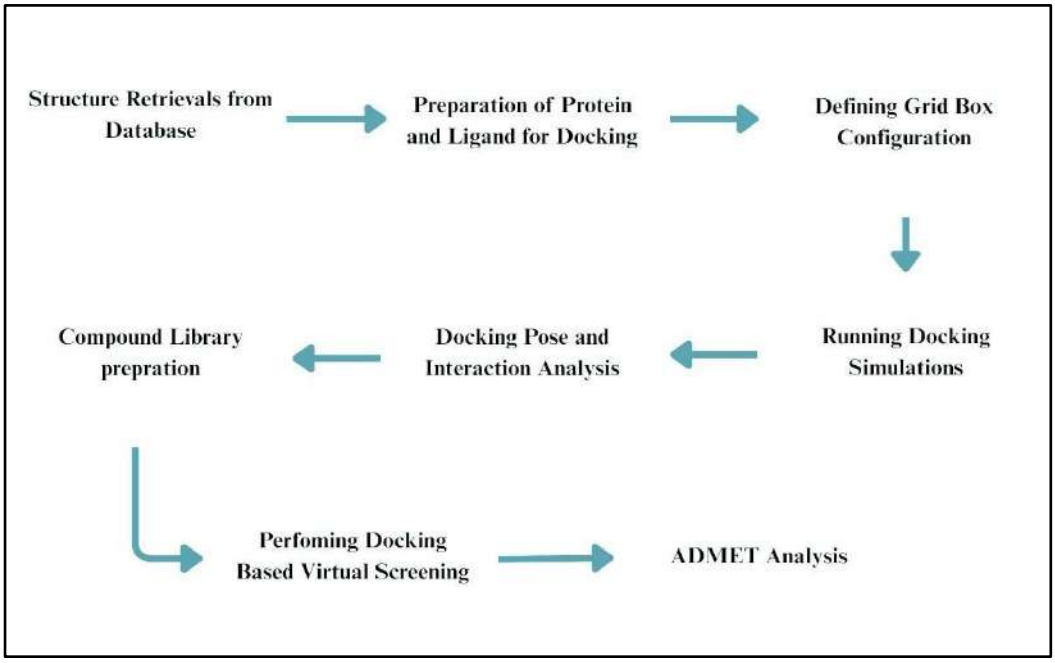


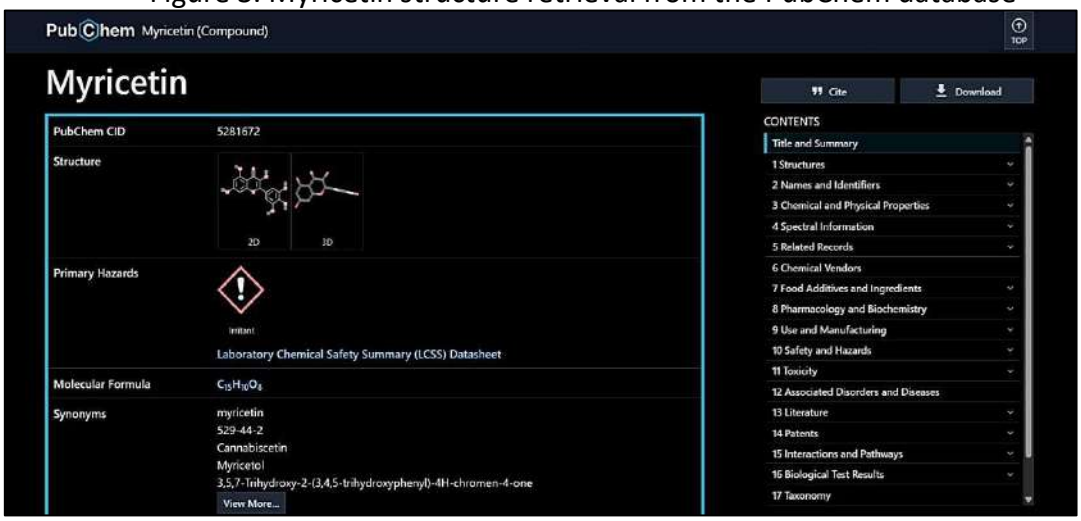
Figure 1. Workflow

Structure retrieval of protein and ligand molecules

The crystal structure of the *E. faecalis* Enterococcal Surface Protein (Esp) N-terminal domain was retrieved from PDB: 6ORI (Kim et al., 2021), as shown in Figure1. This structure, resolved through X-ray crystallography at 1.40 Å resolution, represents residues 50th to 694th, following the 49th-residue signal peptide. The N- terminal domain was selected due to its functional importance in mediating biofilm enhancement through amyloid-like fibril formation under acidic conditions

Figure 2. Enterococcal surface protein with partial N-terminal region.

Figure 3. Myricetin structure retrieval from the PubChem database



(Spiegelman et al., 2022). It is rich in β- sheet content, exhibiting a globular fold with multiple domains involved in fibrillation. The ligand structures were retrieved from the PubChem Compound Database in 3D SDF format, as shown in Figure 3. The selected compounds included Myricetin, Quercetin, Parietin, Epigallocatechin Gallate (EGCG), 11-buffer ions, and any unused heteroatoms, and any missing atoms were added. Polar hydrogen atoms were also added to the protein to facilitate hydrogen bonding to better resemble actual conditions, and Kollman charges were automatically assigned to all atoms to facilitate better electrostatic calculations. As for ligands, 3D SDF structures contained within the Deoxyalisol-B, Delphinidin and Propylcholestrol. These ligands were then converted into PDB format using OpenBabel Chemical File Format Converter for further molecular docking studies.

Protein and Ligand Preparation Using AutoDock Tools

The molecular structure of the *E. faecalis* Esp N-terminal domain and selected ligand molecules were prepared for molecular docking by AutoDock Tools version 1.5.7. (Sanner, 1999) Before commencing the docking process, the protein structure was optimised by removing water molecules. PubChem Compound Database were converted to PDB format initially by using the OpenBabel tool. Then they were converted to PDBQT format via AutoDock Tools. Gasteiger charges were calculated and applied during this process to each ligand. Non-polar hydrogens were merged, and all rotatable bonds were determined and designated flexible to permit the ligand to shift shape during the docking process. A separate PDBQT format was saved for each ligand once flexible units were determined. Both protein receptor and ligand molecules were inspected visually within ADT to

ensure that atom type, bonding, and charge allocations were correct. PDBQT files were then utilised as inputs for AutoDock Vina docking simulations.

Molecular docking using AutoDock4.2

In molecular docking analysis, AutoDock version 4.2 (Morris et al., 2009) was used to predict the binding efficiency and interacting behaviour of some screened ligands against the N-terminal domain of Enterococcus faecalis Esp (PDB ID: 6ORI). It comprised the generation of a grid box, the generation of AutoGrid-based grid maps, the generation of AutoDock parameter files, the execution of AutoDock employing the Lamarckian Genetic An algorithm search algorithm, and the validation of AutoDock outputs. These steps have been further described in the sections below.

Grid Box Preparation

The binding site on the Esp N-terminal domain was identified based on structural cavities and previously reported functional domains. A grid box was prepared to cover the binding region completely. The grid box dimensions were typically set to 126 × 126

× 126 points with a grid spacing not more than 0.372, and the grid centre coordinates X, Y, Z were selected to encompass the active site as given in Table 1.

Table 1. Blind dock Grid box Co-ordinates

ID	Spacing	X-centre	Y-centre	Z- centre
PubChem CID 3696342	0.419 Å	1.234	-5.018	1.271

Running AutoGrid

Having defined the grid box, AutoGrid was executed to produce the needed grid map files for docking. AutoDock Tools were utilized to produce a grid parameter file that contained receptor atom types and grid box parameters. Executing AutoGrid resulted in some map files being automatically created corresponding to each atom type in ligands. Pre-calculated interaction energies contained in these maps serve to accelerate the process of scoring while docking. Docking preparation

Docking parameter files were generated for each ligand using AutoDock Tools. These files included algorithm selection, search parameters, and evaluation criteria. The Lamarckian Genetic Algorithm was selected due to its capability to combine global and local search methods for better convergence and accuracy.

Running AutoDock

Docking simulations were performed on all ligands using AutoDock; thus, dpf and map files were created. AutoDock considered several binding conformations, examining their interactions within the Esp protein across consecutive generations with the help of the LGA approach. AutoDock created docking log files that contain information about Estimated Binding Free Energy (ΔG) and conformations determined by RMSD levels. The most ranked conformation associated with each ligand having a minimum binding energy was selected for further evaluation.

Following the preliminary blind docking, ligands exhibiting the lowest binding energies, along with the most favourable interactions with the *E. faecalis* Esp N- terminal domain, were selected for focused docking to improve their binding orientations. The binding site was determined by examining the centroid of the most favourable coordinates seen during the

blind docking process, as listed in Table 2. A novel grid box typically 100 × 100 × 100 points was created around this site employing AutoDock Tools version 1.5.7. AutoGrid was then re-run to generate new grid map files appropriate to this refined area. Parameter files for docking were created employing the Lamarckian Genetic Algorithm (LGA) to maintain consistency with parameter files employed in previous runs. Focused docking was then performed employing AutoDock 4.2, utilising new grid and parameter files. Generated docking log files were then analysed in AutoDock Tools to identify the lowest-energy binding conformations, by comparing the blind docking results were employed to examine improvement in binding affinity.

Table 2. Fixed Docking Grid Co-ordinates

ID	Spacing	X-centre	Y-centre	Z- centre
PubChem CID 3696342	0.405 Å	12.587	-6.397	6.476

Analysis of Ligand-Protein Interactions

After completing molecular docking using AutoDock4.2, detailed analysis of ligand-protein binding interactions was performed to understand the structural basis of binding within the *E. faecalis* Esp N-terminal domain. Docking log files were loaded into AutoDock Tools 1.5.7 for evaluation. The best docking pose for each ligand was selected based on the lowest estimated binding energy ΔG and RMSD clustering, ensuring structural consistency and stability, commonly RMSD ≤ 2.0 Å. Using ADT's Analyse Docking feature, the ligand's spatial orientation inside the binding pocket was visualised, and interaction parameters were identified. Key non-covalent interactions such as hydrogen bonds, hydrophobic interactions, and electrostatic contacts between the

ligand and the amino acid residues of the Esp N-terminal were examined. These contacts were matched against the active site region defined during grid box preparation to confirm specific and meaningful binding.

Ligand-based Virtual Screening through AutoDock Vina

With the identification of high-affinity lead composites through molecular docking against the Esp N-terminal sphere (PDB ID 6ORI), Ligand-Based Virtual Screening (LBVS) was carried out using AutoDock Vina (Trott & Olson, 2010) using Ubuntu background (Canonical Ltd., 2024) via Windows Subsystem for Linux (Microsoft Corporation, 2024). The top docked ligands served as reference structures for retrieving structurally analogous composites from public chemical libraries such as PubChem using similarity searching. To ensure medicine likeness and bioavailability, all recaptured composites were filtered based on Lipinski's Rule of Five, which includes criteria like molecular weight ≤ 500 Da, logp ≤ 5, no further than 5 Hydrogen Bond Donors, and no further than 10 Hydrogen Bond Acceptors. composites meeting these criteria were converted from SDF to PDBQT format using Open Babel with figure optimization enabled (Bandyopadhyay et al., 2016). The Esp N- terminal protein was pre-processed in AutoDock Tools by adding polar hydrogens, assigning Kollman charges, and saving the final structure as a PDBQT file. A focused grid box was created, and a configuration file was prepared specifying parameters such as centre equals, grid size, exhaustiveness, and number of affair binding modes. A Bash script was written to automate docking of all ligands, and AutoDock Vina was used to calculate binding affinities and induce affair conformations. The performing docking scores (ΔG) were extracted

from log lines and ranked to prioritise composites with the most favourable list parameters. This virtual webbing approach enabled the identification of new Esp N-terminal impediments that not only showed structural similarity to known binders but also satisfied medicine-likeness criteria, making them suitable campaigners for further ADME/ Tox profiling.

ADMET Profiling Using SwissADME

After performing virtual screening through AutoDock Vina, ligands with high binding affinities typically between –8.0 to –9.5 kcal/mol were shortlisted for *in silico* ADMET profiling using the Swiss ADME. This step aimed to evaluate the drug-likeness, pharmacokinetic behaviour, and physicochemical properties of structurally similar analogues of the lead compound to facilitate effective prioritisation for further biological validation. First, the canonical SMILES format of each compound was retrieved from PubChem or ZINC databases. These SMILES strings were individually submitted to SwissADME (Pandey et al., 2017, Verma & Chouhan, 2018b). This tool provides parameters such as lipophilicity logP, water solubility logS, topological polar surface area TPSA, number of rotatable bonds, gastrointestinal (GI) absorption, blood-brain barrier (BBB) permeability, and cytochrome P450 (CYP) enzyme inhibition profiles. Compounds were also evaluated for drug-likeness using Lipinski's Rule of Five, Veber, Ghose, and Egan filters, along with BOILED-Egg plots to visualise passive absorption and brain penetration capabilities. Additional properties such as bioavailability score, synthetic accessibility, and P-glycoprotein substrate behaviour were also reviewed. The pharmacokinetic evaluation focused on drug-likeness, physicochemical properties, and oral bioavailability to prioritise compounds for further development.

All selected ligands satisfied Lipinski's Rule of Five, indicating favourable drug-likeness with molecular weights

below 500 Da, logP values under 5, no more than 5 hydrogen bond donors, and fewer than 10 hydrogen bond acceptors. Additionally, all candidates demonstrated optimal topological polar surface area < 140 Å² and ≤10 rotatable bonds, adhering to have low CYP inhibition. Synthetic accessibility scores ranged from 2.5 to 4.5, suggesting that these molecules can be synthesised with minimal complexity. In conclusion, the ADMET analysis validated the molecular docking findings, confirming that the top- performing Esp N-terminal inhibitors not only exhibit strong binding affinity but also possess desirable pharmacokinetic and drug-likeness properties (Kumar Suryawanshi et al., 2022). These compounds present promising scaffolds for further optimization and experimental validation against *E. faecalis* biofilm-associated infections.

Results and Discussion

Veber's rule, which confirms good oral bioavailability. Furthermore, all ligands displayed logS values better than -4, suggesting acceptable water solubility, and bioavailability scores ≥ 0.55, indicating favourable oral drug potential. Additionally, they should

Blind Docking and Fixed Docking using AutoDock4

Blind docking of named ligands with *E. faecalis* Esp N-terminal PDB: 6ORI was carried out using AutoDock 4.2 to identify binding sites across the entire protein face. Docking scores binding energy, ΔG ranged from -6.00 to -9.9 Kcal/mol, particularly Myricetin, Quercetin, dihydrate, 11DeoxyAlisol-B and 1,3,4-Oxadiazole,2- Napthol demonstrated the highest binding affinity. Based on blind docking results, fixed docking is carried out using a defined grid box centred around the best binding coordinates. This type of fixed docking improved binding site precision and confirmed interaction consistency. The binding energies remained stable, with slight

improvements observed. The results of blind docking and fixed docking of compounds with good binding energy are shown in Table 3.

Table 3. Blind docking and Fixed Docking Results

S.NO.	Source	Ligand name	PubChem CID	Blind dock Binding energy	Fixed dock Binding energy
L1	**PubChem**	**MYRICETIN**	**5281672**	**-9.84**	**-9.26**
L2	PubChem	QUERCETIN DIHYDRATE	5284452	-9.38	-9.35
L3	PubChem	PARIETIN	10639	-7.14	-7.08
L4	PubChem	ECGC	65064	-8.04	-7.05
L5	PubChem	11-DEOXYALISOL-B	3696342	-8.88	-9.59
L6	PubChem	DELPHINIDIN	68245	-7.84	-7.56
L7	PubChem	ROSMARINIC ACID	5281792	-6.98	-6.27
L8	PubChem	THYMOQUINONE	10281	-6.49	-6.54
L9	PubChem	1,3,4-OXADIAZOLE, 2 NAPHTHOL	11995217	-9.62	-9.79
L10	PubChem	PROPYLCHOLESTROL	14213727	-6.84	-7.98

Protein-Ligand Interaction Analysis

After performing molecular docking, the analysis of compounds or ligands' binding site is carried out to define their interaction and conformations with the Esp N-terminal protein molecule. Here, a protein-ligand complex of Myricetin-Esp N-terminal protein is taken, showing interaction between myricetin and Esp protein, having a good binding energy of -9.84 Kcal/mol in Figure

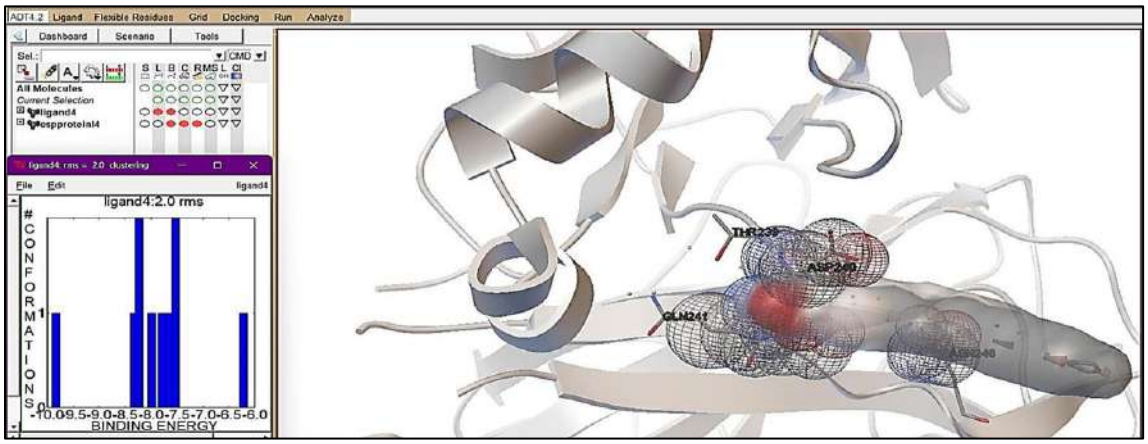


Figure 4. Analysis of Esp N-terminal-Myricetin complex in AutoDock4.2

Ligand-Based Virtual Screening (LBVS) and Binding Analysis in AutoDock Vina

The virtual screening of structurally similar compounds was performed using AutoDock Vina at the fixed binding site. Among the screened compounds, the top hits were shortlisted based on binding energy and are presented in Table 4 with their fixed co-ordinates (X, Y, Z).

Table 4. Coordinates of Compounds and Docking Results from Virtual Screening

Compound name	Binding Energy (Kcal)	X coordinate	Y coordinate	Z coordinate
11-deoxyalisol-B	-8.1	12.587	-6.397	6.476
Thymoquinone	-7.0	-7.049	-14.580	1.965
Myricetin	-9.0	11.724	-5.341	4.664
Quercetin dihydrate	-9.8	6.058	-3.333	1.799
Parietin	-9.4	7.751	-3.369	2.725
ECGC	-9.6	13.343	-6.930	4.019

Table 5. SwissADME Physicochemical Properties

Compound ID	C log p (-5 to +5)	Mol. wt. (150 to 500 g/mol)	HBD (0-5)	HBA (0-10)	Rotatable Bonds (<10)	TPSA (60-140 Å)	Solubility (mg/ml)
90643991	3.19	410.35	4	8	4	120.36	5.90E-03
127031525	2.79	322.31	3	5	3	90.9	1.71E-02
91162275	2.2	314.29	4	6	3	111.13	2.67E-02
127030288	2.86	320.3	3	5	1	90.9	8.16E-03
168298172	1.77	272.25	3	5	1	86.99	1.68E-01
90643984	3.23	410.35	4	8	4	120.36	1.72E-03

118713519	2.32	330.33	4	6	2	107.22	1.83E-02
156891827	3.31	376.36	3	6	6	104.06	6.08E-03

ADMET analysis using SwissADME

The top ligands identified from AutoDock Vina virtual screening, having binding energy between –8.0 to –9.7 kcal/mol were further evaluated for their drug-likeness and pharmacokinetic profiles using SwissADME (Daina, Michielin, & Zoete, 2017). This analysis included Lipinski's Rule of Five, physicochemical properties, and predicted pharmacokinetics. Among all compounds, quercetin and its similar structures have shown optimal physicochemical properties and best pharmacokinetics predictions, as shown in Tables 5 and 6.

Table 6: SwissADME Pharmacokinetic predictions

CID	Formula	GI absorption	BBB permeability	P-gp substrate	Skin Permeation	Bioavailability Score	Synthetic Accessibility
90643991	C22H15FO7	High	No	No	-6.35	0.55	3.67
127031525	C19H14O5	High	No	No	-5.85	0.55	3.57
91162275	C17H14O6	High	No	No	-6	0.55	3.14
127030288	C19H12O5	High	No	No	-5.75	0.55	3.07
168298172	C15H12O5	High	No	Yes	-6.49	0.55	3.03
90643984	C22H15FO7	High	No	No	-5.74	0.55	3.62
118713519	C18H18O6	High	No	Yes	-5.92	0.55	2.95
156891827	C22H16O6	High	No	No	-5.73	0.55	3.19
9116275	C17H14O5	High	No	No	-5.56	0.55	3.07

Conclusion

This report presents our study, which looked at the role of computation methods in identifying inhibitors of the N-terminal domain of *E. faecalis* Esp, which is a biofilm-associated virulence factor. We used AutoDock 4.2 and Vina for docking studies, which reported that Myricetin, Quercetin derivatives, 11-Deoxyalisol-B, and 1,3,4-Oxadiazole,2-Napthol have strong binding to Esp with affinities which range from –9.9 to –8.0 kcal/mol. What we saw is that these interactions, which are stabilised by hydrogen bonding and hydrophobic contacts, are very specific and favourable within the protein's active site. We did virtual screening, which expanded our field of potential inhibitors which we then ran through SwissADME which in turn confirmed that we do indeed have favourable pharmacokinetic properties, which include good solubility, oral bioavailability and compliance with drug-likeness rules. Out of all the candidates, Myricetin and Quercetin analogues did the best, which makes them very strong leads for development.

In summary, this integrated docking, screening and ADMET approach identified stable compounds capable of disrupting Esp function and interfering with biofilm formation in *E. faecalis*. These findings support anti-virulence strategies as a promising alternative to conventional antibiotics and lay the groundwork for experimental validation and future drug discovery efforts.

References

Al-Ahmad, A., et al. (2024). Structural and functional analysis of amyloidogenic domains in Enterococcus faecalis Esp protein. Journal of Molecular Biology, 436(5), 168733.

Kumar Suryawanshi, S., Chouhan, U., & Kant Choudhari, J. (2022). Insilico Anticancer Peptide Prediction from Curcuma longa. *Molecular & Cellular Biomechanics*, *19*(4), 191–208.

Verma, N., & Chouhan, U. (2018a). In-silico analysis of phenyl propanoic acid derivatives to design potent peroxisome proliferator-activated receptor (PPAR) dual agonists for type 2 diabetes mellitus therapy. *Oriental*

Journal of Chemistry, *34*(3), 1400.

Verma, N., & Chouhan, U. (2018b). ADMET Prediction of Dual PPARα/γ Agonists for Identification of Potential Anti-diabetic Agents. In *Machine Intelligence and Signal Analysis* (pp. 355-362). Singapore: Springer Singapore.

Pandey, D., Chouhan, U., & Verma, N. (2017). Identification and analysis of lead compounds against HIV-1 integrase enzyme using ADMET prediction and docking analysis. *International Journal of Pharmaceutical Sciences and Research*, *8*(10), 4129-4137.

Mohamed, J. A., Huang, D. B., Jiang, Z. D., DuPont, H. L., & Murray, B. E. (2005) N-terminal domain of Esp enhances biofilm formation by Enterococcus faecalis. Infection and Immunity, 73(9), 4901–4903.

Tendolkar, P. M., Baghdayan, A. S., & Shankar, N. (2005). The N- terminal domain of enterococcal surface protein Esp is sufficient for Esp-mediated biofilm enhancement. FEMS Microbiology Letters, 253(1), 127–132.

Toledo-Arana, A., Valle, J., Solano, C., Arrizubieta, M. J., Cucarella, C., Lamata, M., Amorena, B., Leiva, J., Penadés, J. R., & Lasa, I. (2001). The enterococcal surface protein, Esp, is involved in Enterococcus faecalis biofilm formation. Applied and Environmental Microbiology, 67(10), 4538–4545.

Spiegelman, J., Kline, K. A., & Nielsen, H. V. (2022). Structural insights into Esp fibril formation and its role in biofilm mechanics in Enterococcus faecalis. Nature Communications, 13, Article 3792.

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of Computational Chemistry, 30(16), 2785–2791.

Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry, 31(2),455–461.

Bandyopadhyay, B., Verma, N., & Chouhan, U. (2016). In Sillico analysis of newly identified potential drug lead compound against VP40 for the treatment of ebola virus infection. *International Journal of Advanced Biotechnology and Research*, *7*(4), 1357-1365.

Python Software Foundation. (2010). Python version 2.7.

Kristich, C. J., Rice, L. B., & Arias, C. A. (2014). Enterococcal infection—treatment and antibiotic esistance. *Infectious Disease inics of North America*, *28*(1), 1–17.

Jett, B. D., Huycke, M. M., & Gilmore, M. S. (1994). Virulence of enterococci. *Clinical Microbiology Reviews*, *7*(4), 462–478.

Heikens, E., Bonten, M. J., & Willems, R. J. (2007). Enterococcal surface protein Esp is important for biofilm formation of *Enterococcus faecium* E1162. *Journal of Bacteriology*, *189*(22), 8233–8240.

Rice, L. B. (2001). Emergence of vancomycin-resistant enterococci. *Emerging Infectious Diseases*, *7*(2), 183–187.

Zhou, Y., & Rui, L. (2010). Bioinformatics methods in

drug discovery. *Current Topics in Medicinal Chemistry*, *10*(2), 123–135.

Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, rug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, *7*, 42717.

Ferreira, L. G., dos Santos, R. N., Oliva, G., & Andricopulo, A. D. (2015). Molecular docking and structure-based drug design strategies. *Molecules*, *20*(7), 13384– 13421.

Sanner, M. F. (1999). Python: A programming language for software integration and development. *Journal of Molecular Graphics and Modelling*, *17*(1), 57–61.

Berman, H. M., et al. (2000). The Protein Data Bank. *Nucleic Acids Research*, *28*(1), 235–242.

Pires, D. E. V., Blundell, T. L., & Ascher, D. B. (2015). pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *Journal of Medicinal Chemistry*, *58*(9), 4066–4072.