

A multi-level computational analysis identifies *Ocimum basilicum* phytochemicals as potent inhibitors of the Marburg virus VP40 matrix protein

Md. Foyzur Rahman · Md. Mustak Khan · Md.Hridoy Ahmed · Shishir Dutta · Farah Naima Tapti · Umme Marzia Tasnim · Marwah Talukder · Md. Mahfuzur Rahman · Md. Foisal Imtiaz · Md. Taleb Hossain Roman · Laila Khaleda Ankhi · Mohammad Al-Forkan

Received: 15 February 2026 / Accepted: 14 September 2026
© The Author(s) 2026

Abstract

Marburg virus disease (MVD) is an extremely lethal hemorrhagic fever with no approved antiviral treatment available, which emphasizes the need to develop newer and safer drug candidates. In the current study, we used an in silico approach to identify candidate inhibitors of the Marburg virus matrix protein VP40, which plays a vital role in the viral assembly and budding process using bioactive compounds derived from *Ocimum basilicum*. A homology model of the Marburg virus VP40 protein was built and assessed by Ramachandran, QMEAN, ProSA, Verify3D and ERRAT analysis, and 156 bioactive compounds derived from *Ocimum basilicum* were then screened against it by molecular docking. Rutin ranked first, with a docking score of -10.4 kcal/mol and hydrogen-bonded and van der Waals contacts to residues of the predicted pocket. A 500 ns all-atom molecular dynamics simulation indicated a stable VP40-rutin complex, with a backbone RMSD of 1.98 ± 0.19 Å, a mean C α RMSF of 1.10 ± 0.52 Å, and a compact, well-preserved fold (radius of gyration 19.77 ± 0.15 Å). End-state binding free energies calculated over 51 uniformly sampled trajectory frames were favourable under both solvation models, at -46.08 ± 5.96 kcal/mol (MM/GBSA) and -27.92 ± 7.63 kcal/mol (MM/PBSA), driven primarily by van der Waals and electrostatic contributions. ADMET profiling, however, revealed a substantial pharmacokinetic liability: rutin violates three of Lipinski's criteria (molecular weight 610.52 g/mol, topological polar surface area 269.43 Å², 10 hydrogen-bond donors and 16 acceptors) and its predicted human intestinal absorption is only 23.45%, so it is best regarded as a lead scaffold requiring structural or formulation optimisation rather than a ready-to-use agent. Validation of the docking protocol against independent molecular dynamics reference poses gave a mixed result. The predicted affinity was reproducible, and the predicted pocket was clearly preferred over decoy boxes placed elsewhere on the protein surface, but pose recovery failed. No top-ranked pose fell within the conventional 2 Å success threshold of the reference pose. Docking was therefore used only as an initial screening filter to rank compounds, and neither the binding geometry nor any mechanism of inhibition is inferred from it. All structural and energetic conclusions reported here rest

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.



on the molecular dynamics ensemble and the end-state free-energy analyses. Gas-phase DFT analysis gave a narrow HOMO–LUMO gap (4.13 eV), high softness and a moderate electrophilicity index, which offer supportive electronic evidence of charge-transfer capacity rather than independent confirmation of affinity. Taken together, these findings identify rutin as a natural lead compound against VP40, subject to the pharmacokinetic and docking limitations set out above, and provide a critically validated computational framework for anti-Marburg drug discovery.

Keywords Marburg virus (MARV) · VP40 matrix protein · Natural antiviral source · Phytochemicals · Structure-based molecular dynamics · DFT

Introduction

The Marburg virus (MARV) is a highly pathogenic filovirus that causes Marburg virus disease (MVD), a severe hemorrhagic fever with a high fatality rate, ranging from 24 to 88%, depending on the outbreak and the availability of healthcare interventions¹. MARV is closely related to the Ebola virus (EBOV), sharing similar transmission mechanisms, clinical manifestations, and genetic features. The virus is primarily transmitted through contact with infected bodily fluids, contaminated surfaces, or direct exposure to infected non-human primates or fruit bats, which serve as natural reservoirs². The disease is characterized by high fever, vascular leakage, multi-organ failure, and severe hemorrhagic manifestations³. The lack of effective antiviral treatments or approved vaccines underscores the urgent need for novel therapeutic interventions to combat this deadly pathogen.

Among the viral proteins, VP40 is the major matrix protein of MARV and plays a critical role in virion assembly, membrane trafficking, oligomerization, and viral budding from the host cell membrane⁴. As VP40 is essential for the late stage of the viral life cycle, disruption of its structural or functional activity may impair viral particle formation and release⁵. Therefore, VP40 represents a promising molecular target for structure-based antiviral drug discovery.

The development of antiviral agents against MARV has been hindered by the high biosafety level requirements for laboratory studies, limiting experimental approaches to high-containment facilities. In this context, computational drug discovery approaches, including molecular docking, virtual screening, and molecular dynamics simulations, have emerged as powerful tools for identifying potential drug candidates with high specificity and binding affinity. These *in silico* approaches allow rapid screening of large compound libraries against viral targets, accelerating the identification of lead molecules with potential therapeutic applications. Previous studies have demonstrated the effectiveness of computational techniques in identifying natural products with antiviral potential, particularly against RNA viruses such as SARS-CoV-2, Dengue virus, and filoviruses^{6–8}. Additionally, Density Functional Theory (DFT) has become an important computational tool for investigating the electronic properties of bioactive compounds, including frontier molecular orbitals, chemical reactivity, and charge-transfer behavior⁹. These descriptors complement molecular docking and molecular dynamics simulations by providing supportive mechanistic insight into ligand stability and interaction potential with biological targets.

Natural products have played a critical role in antiviral drug discovery, offering structurally diverse bioactive compounds with promising pharmacological properties. Among these, medicinal plants have been widely studied for their ability to inhibit viral replication, block viral entry, or modulate host immune responses. *Ocimum basilicum* L., commonly known as basil or sweet basil, is an ornamental and therapeutic herb belonging to the Lamiaceae family^{10,11}. In Bengali, it is referred to as *Babui Tulsi*¹². This annual herb is native to Southeast Asia but is now widely cultivated in tropical and subtropical regions for its culinary and medicinal applications¹³. This herb is rich in bioactive phytochemicals, including flavonoids, phenolic acids, terpenoids, and essential oils, which contribute to its broad pharmacological potential. Traditionally, it has been used for the management of kidney disorders, earache, menstrual irregularities, arthritis, anorexia, colds, and malaria^{14,15}. Its extracts and essential oils have demonstrated antibacterial, antifungal, antioxidant, antiradical, immunomodulatory, and antiviral

properties^{16–21}. Phytochemicals such as ursolic acid, rosmarinic acid, eugenol, linalool, and related constituents have been reported to inhibit viral replication, interfere with viral protein synthesis, and modulate host immune responses, making *O. basilicum* an attractive candidate for antiviral drug discovery²².

This study aims to investigate the antiviral potential of bioactive compounds from *O. basilicum* against the Marburg virus using an integrated computational approach. Molecular docking was conducted to predict the binding affinity of selected phytochemicals with key MARV proteins, primarily focusing on the matrix protein VP40 as the main antiviral target. The most promising ligand–protein complexes were further analyzed through molecular dynamics simulations to assess their stability, interaction profiles, and binding energies. Additionally, pharmacokinetic and toxicity assessments were performed to evaluate the drug-likeness properties of the identified compounds. The findings of this study may provide valuable insights into the development of plant-derived antiviral agents with potential therapeutic applications against the Marburg virus and other emerging viral infections.

Materials and methods

Sequence retrieval and analysis

The amino acid sequence of Marburg virus (MARV) VP40 protein (Accession ID: APQ46231.1) was obtained from the NCBI database in FASTA format. The VP40 sequence was also compared with sequences of Chikungunya, Ebola, and Monkeypox viruses to evaluate evolutionary relationships, revealing the highest sequence identity with Ebola virus VP40.

Homology modeling

Although experimentally determined structures of homologous filoviral VP40 proteins are available, no complete high-resolution structure of the specific Marburg virus VP40 sequence used in this study was available in a directly usable form. Homology modeling was therefore performed with the SWISS-MODEL server, using the most suitable template selected on sequence identity, GMQE score and structural quality. Among the available homologous VP40-related structures, template 5B0V.1 showed the highest sequence similarity and the most favourable structural quality²³. Template-based modeling was chosen in preference to AlphaFold2 and comparable deep-learning predictors for three reasons. First, a close experimental template exists. Marburg virus VP40 has been solved crystallographically and deposited as Protein Data Bank entry 5B0V, and that structure shares very high sequence identity with the target sequence. In this regime comparative modeling and deep-learning prediction converge on essentially the same backbone, and comparative modeling has the advantage of an explicit and traceable link to an experimental structure. Second, the feature that matters for the present work is the geometry of the flat C-terminal basic patch that mediates membrane association and oligomerisation in the Marburg lineage. That surface is directly resolved in the template. A de novo prediction would reproduce it only indirectly and would still require the same stereochemical validation before use. Third, deep-learning predictors return a single static conformer with per-residue confidence estimates but no companion experimental structure, so the comparative measures used here, namely RMSD to the template, QMEAN, the ProSA Z-score, Verify3D and ERRAT, would lose their point of reference. The model was in any case subjected to the full set of stereochemical and energetic checks described below, and its quality scores fall within the range expected of experimentally resolved structures of similar size. We recognise that a deep-learning model would be the preferred starting point for filoviral targets that lack a close template, and that a side-by-side comparison of an AlphaFold2 model of the same sequence with the present template-based model would be a useful extension of this work.

Model evaluation and energy minimization

The qualities of the modeled structures were evaluated by Ramachandran plot analysis using the SAVES v6.0 server (UCLA-DOE Institute), which validated the stereochemistry and backbone conformations of the homology model²⁴. To further refine the structural geometry, energy minimization was carried out using SwissPdb-Viewer (DeepView) version 4.1.0 with the GROMOS96 force field. This energy minimization step removed steric clashes, optimized bond angles, and improved the stability of the modeled protein structure²⁵.

Literature survey and compound selection

A thorough literature search was carried out to identify the phytochemical constituents of *Ocimum basilicum*. A total of 156 unique small molecules were retrieved from the compiled literature reported from the species. In addition, Laurifolin, a previously reported antiviral compound investigated in filoviral computational studies, was included as a reference ligand to provide a comparative benchmarking framework for docking validation²⁶. All the compounds were cross-checked against the PubChem database <https://pubchem.ncbi.nlm.nih.gov/> to confirm chemical identity as well as to retrieve standard structure files for further processing Table 4. The reported compounds were downloaded in SDF format and converted into PDBQT format for docking studies.

Active site prediction

The binding pocket predicted using CASTp v3.0 (<http://sts.bioe.uic.edu/castp/calculation.html>) was further evaluated for biological relevance by comparing it with reported functional regions of Marburg virus VP40. To avoid blind docking, the precise amino acid residues involved in protein–ligand interactions were identified using CASTp-based pocket mapping. The selected pocket includes residues implicated in membrane association and oligomerization, indicating its potential role as a functionally relevant interface. Key interacting residues identified in docking and molecular dynamics analyses further support the functional significance of the predicted active site. The predicted pocket overlaps with reported VP40 membrane-binding and oligomerization interface regions, confirming its biological relevance for docking studies.

Molecular docking analysis

Molecular docking was performed using AutoDock Vina version 1.2.0 to predict the binding affinity of selected bioactive ligands with MARV VP40 proteins. The protein structures were prepared using MGLTools version 1.5.7 by optimizing protonation states, adding Kollman charges, and incorporating polar hydrogens to improve electrostatic interactions. A rectangular grid box was defined around the active site of each protein to ensure optimal ligand positioning during docking simulations^{27,28}. The target protein was loaded in Pyrx 0.8 and converted to macromolecules. All the phytochemicals were converted into a pdbqt file after energy minimization. The docking grid center coordinates were defined as X=30.6590, Y=1.2180, and Z=33.0220, and the grid box dimensions were set to 25 × 25 × 25 Å to ensure complete coverage of the predicted active-site pocket. The docking grid box was defined based on the active site residues predicted by CASTp v3.0. The center coordinates and grid dimensions were selected to cover the predicted binding pocket fully, ensuring inclusion of all key functional residues involved in ligand interaction. This setup enabled sufficient conformational sampling of ligands within the biologically relevant active-site region. The docking results were then filtered based on binding affinity, and all the generated possible docked conformations were saved in CSV format. Only the conformations that specifically interacted with the active-site residues of the target protein were selected for further analysis. The docking results were ranked based on binding energy (ΔG kcal/mol), hydrogen bonding, and non-covalent interactions²⁹. To ensure reproducibility, molecular docking was performed in multiple independent runs using AutoDock Vina with identical grid parameters and increased exhaustiveness. Reproducibility assessment of molecular docking

through independent triplicate docking runs is illustrated in Supplementary Table 1. Consistent binding poses and stable ranking of top ligands were observed across all runs, confirming the reliability of the docking protocol. From the initial set of 156 phytochemicals, compounds were prioritized based on molecular docking results against VP40. The top 20 compounds showing the most favorable binding affinities and stable interaction profiles with active site residues were selected for further detailed analyses, including ADMET evaluation, molecular dynamics simulations, and DFT calculations. These are listed in Table 1. Two independent limitations preclude the standard external validation routes for this target. First, no co-crystallised MARV VP40-ligand complex has been deposited, so there is no crystallographic reference pose against which a classical redocking RMSD could be measured. Second, no small molecule has been experimentally confirmed as a MARV VP40 inhibitor, so no labelled set of actives exists from which to construct the positive class that a ROC/AUC enrichment calculation requires; any enrichment curve computed here would be defined by the very scoring function it purports to test. The docking protocol was therefore validated in three ways that are well defined for this system. (i) Reproducibility: independent triplicate docking runs with identical grid parameters and increased exhaustiveness (Supplementary Table 1), together with pose-overlap analysis (Supplementary Figure S2). (ii) Pose recovery against an independent reference: rutin was removed from five snapshots taken from the equilibrated portion of the 500 ns molecular dynamics trajectory (100, 200, 300, 400 and 500 ns) and redocked into the corresponding rigid receptor conformation using the published grid, with three random seeds and exhaustiveness values of 8 (the setting used throughout this work) and 32. The symmetry-corrected heavy-atom RMSD between the docked pose and the molecular dynamics reference pose was computed in place, without superposition. Because rutin relaxed 2.28 Å away from the original docked pose during the simulation, these reference poses were not themselves generated by the docking protocol. (iii) Site specificity: rutin was docked into the CASTp-predicted pocket and into five decoy boxes of identical dimensions centred on C α atoms lying at least 18 Å away on the protein surface. The outcomes of (ii) and (iii) are presented in Supplementary Figure S4 and described in the Results.

ADMET screening

The pharmacokinetic properties and drug-likeness of the selected bioactive compounds were assessed using SwissADME and pkcsms <https://biosig.lab.uq.edu.au/pkcsms/prediction>. The ADMET screening evaluated absorption, distribution, metabolism, excretion, and toxicity (ADMET) profiles to ensure the selected ligands met drug development criteria³⁰. Drug-likeness properties were analyzed using Lipinski's rule of five. Lipinski's rule states that an oral drug should have no more than five hydrogen bond donors, ten hydrogen bond acceptors, a molecular weight of ≤ 500 Da, and a logP of ≤ 5 . Veber's rule specifies that a drug should have no more than ten rotatable bonds and a total polar surface area (TPSA) below 140 Å².

Molecular dynamics simulation analysis

Stability of protein–ligand interaction on macromolecule structure-to-function transitions was explored using molecular dynamics. The Desmond Molecular Dynamics package (Schrödinger Release 2023–1, Schrödinger

Table 1 Standardized parameters for molecular docking reproducibility.

Parameter	Value / Setting
Docking engine	AutoDock Vina 1.2.0
Search exhaustiveness	8 (Standard for Vina 1.2.0)
Number of modes	9 (Maximum generated poses)
Grid center (X, Y, Z)	30.6590, 1.2180, 33.0220
Grid dimensions (Å)	25 × 25 × 25
Force field	Kollman Charges (Protein); Gasteiger Charges (Ligands)
Protein preparation	MGLTools 1.5.7 (Protonation & Hydrogen addition)
Ligand minimization	UFF (Universal Force Field) via PyRx 0.8

LLC, New York, NY, USA) was used to perform an all-atom molecular dynamics (MD) simulation of the VP40-rutin complex for 500 ns. The simulations, based on Newton's classical equations of motion, followed the trajectories of atoms over time. Preprocessing of the receptor-ligand complex was done using Maestro's Protein Preparation Wizard, which consisted of optimization and minimization protocols. The system was constructed using the System Builder tool, using the Transferable Intermolecular Interaction Potential 3 Points (TIP3P) solvent model in an orthorhombic box of $90.25 \times 58.65 \times 68.42$ Å containing 10,604 water molecules. The simulation was managed by the OPLS_2005 force field, and 69 counterions were added to make the model neutral. A 0.15 M sodium chloride (NaCl) solution was added to simulate in vivo conditions in the body. The simulations were performed in the Number of particles (N), Pressure (P), and Temperature (NPT) ensemble at a temperature of 300 K and pressure of 1 atm. The models were equilibrated before the simulation. Trajectory frames were written every 1 ns, yielding 501 frames for analysis. All reported quantities were computed directly from the production trajectory and the accompanying force-field parameters: backbone root mean square deviation (RMSD) and per-residue root mean square fluctuation (RMSF) by least-squares superposition; radius of gyration; solvent-accessible surface area (SASA) by the Shrake-Rupley algorithm; hydrogen bonds and non-bonded contacts by geometric criteria; the protein-ligand non-bonded interaction energy by Coulomb and Lennard-Jones summation; and principal component analysis (PCA) together with the dynamic cross-correlation matrix (DCCM) from the C α coordinates^{31–34}.

Before production analysis, the system was subjected to standard relaxation and equilibration protocols implemented in the Desmond workflow. Trajectory convergence was assessed by monitoring time-dependent RMSD profiles of the protein backbone and ligand, together with residue-wise RMSF behaviour and the radius of gyration. A stable RMSD plateau without systematic drift was considered indicative of equilibration and convergence (Fig. 6). The backbone RMSD rose over the first 15–20 ns and thereafter remained on a flat plateau, with a standard deviation of 0.15 Å across the equilibrated portion of the trajectory and 0.10 Å over the final 50 ns, comfortably below the 0.3 Å criterion adopted here. The residual upward slope of the plateau (+0.04 Å per 100 ns) amounts to less than 0.2 Å over the whole production run. Convergence was further quantified by block averaging of the backbone RMSD in consecutive 50 ns windows, by a Flyvbjerg-Petersen blocking analysis combined with the integrated autocorrelation time to obtain a correlation-corrected standard error, by comparison of the RMSD distributions of the two trajectory halves, and by the cosine content of the leading principal components (Supplementary Figure S3). The full 500 ns trajectory was therefore analysed as the production ensemble.

The production length of 500 ns was selected on three grounds. It exceeds by roughly fivefold the 50 to 100 ns commonly reported for filoviral VP40 complexes, which makes the present trajectory longer than, and directly comparable with, the published studies summarised in Table 14. It is long enough for the scalar observables on which the conclusions rest, namely the backbone RMSD, the radius of gyration, the solvent-accessible surface area, the contact occupancies and the interaction energy, to reach stationary values, as verified by block averaging and by the Flyvbjerg-Petersen blocking analysis. And it is long relative to the integrated autocorrelation time of the backbone RMSD, which is 21 ns, so the run yields roughly 23 effectively independent samples and a correlation-corrected standard error of 0.03 Å on the mean.

The 500 ns window is not, however, sufficient for the slowest collective motions. The cosine content of the two leading principal components is 0.79 and 0.57, close to the values expected of free diffusion, which shows that the essential subspace is still being explored rather than sampled at equilibrium. Additional simulation time would improve the convergence of these slow modes. For a diffusive mode the cosine content of the leading component falls roughly in inverse proportion to the length of the trajectory, so a run of the order of several microseconds would be needed before the principal components could be read as resolved conformational substates. An ensemble of independent replicas started from different velocity seeds would be the more efficient route to the same end, since it also provides a direct estimate of the between-run variance. Calculations on that scale lay outside the computational budget of the present study and are proposed as future work. They would not be expected to alter the scalar observables reported here, which are already stationary, but they would be necessary before any claim could be made about ligand-induced modulation of the slow collective dynamics of VP40.

Binding free energy calculation (MM/GBSA and MM/PBSA)

To provide a more accurate estimation of the binding affinity than traditional docking scores, end-state binding free energies were computed using both the Molecular Mechanics Generalized Born Surface Area (MM/GBSA) and the Molecular Mechanics Poisson-Boltzmann Surface Area (MM/PBSA) methods, following a single-trajectory protocol in which the complex, receptor, and ligand ensembles are all extracted from the trajectory of the complex. The binding free energy (ΔG_{bind}) was calculated using the following equation:

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{solv} + \Delta G_{SA}$$

where ΔE_{MM} is the gas-phase molecular-mechanics interaction energy (the sum of the van der Waals and electrostatic terms), ΔG_{solv} is the polar solvation free energy, obtained from the Generalized Born model for MM/GBSA and from numerical solution of the Poisson-Boltzmann equation for MM/PBSA, and ΔG_{SA} is the non-polar solvation free energy, estimated from the solvent-accessible surface area. Calculations were performed with gmx_MMPBSA version 1.6.5 built on AmberTools (CPPTRAJ 6.18.1), using a single-trajectory protocol in which the complex, receptor, and ligand ensembles are all extracted from the trajectory of the complex. Snapshots were taken from the production trajectory between frames 1 and 502 at an interval of 10 frames, giving 51 snapshots spaced 10 ns apart and spanning the full 500 ns. The Generalized Born calculation used the OBC model (igb=5) with a salt concentration of 0.150 M; the Poisson-Boltzmann calculation used an ionic strength of 0.150 M, interior and exterior dielectric constants of 1.0 and 80.0, respectively, radiopt=0 and fillratio=4.0. The reference temperature was set to 310 K; because conformational entropy was not computed, this setting does not enter any reported quantity. Conformational entropy contributions $-T\Delta S$ were not calculated via normal mode analysis or quasi-harmonic analysis, owing to the high computational expense and the known tendency of absolute entropy estimates to introduce substantial statistical noise without improving correlation with experimental binding affinities; the reported values are therefore effective binding free energies suitable for comparative ranking. Per-residue energy decomposition was performed as a separate Generalized Born calculation on the same trajectory using idecomp=2 (1–4 electrostatic and van der Waals terms folded into the electrostatic and van der Waals contributions respectively), dec_verbose=3 and print_res="within 5", which restricts the printed output to residues lying within 5 Å of the ligand; that calculation used frames 1 to 502 at an interval of 20, giving 26 snapshots spaced 20 ns apart. For every quantity, the reported value is the arithmetic mean $\pm$ standard deviation over the sampled snapshots, together with the standard error of the mean and the 95% confidence interval.

DFT analysis

Density functional theory (DFT) calculations were performed using Gaussian 09 software. Geometry optimization of rutin was carried out using the B3LYP functional with the 6-31G(d,p) basis set. Frequency calculations were subsequently performed at the same level of theory to confirm that the optimized structure corresponded to a true minimum energy state with no imaginary frequencies. Frontier molecular orbital energies (HOMO and LUMO) were obtained from the optimized geometry. Following the Parr and Pearson formulation of conceptual density functional theory and Koopmans' theorem, global reactivity descriptors were derived from the frontier orbital energies. These comprise the energy gap (ΔE), the chemical hardness (η) and the softness (S), obtained from the equations below, together with the vertical ionisation potential ($IP = -\epsilon_{HOMO}$), the vertical electron affinity ($EA = -\epsilon_{LUMO}$), the electronegativity ($\chi = (IP + EA)/2$), the chemical potential ($\mu = -\chi$), the electrophilicity index ($\omega = \mu^2/2\eta$) and the maximum charge transfer ($\Delta N_{max} = -\mu/\eta$)^{35,36}.

$$\eta = \frac{(E_{LUMO} - E_{HOMO})}{2}$$

$$S(\text{Softness}) = \frac{1}{\eta}$$

where E_{LUMO} and E_{HOMO} represent the orbital energies of HOMO and LUMO, respectively. Lower hardness and higher softness values generally indicate greater electronic polarizability and chemical reactivity, which may facilitate non-covalent interactions with biological targets.

All calculations were performed on the isolated molecule in the gas phase. Implicit solvation was not applied, and the descriptors reported here should therefore be read as gas-phase quantities. Placing rutin in a polarisable continuum such as the IEF-PCM or SMD water model would stabilise both frontier orbitals, would be expected to widen the gap modestly and to lower the electrophilicity index, and would matter most for the charged and strongly polar hydroxyl and glycosidic groups of the molecule. Because the descriptors are used here only in a comparative and supporting role, and because no comparison is drawn between molecules calculated at different levels of theory, the gas-phase values are adequate for that purpose. A solvent-phase treatment, together with molecular electrostatic potential mapping, natural bond orbital analysis and Fukui functions for site-resolved reactivity, would be required before the electronic descriptors could be used quantitatively, and this is proposed as an extension of the present work.

Results

Homology modeling and structural validation of VP40 protein

A homology model of the VP40 protein (PDB ID: 5B0V, chain A; 258 amino acids) was constructed and subjected to stereochemical and structural validation. Ramachandran plot analysis revealed that 91.9% of residues were in the most favored regions, 7.9% in the additionally allowed regions, and 0.2% in the generously allowed regions, with none in the disallowed regions (Fig. 1 and Table 2). Altogether, 96.78% of the residues were located

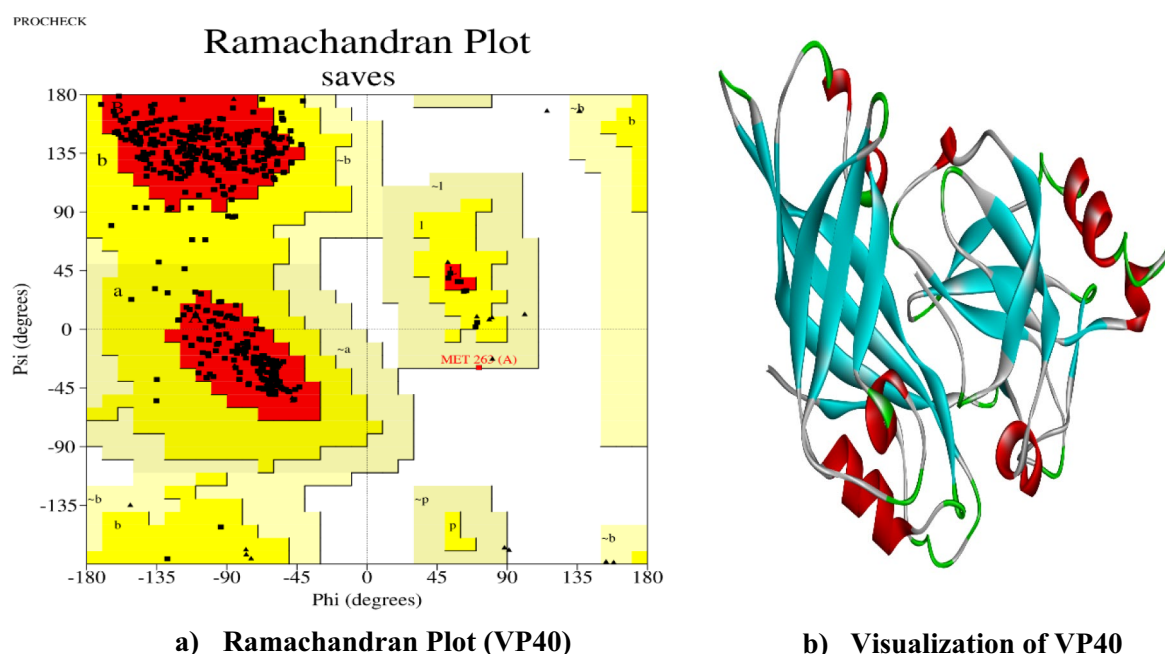
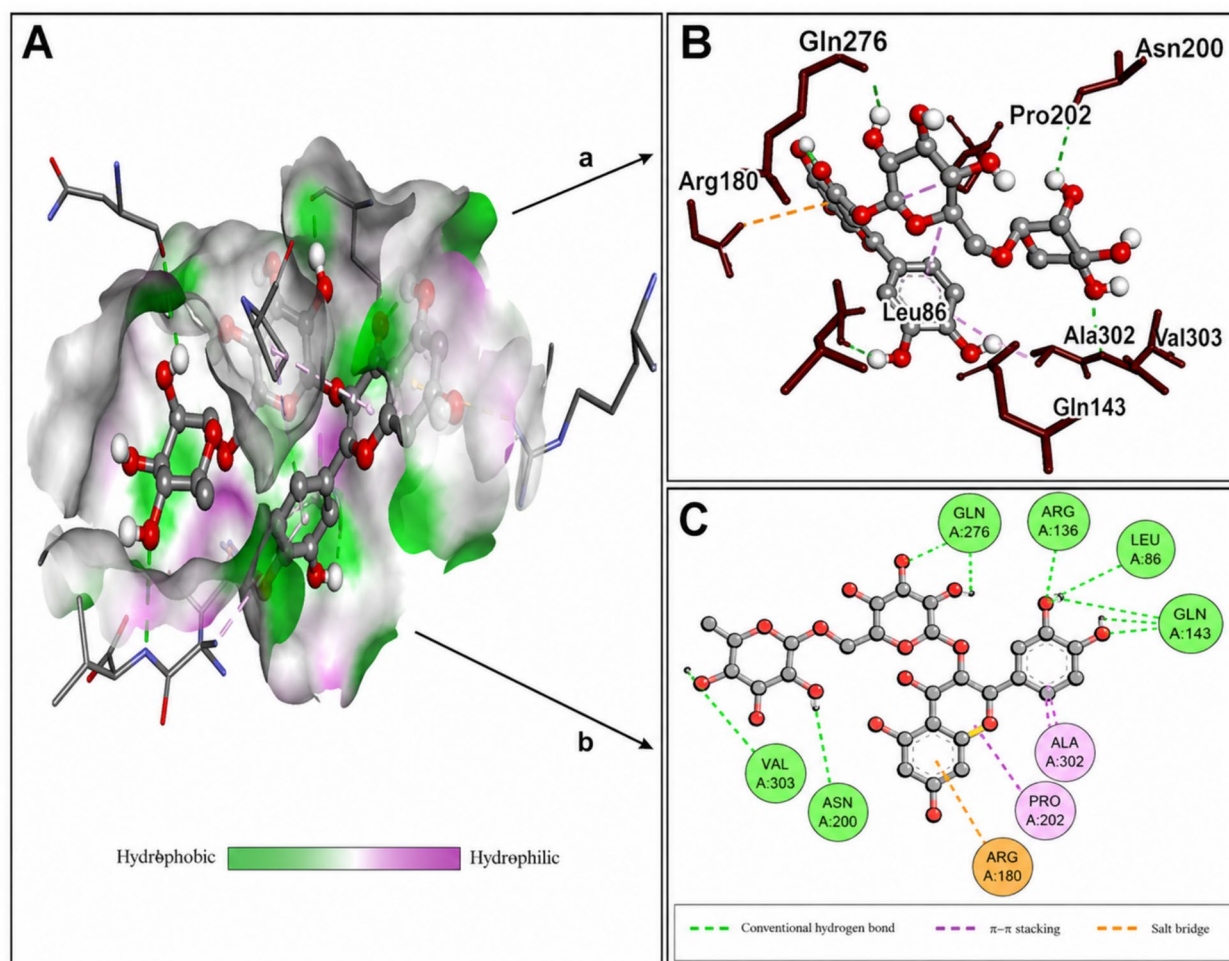


Fig. 1 Ramachandran plot (a) and visualization (b) of the identified protein.

Table 2 Structure corroboration and quality assessment of target proteins.

Protein	Most favoured region (%)	Additional allowed region (%)	Generously allowed region (%)	Disallowed region (%)	Total residues	QMEAN score
VP40	91.9	7.9	0.2	0.0	258	0.88


Fig. 2 Three-dimensional (a) and two-dimensional (b) interaction profiles of the rutin-VP40 complex.

in energetically favorable phi-psi conformational space, confirming the high stereochemical quality and accurate backbone geometry of the predicted model (Fig. 2).

The QMEAN score of 0.88 lies within the range recorded for experimentally determined structures of similar size, indicating that the model has realistic structural features. The root mean square deviation (RMSD) analysis also supported high structural congruence between the model developed for the VP40 protein and its target template, and the template resolution of 2.81 Å is adequate for comparative modeling.

To ensure structural accuracy, the mVP40 homology model was compared with existing filoviral structures. While the NTD shares high structural homology with the Ebola virus VP40 (eVP40) to support conserved dimerization, the CTD exhibits a distinct, flat cationic surface characteristic of the Marburg lineage. This unique basic patch is essential for membrane binding and represents the functionally relevant target for our inhibitor screening. Our model (RMSD analysis supported) maintains high congruence with the experimental template 5B0V, specifically capturing the domain orientations necessary for understanding ligand-induced stabilization.

To further strengthen the structural confidence of the modeled VP40 protein, additional validation analyses were performed using ProSA, Verify3D, and ERRAT. The ProSA analysis yielded a Z-score of -7.99, which falls within the range of experimentally resolved protein structures of similar size. Verify3D assessment demonstrated that approximately 85% of the residues had an average 3D-1D score ≥ 0.2 , confirming favorable compatibility between the modeled structure and its amino acid sequence. Furthermore, the ERRAT analysis showed an overall quality factor of 89.45, reflecting reliable non-bonded atomic interactions within the structure. Taken together, these results support the structural reliability of the predicted VP40 homology model.

To further benchmark the homology model beyond standard validation scores, structural superposition was performed against the primary template 5B0V.1. The analysis yielded a QMEAN score of 0.88, reflecting a model quality comparable to high-resolution experimental structures of similar size. Comparative benchmarking with the homologous Ebola virus VP40 (eVP40) highlighted that our model correctly adopts the ‘flat’ C-terminal domain architecture unique to the Marburg lineage, which is structurally distinct from the more globular C-terminal folds of other filoviruses. This architectural fidelity ensures that the predicted binding pocket, which overlaps with known membrane-binding and oligomerization interfaces, is biologically relevant for drug discovery efforts (Table 3).

These findings indicate that the VP40 homology model has acceptable stereochemical quality and agrees closely with the experimental template, and that it is a suitable framework for the computational studies that follow.

Molecular docking

The molecular docking of the selected 20 phytochemicals from *Ocimum basilicum* and one reference drug namely Laurifolin into the catalytic pocket of MARV VP40 protein demonstrated variable binding affinities ranging from -10.4 to -8.0 kcal/mol. Among the top-ranked compounds, Rutin obtained the most favourable docking score, at -10.4 kcal/mol, which placed it first in the screening rank order. This was closely followed by the (+)-Ursolic acid and Orientin, showing a docking score of -10.1 kcal/mol, hence both compounds have comparable binding potential.

The compounds like Oleanolic acid, Isoquercetin, and Vicenin-2 had a moderate binding affinity ranging from -9.9 to -9.3 kcal/mol. The compounds Oleanolic acid, Isoquercetin, and Vicenin-2 showed docking scores ranging from -9.9 to -9.3 kcal/mol. The reference compound Laurifolin obtained a docking score of -9.9 kcal/mol, comparable to that of Oleanolic acid. Within the present docking workflow, this score indicates its relative position in the screening rank order only and does not establish binding occupancy, inhibitory activity, or a mechanism of action against VP40. Similarly, Stigmasterol, Riboflavin, and Campesterol received comparatively less favourable docking scores, ranging from -8.9 to -8.0 kcal/mol. Within the scoring framework used in this study, these compounds therefore ranked below the higher-scoring phytochemicals and were given lower priority for subsequent computational analyses. These docking scores alone do not establish the strength, stability, or biological relevance of their interactions with VP40 (Table 4).

Flavonoid glycosides such as rutin and triterpenoids such as ursolic acid therefore rank highest against the predicted VP40 pocket under the scoring function used. These scores served a single purpose, which was to

Table 3 Structural validation and comparative quality assessment of the predicted homology model against the reference template and a benchmark ebola VP40 structure.

Parameter	Homology model (This Study)	Template protein data bank: 5B0V	Benchmark protein data bank: 4LAV (Ebola)
Residue Count	258	258	~260–290
Z-score (ProSA)	-7.99	-8.1 (approx.)	-7.5 to -8.5
RMSD to Template	N/A	0.88 Å	~1.5–2.0 Å
Favored Regions	91.90%	92.50%	> 90%

Table 4 Top 10 docked molecules in the catalytic pocket of VP40 protein.

SL NO	Name	Docking score
1	Rutin	-10.4
2	(+)-Ursolic Acid	-10.1
3	Orientin	-10.1
4	Oleanolic Acid	-9.9
5	Laurifolin	-9.9
6	Isoquercetin	-9.7
7	Vicenin-2	-9.3
8	Stigmasterol	-8.9
9	Riboflavin	-8.9
10	4-Hydroxy-3-methoxyestra-1,3,5(10)-trien-17-one	-8.8
11	Luteolin	-8.8
12	(+)-Eriodictyol	-8.7
13	Quercetin	-8.4
14	Apigenin	-8.4
15	Esculin	-8.4
16	2-(3,4-Dihydroxyphenyl)-3,5,7-trihydroxy-1-Benzopyrylium	-8.3
17	Rosmarinic Acid	-8.3
18	Campesterol	-8.2
19	Kaempferol	-8.2
20	Xanthomicrol	-8.1
21	Phytosterols	-8

prioritise compounds for the later stages of the workflow. Because the protocol reproduces its own scores but not the reference pose geometry, as shown in the following section, no inference about occupancy of the pocket or about a mode of inhibition is drawn at this stage. The top-scoring compounds were carried forward for interaction analysis, ADMET profiling and, in the case of rutin, molecular dynamics simulation and end-state free-energy analysis. The reproducibility of docking poses and ligand ranking across independent runs is summarized in Supplementary Figure S2. Supplementary Table 1 and Supplementary Figures S2-S4 are supplied as separate files with this revision and document, respectively, triplicate docking reproducibility, pose consistency, molecular dynamics convergence, and validation of the docking protocol.

Validation of the docking protocol

Because MARV VP40 has neither a co-crystallised ligand complex nor an experimentally confirmed small-molecule inhibitor, neither a classical redocking RMSD nor a ROC/AUC enrichment curve can be defined for this target (see Methods). The docking protocol was therefore assessed against reference poses drawn from the independent 500 ns molecular dynamics trajectory, together with a binding-site specificity control (Supplementary Figure S4).

The predicted affinity proved highly reproducible. Across five equilibrated receptor snapshots, three random seeds and two exhaustiveness settings, the top-ranked Vina affinity was -10.37 ± 0.33 kcal/mol, in close agreement with the reported docking score of -10.4 kcal/mol, and independent triplicate runs on the original receptor differed by no more than 0.06 kcal/mol (Supplementary Table 1). The predicted pocket was also strongly favoured over the remainder of the protein surface: rutin scored -10.74 kcal/mol in the CASTp-defined pocket versus -6.35 ± 0.47 kcal/mol in five decoy boxes of identical dimensions placed at least 18 Å away ($\Delta = 4.4$ kcal/mol; one-sample t test, $p = 5 \times 10^{-5}$). This supports the choice of binding site, although it remains computational rather than experimental evidence.

Pose recovery, in contrast, was poor. The top-ranked pose never fell within the conventional 2 Å success threshold of the molecular dynamics reference pose (mean heavy-atom RMSD 5.59 ± 2.22 Å; 0 of 30 runs), and the closest of the nine returned poses reached only 2.74 ± 0.72 Å, with just 2 of 30 runs producing any pose within

2 Å. Increasing the exhaustiveness from 8 to 32 changed the mean RMSD by less than 0.01 Å, so the shortfall does not stem from insufficient search effort alone; moreover, the pose closest to the reference was returned at a median rank of 6 out of 9, showing that the scoring function does not rank it first. As an internal control, when redocked into the original energy-minimised complex the protocol reproduced its own starting pose to 2.25 Å.

These findings bear directly on how the docking output should be interpreted. The perfect run-to-run reproducibility of the docking scores (standard deviation 0.000 kcal/mol for rutin across triplicate runs) measures the precision of the conformational search, not the accuracy of the resulting pose: the same non-native pose is returned reproducibly. Rutin is a large flavonoid glycoside with sixteen rotatable bonds, a regime in which empirical scoring functions are known to rank poses unreliably. Accordingly, the docking calculations in this study serve only as a coarse prioritisation filter, and the mechanistic conclusions rest on the 500 ns molecular dynamics simulation and the MM/GBSA and MM/PBSA analyses rather than on the static docked pose. Two caveats apply to this assessment: the reference is a force-field-relaxed molecular dynamics pose rather than an experimental structure, and the receptor was held rigid within each snapshot.

Physicochemical, ADME(T) and drug-likeness properties

Physicochemical profiles of the selected phytochemicals showed a wide chemical space of flavonoid glycosides, aglycones, triterpenoids, sterols, and phenolic acids, as shown in Table 5. Molecular weight of the selected phytochemicals varied from 270.24 g/mol of Apigenin to 610.52 g/mol of Rutin.

Flavonoid glycosides Rutin, Vicenin-2, Isoquercetin, and Orientin have been characterized by their high hydrogen bond donor ($\text{HBD} \geq 8$) and hydrogen bond acceptor ($\text{HBA} \geq 11$) values and their very high topological polar surface area ($\text{TPSA} > 200 \text{ Å}^2$). Rutin, in particular, showed the highest TPSA (269.43 Å^2) and HBD/HBA values (10/16), resulting in three Lipinski rule violations, consistent with its glycosidic nature. On the other hand, aglycone flavonoids such as Apigenin, Luteolin, Quercetin, Kaempferol, and Eriodictyol complied with Lipinski's rule on all four parameters, and these molecules also showed moderate lipophilicity, with LogP values ranging from 1.6 to 1.9. Moreover, TPSA was within an optimal range for oral bioavailability, i.e., between 90 and 130. Triterpenoids and phytosterols, including molecules such as Ursolic acid, Oleanolic acid, Stigmasterol, Campesterol, exhibited high lipophilicity ($\text{LogP} \approx 4.9\text{--}5.1$) with minimal polar surface areas, favouring membrane permeability but suggesting solubility limitations.

Caco-2 permeability and intestinal absorption were strongly influenced by the physicochemical properties of rutin and related flavonoids. The extremely poor absorption of Rutin (0.00%) with low Caco-2 permeability at -2.132 was observed to result from its extreme polarity and its high molecular weight. The absorption limitation was observed for Vicenin-2 and Orientin.

Meanwhile, Apigenin (93.25%), Oleanolic acid (99.56%), Ursolic acid (100%), and Stigmasterol (94.97%) showed high predicted human intestinal absorption, reflecting favourable lipophilicity and low TPSA. The permeability of the BBB was generally low for the polyphenolic compounds, while sterols, represented by Stigmasterol, Campesterol, and Phytosterols, showed positive values of ~ 0.77 for BBB permeability, suggesting that they could be exposed to the CNS. Significantly, none of the compounds studied displayed hERG inhibition, a characteristic indicating low cardiotoxicity. The hepatotoxicity prediction list identified the potential issues with Ursolic acid, Oleanolic acid, Riboflavin, and Esculin. Rutin was identified as not being hepatotoxic, even though it has poor absorption. All compounds fall in rat acute toxicity classes II to III, as listed in Table 6.

Molecular interaction analysis

The interaction patterns reported in this section are those of the static docked poses. They are presented to document the output of the screening stage and to allow comparison with the reference ligand Laurifolin. As set out in the preceding section, the docking protocol does not reproduce the reference pose geometry to within 2 Å, so these contacts describe the pose returned by the scoring function rather than a determined binding mode. The

Table 5 Physicochemical and drug-likeness properties of the selected compounds.

Compound	LogP	HBD	HBA	Molar Refractivity	TPSA (Å ²)	Lipinski Rule	Water Solubility	Molecular weight (g/mol)
Rutin	0.46	10	16	141.38	269.43	No ; 3 violations	-2.892	610.52
(+)-Ursolic Acid	3.95	1	3	136.91	57.53	Yes; 1 violation	-3.072	456.70
Orientin	1.00	8	11	108.63	201.28	No; 2 violations	-2.905	448.38
Oleanolic Acid	3.94	2	3	136.65	57.53	Yes; 1 violation	-3.261	456.70
Isoquercetin	0.94	8	12	110.16	210.51	No; 2 violations	-2.925	464.38
Vicenin-2	1.73	11	15	139.23	271.2	No; 3 violation	-2.844	594.52
Stigmasterol	5.08	1	1	132.75	20.23	Yes; 1 violation	-6.682	412.69
Riboflavin	1.63	5	8	96.99	161.56	Yes; 0 violation	-2.029	376.36
4-Hydroxy-3-methoxyestra-1,3,5(10)-trien-17-one	2.87	1	3	86.56	46.53	Yes; 0 violation	-3.830	300.39
Luteolin	1.86	4	6	76.01	111.13	Yes; 0 violation	-3.094	286.24
(+)-Eriodictyol	1.62	4	6	73.59	107.22	Yes; 0 violation	-3.253	288.25
Quercetin	1.63	5	7	78.03	131.36	Yes; 0 violation	-2.925	302.24
Apigenin	1.89	3	5	73.99	90.9	Yes; 0 violation	-3.329	270.24
Esculin	1.33	5	9	78.65	149.82	Yes; 0 violation	-2.031	340.28
2-(3,4-Dihydroxyphenyl)-3,5,7-trihydroxy-1-Benzopyrylium	-2.62	5	6	76.17	114.29	Yes; 0 violation	-2.935	287.24
Rosmarinic Acid	1.48	5	8	91.4	144.52	Yes; 0 violation	-3.059	360.31
Campesterol	4.97	1	1	128.42	20.23	Yes; 1 violation	-6.818	400.68
Kaempferol	1.70	4	6	76.01	111.13	Yes; 0 violation	-3.040	286.24
Xanthomicrol	2.99	2	7	91.44	98.36	Yes; 0 violation	-3.377	344.32
Phytosterols	5.05	1	1	133.23	20.23	Yes; 1 violation	-6.773	414.71

binding mode used for every mechanistic statement in this work is the molecular dynamics ensemble described later. Several of the docked contacts are not maintained under dynamic, solvated conditions, the pi-cation interaction with ARG180 being the clearest example, and the docked contact list should be read with that in mind.

In the docked poses, both ursolic acid and rutin form multiple contacts with residues of the predicted VP40 pocket. Ursolic acid formed five hydrogen bonds with various amino acid residues of VP40, ranging from 2.201 Å to 3.773 Å. Conventional hydrogen bonds were formed between ursolic acid and ARG136 (3.211 Å), PRO134 (2.461 Å), PRO298 (2.201 Å), and TYR274 (2.619 Å). Additionally, GLY87 formed a carbon-hydrogen bond at a distance of 3.773 Å. The shortest of these distances are those to PRO298 and PRO134, as presented in Table 7 and illustrated in Fig. 3. Rutin demonstrated the most diverse interaction profile with VP40. It formed four conventional hydrogen bonds with VAL303 (2.879 Å), LEU86 (2.324 Å), GLN276 (1.816 Å), and ASN200 (2.800 Å). In addition, rutin exhibited an electrostatic pi-cation interaction with ARG180 at 4.617 Å, together with hydrophobic pi-alkyl interactions involving PRO202 (4.681 Å and 5.244 Å) and ALA302 (5.223 Å). Rutin therefore

Table 6 Pharmacokinetic and toxicological (ADMET) profiles of the selected compounds.

Compound	CaCO ₂ Perm	Intes- tinal Abs. (%)	BBB Perm	CNS Perm	Metabolism	Clearance	Rat Acute Toxicity	Hepatotoxicity	hERG Inhib
Rutin	-2.13	23.45	-14.51	-5.56	No	-45.64	2.48	No	No
(+)-Ursolic Acid	1.17	100.00	-0.14	-1.19	No	0.08	2.35	Yes	No
Orientin	-1.25	43.73	-1.64	-4.02	No	0.37	2.57	No	No
Oleanolic Acid	1.17	99.56	-0.14	-1.18	No	-0.08	2.20	Yes	No
Isoquercetin	0.24	47.90	-1.69	-4.09	No	0.39	2.54	No	No
Vicenin-2	-1.13	14.66	-1.93	-4.77	No	-0.11	2.48	No	No
Stigmasterol	1.21	94.97	0.77	-1.65	No	0.62	2.54	No	No
Riboflavin	-0.68	36.11	-1.55	-4.14	No	0.70	2.17	Yes	No
4-Hydroxy-3-methoxyestra-1,3,5(10)-trien-17-one	1.34	96.32	-0.01	-1.91	No	0.74	2.42	No	No
Luteolin	0.10	81.13	-0.91	-2.25	No	0.50	2.46	No	No
(+)-Eriodictyol	-0.09	74.69	-0.83	-3.14	No	-0.01	2.03	No	No
Quercetin	-0.23	77.21	-1.10	-3.07	No	0.41	2.47	No	No
Apigenin	1.01	93.25	-0.73	-2.06	No	0.57	2.45	No	No
Esculin	0.28	46.09	-1.26	-3.89	No	0.65	2.30	Yes	No
2-(3,4-Dihydroxyphenyl)-3,5,7-trihydroxy-1-Benzopyrylium	-0.35	87.30	-1.23	-2.22	No	0.53	2.46	No	No
Rosmarinic Acid	-0.94	32.52	-1.38	-3.35	No	0.25	2.81	No	No
Campesterol	1.21	94.76	0.77	-1.81	No	0.57	2.28	No	No
Kaempferol	0.03	74.29	-0.94	-2.23	No	0.48	2.45	No	No
Xanthomicrol	0.52	96.28	-0.58	-3.30	Yes	0.66	2.34	No	No
Phytosterols	1.20	94.46	0.78	-1.71	No	0.63	2.55	No	No

Table 7 Binding interactions of VP40 and Ursolic acid, Rutin and Laurifolin.

Compound name	Amino acid	Distance	Categories	Types
Ursolic acid with VP40	ARG136	3.211	Hydrogen Bond	Conventional Hydrogen Bond
	PRO134	2.461	Hydrogen Bond	Conventional Hydrogen Bond
	PRO298	2.201	Hydrogen Bond	Conventional Hydrogen Bond
	TYR274	2.619	Hydrogen Bond	Conventional Hydrogen Bond
Rutin with VP40	GLY87	3.773	Hydrogen Bond	Carbon Hydrogen Bond
	VAL303	2.879	Hydrogen Bond	Conventional Hydrogen Bond
	LEU86	2.324	Hydrogen Bond	Conventional Hydrogen Bond
	GLN276	1.816	Hydrogen Bond	Conventional Hydrogen Bond
	ASN200	2.800	Hydrogen Bond	Conventional Hydrogen Bond
	ARG180	4.617	Electrostatic	Pi-Cation
	PRO202	4.681	Hydrophobic	Pi-Alkyl
	PRO202	5.244	Hydrophobic	Pi-Alkyl
Laurifolin with VP40	ALA302	5.223	Hydrophobic	Pi-Alkyl
	GLN276	1.834	Hydrogen Bond	Conventional Hydrogen Bond
	VAL205	1.876	Hydrogen Bond	Conventional Hydrogen Bond
	ARG180	4.067	Electrostatic	Pi-Cation
	GLY87	3.975	Hydrophobic	Pi-Sigma
	PRO202	4.545	Hydrophobic	Pi-Alkyl

returns the most varied contact list of the compounds examined, although the persistence of these contacts is assessed only in the simulation reported below. The detailed three-dimensional and two-dimensional interaction patterns of the rutin-VP40 complex are shown in Fig. 2, while the interacting residues and bond categories are summarized in Table 7.

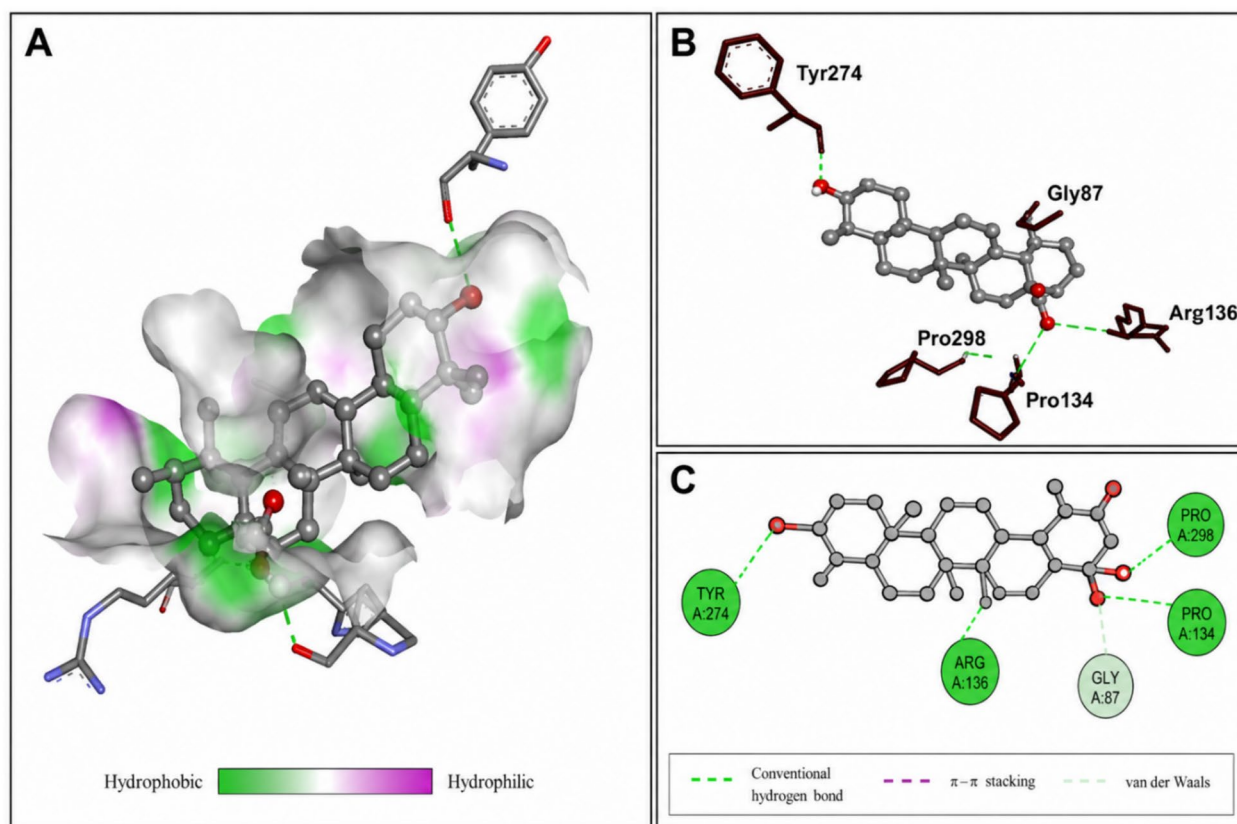


Fig. 3 Three-dimensional (a) and two-dimensional (b) interaction profiles of the ursolic acid-VP40 complex.

Similarly, the reference compound Laurifolin also showed significant interactions with the VP40 binding pocket through a combination of hydrogen bonding, electrostatic, and hydrophobic interactions. Laurifolin formed two conventional hydrogen bonds with GLN276 (1.834 Å) and VAL205 (1.876 Å), along with an electrostatic pi-cation interaction with ARG180 at 4.067 Å. Furthermore, hydrophobic interactions including pi-sigma and pi-alkyl interactions were observed with GLY87 (3.975 Å) and PRO202 (4.545 Å), respectively. The molecular interaction patterns of Laurifolin with VP40 are illustrated in Fig. 4 and summarized in Table 7. Both rutin and Laurifolin contact GLN276, ARG180 and PRO202 in their docked poses, and rutin returns the larger number of hydrogen bonds. These differences are differences between two poses generated by the same scoring function, and they are recorded here without any claim that they predict relative inhibitory potency. Rutin was taken forward for simulation because it ranked first in the screen, and the comparison with Laurifolin is offered as a benchmark for the screening stage only (Tables 8 and 9).

Molecular dynamics simulation

To further evaluate the dynamic stability of the docked VP40-rutin complex under near-physiological conditions, a 500 ns all-atom molecular dynamics simulation was performed using the Desmond simulation package, producing 501 evenly spaced trajectory frames. Structural stability, conformational flexibility, compactness, solvent exposure, persistence of ligand–protein interactions, non-bonded interaction energy, and collective motions were assessed through RMSD, RMSF, radius of gyration, SASA, contact occupancy, interaction-energy, PCA, and DCCM analyses. Summary statistics for every trajectory-derived metric are collected in Table 10.

The RMSD trajectory of the VP40-rutin complex exhibited an initial adjustment phase over the first 15–20 ns of the simulation, followed by stabilization into a sustained plateau, indicating attainment of dynamic equilibrium.

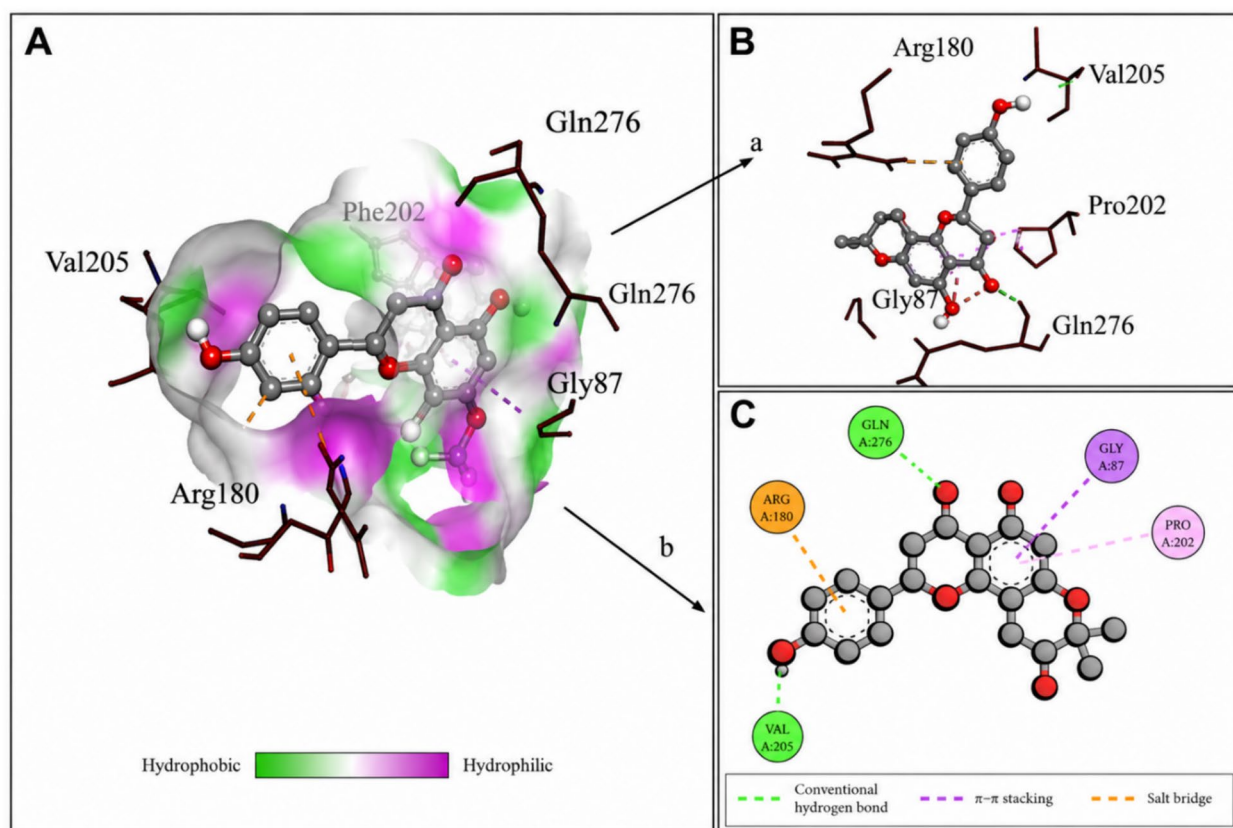


Fig. 4 Three-dimensional (a) and two-dimensional (b) interaction profiles of the Laurifolin-VP40 complex.

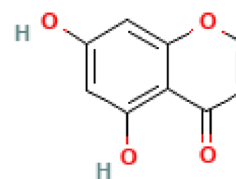
Table 8 Top-ranked phytochemicals from *Ocimum basilicum* showing functional groups and reported biological activities based on literature and docking prioritization.

Sl. No	Compound name	Molecular formula	Major functional groups	Reported biological activities	Reference
1	Rutin	$C_{27}H_{30}O_{16}$	Hydroxyl groups (-OH), Ether linkage (-O-), Carbonyl group (C=O), Sugar moieties	Antibacterial, antiprotozoal, antitumor, anti-inflammatory, antiallergic, antiviral, cytoprotective, vasoactive, hypolipidaemic, antiplatelet, antispasmodic, antihypertensive, anthelmintic, antifungal, anticancer, antioxidant, antidiabetic, neuroprotective, cardioprotective, hepatoprotective, nephroprotective, haematoprotective, antiarthritis, protective against testicular toxicity	37,38
2	Ursolic Acid	$C_{30}H_{48}O_3$	Carboxylic acid (-COOH), Hydroxyl group (-OH), Alkene (C=C)	Anti-inflammatory, antitumor, anticancer, antidiabetic, neuroprotective, antiproliferative, antiviral, antimicrobial, hepatoprotection, inhibits lipoxygenase and cyclooxygenase in human leukaemia HL-60 cells	39–44

After equilibration, backbone RMSD values fluctuated within a narrow range (standard deviation 0.15 Å) without appreciable progressive drift, supporting structural convergence of the protein system. Ligand RMSD similarly stabilized after early conformational adaptation, consistent with persistent occupancy of the binding pocket throughout the simulation. Residue-level RMSF analysis further supported convergence, as fluctuations were mainly confined to solvent-exposed loop and terminal regions, whereas residues constituting the binding pocket remained comparatively rigid during the production trajectory.

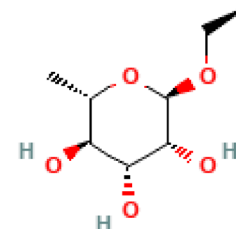
Table 9 Structure of Rutin, Ursolic Acid and Laurifolin.

Sl. No	Compound name	Structure
1	Rutin	



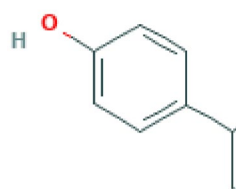
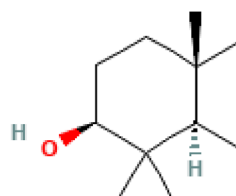
2

Ursolic Acid



3

Laurifolin



Over the final 50 ns of the trajectory the backbone RMSD varied by only 0.10 Å (standard deviation), and its mean across the equilibrated portion (2.00 ± 0.15 Å) was indistinguishable from that over the last 50 ns (1.99 ± 0.10 Å), indicating that the sampling window was sufficient to capture a stable binding mode. Statistical benchmarking of the RMSD trajectory therefore showed that the protein–ligand complex maintained a stable configuration throughout the production phase, and the consistency of the interaction occupancies described below supports the robustness of the VP40–rutin binding network. Additional convergence analyses are presented in Supplementary Figure S3: consecutive 50 ns block means of the backbone RMSD are stationary after the first block, a Flyvbjerg–Petersen blocking analysis together with the integrated autocorrelation time ($\tau_{\text{int}} = 21$ ns, giving ≈ 23 effectively independent samples) places the correlation-corrected standard error of the mean backbone

Table 10 Summary statistics for the 500 ns molecular dynamics trajectory of the VP40-rutin complex (501 frames).

Metric	Mean	SD	Min	Max	95% CI
VP40 backbone RMSD (Å)	1.98	0.19	0.00	2.56	1.97, 2.00
Rutin RMSD (Å)	2.28	0.25	0.00	3.11	2.26, 2.30
C α RMSF (Å)	1.10	0.52	0.47	3.80	1.04, 1.17
Radius of gyration (Å)	19.77	0.15	19.40	20.07	19.76, 19.78
SASA (Å ²)	14,003	228	13,433	14,442	13,958, 14,047
Hydrogen bonds	8.02	1.39	5	13	7.85, 8.20
Hydrophobic contacts	17.91	2.95	10	29	17.54, 18.27
Ionic contacts	3.36	0.77	1	6	3.26, 3.45
Water bridges	6.27	2.45	0	13	5.97, 6.58
π - π stacking	0.00	0.00	0	0	,
π -cation	0.00	0.00	0	0	,
vdW interaction energy (kcal/mol)	-57.35	4.22	-67.27	-43.51	-57.72, -56.98
Coulomb interaction energy (kcal/mol)	-60.24	12.98	-119.77	-33.73	-61.38, -59.10
Total MM interaction energy (kcal/mol)	-117.59	12.50	-168.73	-93.43	-118.69, -116.50

Table 11 MM/GBSA and MM/PBSA binding free-energy components for the VP40-rutin complex, from 51 snapshots sampled at 10 ns intervals over the 500 ns trajectory.

Component	MM/GBSA (mean $\pm$ SD)	MM/PBSA (mean $\pm$ SD)
ΔE_{vdw}	-49.58 $\pm$ 5.00	-49.58 $\pm$ 5.00
ΔE_{elec}	-53.15 $\pm$ 13.30	-53.15 $\pm$ 13.30
$\Delta G_{solv, polar}$	+ 64.89 $\pm$ 8.65	+ 80.60 $\pm$ 9.55
$\Delta G_{solv, nonpolar}$	-8.25 $\pm$ 0.29	-5.79 $\pm$ 0.14
ΔG_{bind} (total)	-46.08 $\pm$ 5.96	-27.92 $\pm$ 7.63
Standard error of the mean	0.84	1.07
95% confidence interval	-47.72 to -44.45	-30.01 to -25.82
Range (min, max)	-58.84, -35.47	-46.02, -14.02
Snapshots analysed (n)	51	51

Values in kcal/mol.

RMSD at only 0.03 Å, the RMSD distributions of the two trajectory halves overlap by 76%, and the cumulative mean of the VP40-rutin interaction energy converges well before the end of the run.

RMSD minima of 0.00 Å correspond to the $t=0$ reference frame. The MM interaction energy is the direct Coulomb + Lennard–Jones energy between VP40 and rutin and is not a binding free energy (see Table 11).

Residue-level flexibility (RMSF)

The root mean square fluctuation (RMSF) analysis was carried out to examine the flexibility of the protein at the residue level during the molecular dynamics simulation (Fig. 5). The RMSF values were computed using the C α atoms of the residues to determine the flexibility of the protein backbone.

In general, the RMSF plot showed that most residues had low fluctuations, with a mean C α RMSF of 1.10 ± 0.52 Å and 82.7% of residues fluctuating by less than 1.5 Å, indicating that the protein is structurally stable throughout the simulation time. The regions that correspond to secondary structural elements like α -helices and β -strands had relatively lower RMSF values, which is an indication of their rigid and ordered nature.

The regions with higher RMSF values were mostly found in the loops and terminal parts of the protein, which is in agreement with their flexible nature. Only twelve residues exceeded the mean + 2 SD threshold of 2.15 Å; these clustered in the 45–48, 72–73, 116–117, and 268–269 loops and at the C-terminal tail. The largest fluctuations were recorded for VAL303 (3.80 Å), ASP46 (3.11 Å), ALA302 (2.92 Å), GLU268 (2.88 Å), and LEU47 (2.86 Å).

Most importantly, the residues that come in contact with the ligand have shown markedly lower RMSF values compared with the highly flexible loop regions. The key contact residues PRO85 (0.62 Å), LEU86 (0.59 Å), ILE88 (0.47 Å), ARG136 (0.64 Å), GLN143 (1.00 Å), ALA144 (0.96 Å), ILE146 (1.11 Å), TRP179 (0.54 Å),

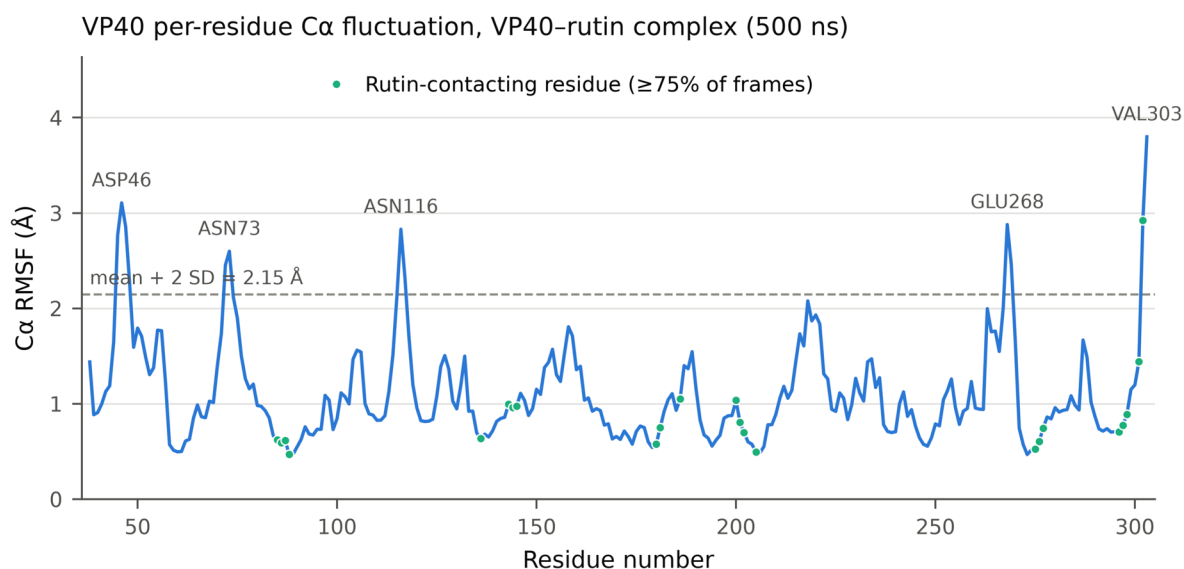


Fig. 5 Per-residue C α root mean square fluctuation (RMSF) of VP40 over the 500 ns simulation of the VP40-rutin complex. The dashed line marks the mean + 2 SD threshold (2.15 Å). Filled markers indicate residues lying within 4 Å of rutin in at least 75% of frames. The five most mobile residues are labelled.

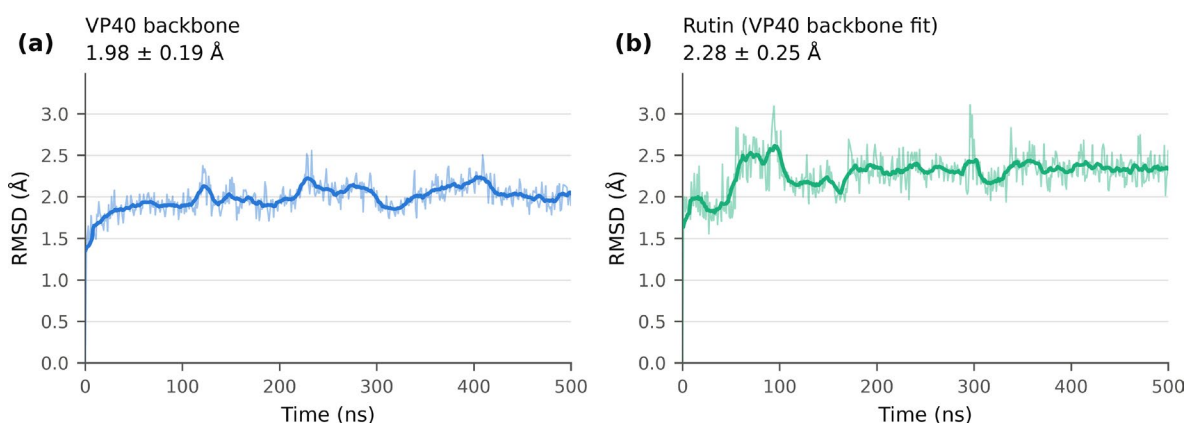


Fig. 6 Time-dependent root mean square deviation (RMSD) during the 500 ns simulation of the VP40-rutin complex. **(a)** VP40 backbone RMSD relative to the initial frame. **(b)** Rutin RMSD after least-squares superposition of the VP40 backbone. Pale traces show frame-wise values; heavy lines show a 15-frame running mean.

ARG180 (0.58 Å), ASN200 (1.04 Å), PRO202 (0.70 Å), VAL205 (0.50 Å), TYR274 (0.51 Å), GLN276 (0.60 Å), and ALA296 (0.71 Å) all fluctuated by less than 1.2 Å. ILE88, one of the principal hydrogen-bonding partners, recorded the lowest fluctuation of any residue in the protein.

These findings suggest that ligand binding contributes to stabilization of the binding pocket. This constrained movement of these residues shows a strong interaction between the protein and the ligand.

Global structural stability (RMSD)

The stability of the protein–ligand complex structure during the 500 ns simulation was analyzed by calculating the root mean square deviation (RMSD) of the protein and ligand atom positions (Fig. 6).

The C α RMSD of the protein increased during the first 15–20 ns of the simulation and then levelled off after the equilibration phase. Across the full trajectory the backbone RMSD averaged 1.98 ± 0.19 Å and never exceeded 2.56 Å. These results indicate that the protein structure was well maintained and preserved during the simulation.

Restricting the analysis to the equilibrated portion of the trajectory ($t \geq 20$ ns) gave a backbone RMSD of 2.00 ± 0.15 Å, and over the final 50 ns the value was 1.99 ± 0.10 Å. The absence of any systematic upward drift across half a microsecond of sampling indicates that no major conformational change of the protein occurred during the simulation.

Ligand RMSD calculated with respect to the protein backbone indicated an initial adaptation phase followed by stabilization. Over the first 80–100 ns the ligand RMSD varied moderately, reflecting refinement of the ligand conformation within the binding pocket. Thereafter it stabilized at 2.31 ± 0.17 Å, and at 2.34 ± 0.16 Å over the final 50 ns, averaging 2.28 ± 0.25 Å across the whole trajectory with a maximum excursion of 3.11 Å. Rutin therefore remained bound within the pocket for the entire simulation without any dissociation.

RMSD analysis clearly indicates that the protein and ligand have reached a stable dynamic equilibrium during the MD simulation.

Binding free energy and structural compactness analysis

To provide a more rigorous quantitative estimation of the binding affinity and structural integrity of the VP40-rutin complex, end-state binding free energy (MM/GBSA and MM/PBSA), radius of gyration (RoG), and solvent-accessible surface area (SASA) analyses were performed, which are illustrated in Figs. 7 and 8.

Thermodynamic profiling via MM/GBSA and MM/PBSA

The binding free energy ΔG_{bind} of the rutin-VP40 complex was calculated over 51 snapshots sampled uniformly at 10 ns intervals across the 500 ns production trajectory (Table 11, Fig. 7). The total binding free energy was strongly favorable under both solvation models: -46.08 ± 5.96 kcal/mol (SEM 0.84; 95% CI -47.72 to -44.45 kcal/mol) for MM/GBSA, and -27.92 ± 7.63 kcal/mol (SEM 1.07; 95% CI -30.01 to -25.82 kcal/mol) for MM/PBSA. The binding free energy was favourable in every one of the 51 snapshots analysed, ranging from -58.84 to -35.47 kcal/mol under the Generalized Born model and from -46.02 to -14.02 kcal/mol under the Poisson-Boltzmann model. Decomposition of the energy components revealed that van der Waals interactions

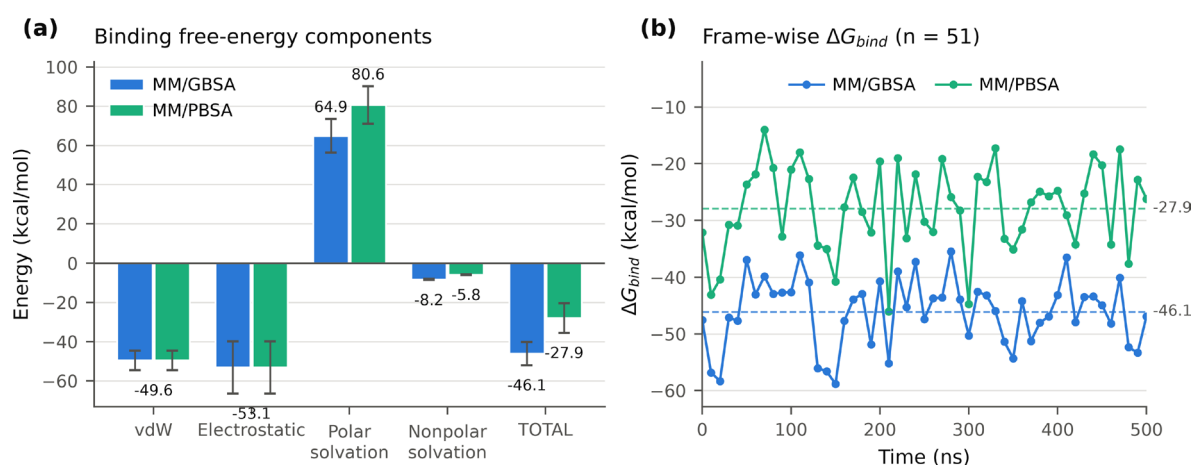


Fig. 7 Binding free energy of the VP40-rutin complex, from 51 snapshots sampled at 10 ns intervals over the 500 ns trajectory. **(a)** Mean energy components for MM/GBSA and MM/PBSA; error bars denote one standard deviation. The van der Waals and electrostatic terms are identical in the two calculations because they derive from the same gas-phase molecular-mechanics energies. **(b)** Frame-wise total binding free energy; dashed lines mark the ensemble means.

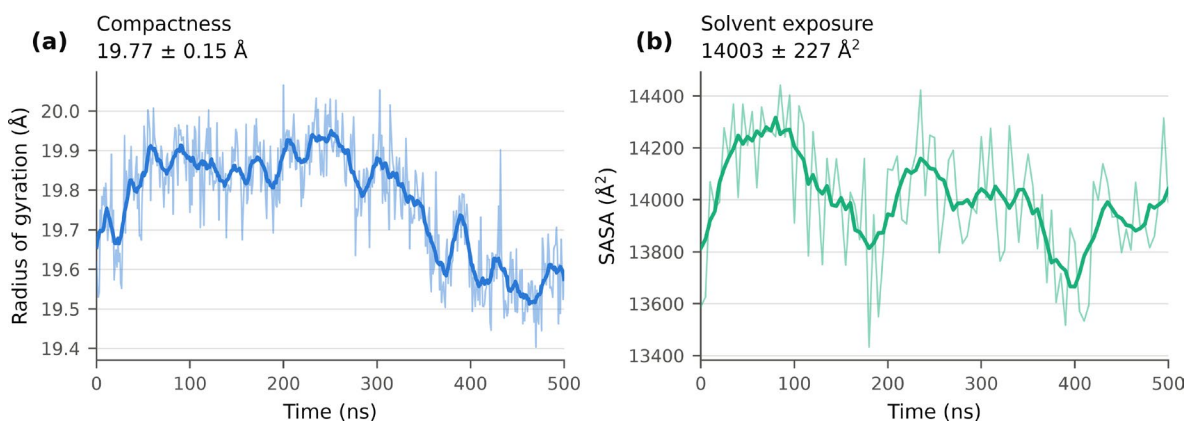


Fig. 8 Structural compactness and solvent exposure of the VP40-rutin complex over 500 ns. **(a)** Radius of gyration of the VP40 backbone. **(b)** Solvent-accessible surface area computed with the Shrake-Rupley algorithm. Pale traces show frame-wise values; heavy lines show a running mean.

($\Delta E_{vdw} = -49.58 \pm 5.00$ kcal/mol) and electrostatic contributions ($\Delta E_{elec} = -53.15 \pm 13.30$ kcal/mol) were the primary driving forces for complex stabilization, together yielding a gas-phase interaction energy of approximately -102.7 kcal/mol. The non-polar solvation term further supported the stability (-8.25 ± 0.29 kcal/mol for GB and -5.79 ± 0.14 kcal/mol for PB), partially offsetting the polar solvation penalty ($+64.89 \pm 8.65$ kcal/mol for GB and $+80.60 \pm 9.55$ kcal/mol for PB). Because the gas-phase molecular-mechanics terms are identical in the two calculations, the difference between the MM/GBSA and MM/PBSA totals arises entirely from the treatment of solvation; the more stringent Poisson-Boltzmann treatment imposes a larger desolvation penalty and therefore yields the more conservative estimate. These results align with the molecular docking and interaction-occupancy data, confirming that the hydrogen bonds and hydrophobic contacts involving residues such as ARG180, ILE88, ARG136, and GLN276 contribute to a robust inhibitory interface. Because the reported values exclude the conformational entropy term, they should be interpreted as effective binding free energies suitable for comparative ranking rather than as absolute affinities.

The van der Waals and electrostatic terms are identical in the two calculations because both are gas-phase molecular-mechanics energies evaluated on the same snapshots; the models differ only in the treatment of solvation. Conformational entropy ($-T\Delta S$) was not computed, so the totals are effective binding free energies intended for comparative ranking.

Assessment of structural compactness and solvent exposure (RoG and SASA)

Figure 8 represents the global stability, compactness, and solvent exposure of the protein backbone throughout the 500 ns simulation, monitored using the radius of gyration (RoG) and the solvent-accessible surface area (SASA). The RoG values for the VP40-rutin complex exhibited an initial stabilization phase within the first 20 ns and thereafter remained confined to a narrow band, averaging 19.77 ± 0.15 Å over a total range of only 19.40 to 20.07 Å. In parallel, the SASA held steady at $14,003 \pm 228$ Å² (range 13,433 to 14,442 Å²). Superimposed on these narrow distributions, both descriptors displayed a weak but systematic downward trend across the trajectory: linear regression gave -0.06 Å per 100 ns for the RoG and -53 Å² per 100 ns for the SASA, corresponding to a 1.2% contraction of the protein and a 1.8% reduction in exposed surface area over the full 500 ns. Because the backbone RMSD remained on a flat plateau near 2.0 Å and rutin stayed bound throughout, this trend reflects a gradual compaction and settling of solvent-exposed loops rather than a conformational transition or a loss of structural integrity. The VP40 protein therefore maintained its native, compact globular fold upon ligand binding, with no substantial opening of the hydrophobic core, indicating that rutin remains favorably accommodated

within the active site without inducing detrimental long-range conformational transitions or unfolding of the protein matrix.

Interaction, persistence, and binding behavior

The molecular interaction profile between the ligand and the target protein was characterized by counting the number and type of non-covalent contacts in every trajectory frame (Fig. 9). Across the 500 ns trajectory the complex maintained 8.02 ± 1.39 direct VP40-rutin hydrogen bonds, reaching a maximum of 13 in a single frame, embedded within an extensive non-polar envelope of 17.91 ± 2.95 simultaneous hydrophobic contacts. Water-mediated bridges contributed a smaller but steady 6.27 ± 2.45 contacts per frame, and charge-mediated (ionic) contacts averaged 3.36 ± 0.77 . Notably, neither π - π stacking nor π -cation interactions were detected in any frame of the trajectory. The π -cation contact with ARG180 predicted from the static docking pose is therefore not maintained under dynamic, solvated conditions; instead, ARG180 stabilises the ligand through persistent hydrogen bonding, as described below. This combination of a durable polar anchor and a large hydrophobic interface is characteristic of a specific and long-lived association.

At the residue level (Fig. 10), the ligand was stabilized by hydrogen bonds to four principal partners. ARG180 was the dominant polar anchor, contributing on average 1.21 hydrogen bonds per frame, more than one bond at a time, since its guanidinium group frequently donated to two hydroxyl acceptors of rutin simultaneously. ILE88 (0.97 bonds per frame), ARG136 (0.91), and GLN276 (0.90) formed the remaining high-occupancy hydrogen bonds, while ALA144 (0.30) and GLN143 (0.18) contributed at moderate frequency and VAL303, ARG139, and SER301 made only sporadic contributions (≤ 0.05 bonds per frame). Twelve residues, namely GLY87, PRO85, ILE88, PHE145, GLN143, ARG180, LEU186, PRO202, PHE275, GLN276, ALA296, and PRO298, lay within 4 Å of rutin in every frame of the simulation, and 23 residues did so in at least 75% of frames, defining a large and continuously occupied binding interface.

The residue-wise interaction profile, as depicted in Fig. 10, indicates the frequency with which each amino acid residue contacted the ligand over the course of the molecular dynamics simulation. The pattern of interaction is diverse, with a subset of residues in continuous contact and others engaging intermittently. Twelve residues remained within 4 Å of rutin in 100% of frames, and ALA144 (99.6%), ARG136 (99.2%), LEU201 (98.0%), and SER301 (96.0%) were only marginally less persistent.

Residues displaying high but sub-saturating contact frequencies included PRO181 (94.8%), ASN200 (88.5%), VAL205 (85.3%), ALA302 (78.5%), ASN297 (76.5%), LEU86 (76.5%), ALA277 (75.3%), and ILE146 (72.9%). Intermediate frequencies were recorded for VAL303 (59.0%), PHE120 (57.4%), ILE204 (40.6%), PRO203

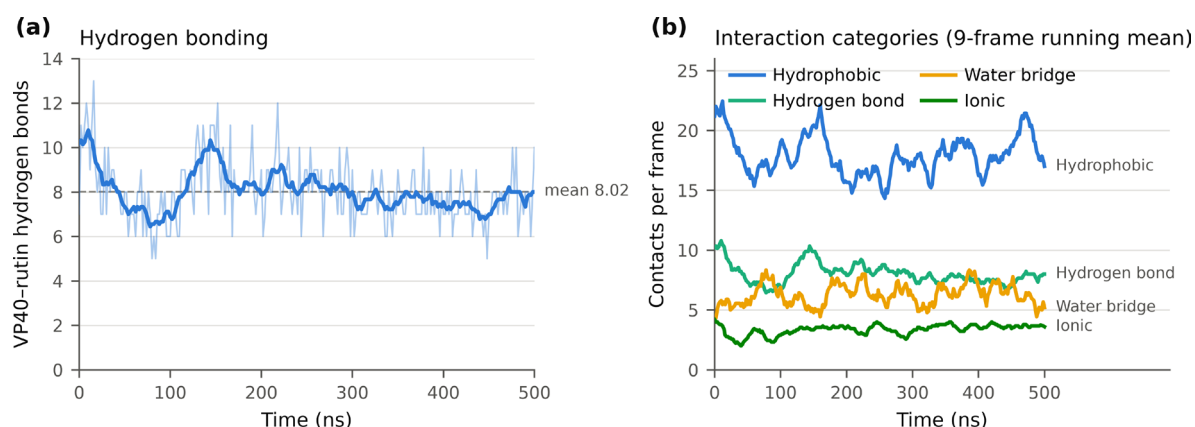


Fig. 9 Non-covalent contact profile of the VP40-rutin complex during the 500 ns simulation. **(a)** Number of direct VP40-rutin hydrogen bonds per frame; the dashed line marks the trajectory mean. **(b)** Nine-frame running mean of each interaction category. Neither π - π nor π -cation contacts were detected at any point in the trajectory.

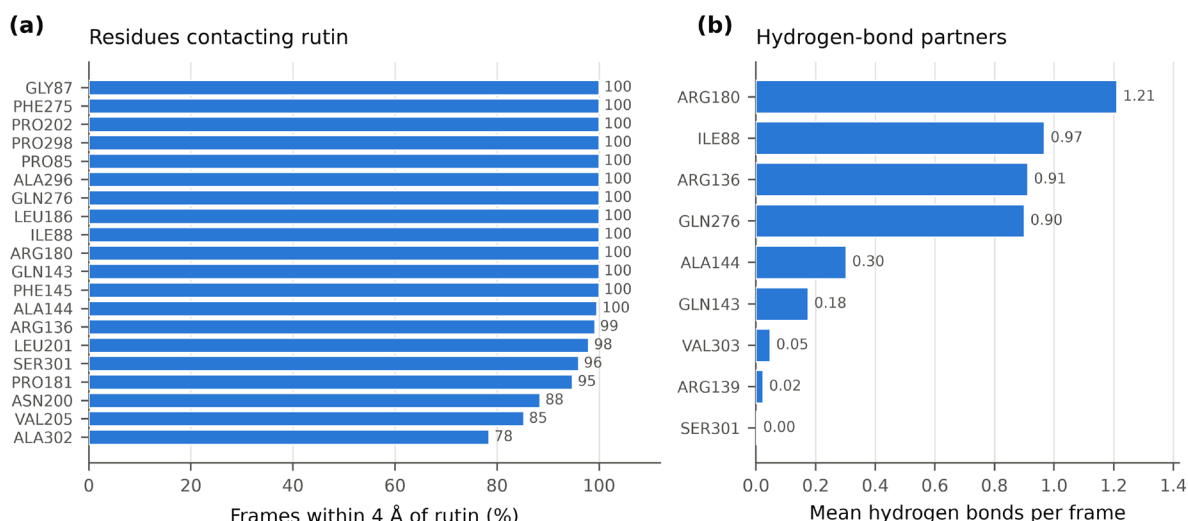


Fig. 10 Residue-wise interaction profile of the VP40-rutin complex during the 500 ns simulation. **(a)** The twenty residues most frequently found within 4 Å of rutin. **(b)** Mean number of hydrogen bonds per frame for each hydrogen-bonding partner; values above 1.0 indicate that a residue frequently forms more than one hydrogen bond to the ligand at the same time.

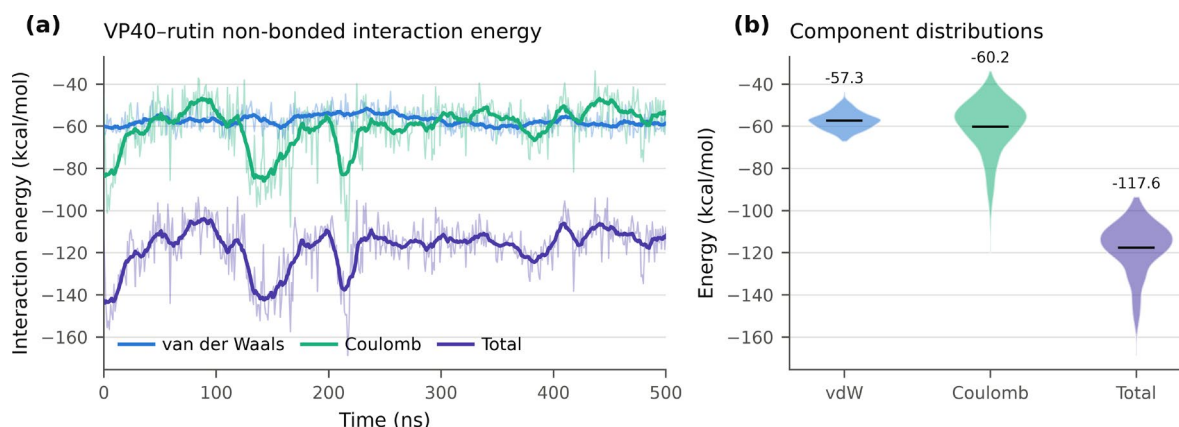


Fig. 11 VP40-rutin non-bonded molecular-mechanics interaction energy over 500 ns. **(a)** Van der Waals, Coulomb, and total interaction energy; pale traces show frame-wise values and heavy lines a 15-frame running mean. **(b)** Distribution of each component; horizontal bars mark the means. These are interaction energies, not binding free energies.

(26.3%), and TYR295 (21.5%). Residues in only occasional contact included ARG139 (12.8%), LEU135 (11.6%), MET149 (10.4%), PRO134 (7.6%), PRO278 (7.6%), LEU300 (5.6%), LEU84 (2.8%), MET89 (2.8%), TYR274 (2.4%), THR299 (1.2%), and MET137 (1.2%).

Non-bonded interaction energy

The direct, non-bonded interaction energy between VP40 and rutin was decomposed into van der Waals and Coulomb contributions over the full trajectory (Fig. 11). The van der Waals term averaged -57.35 ± 4.22 kcal/mol and was both strongly favourable and remarkably steady, reflecting close shape complementarity within the pocket. The Coulomb term averaged -60.24 ± 12.98 kcal/mol and was considerably more variable, consistent with the intermittent making and breaking of individual hydrogen bonds superimposed on the persistent ARG180 anchor. Together these gave a total molecular-mechanics interaction energy of -117.59 ± 12.50 kcal/mol, which remained favourable in every frame. It should be emphasized that this quantity is the direct interaction energy between the

two binding partners and not a binding free energy: it contains neither the desolvation penalty nor the entropic cost of association, which is why its magnitude greatly exceeds the MM/GBSA and MM/PBSA values reported in Table 11. It is presented here as a measure of the strength and temporal stability of the direct intermolecular contacts.

Essential dynamics (PCA)

Principal component analysis of the C α coordinates was used to describe the collective motions sampled during the simulation (Fig. 12). The first two principal components accounted for 21.2% and 15.1% of the total positional variance respectively, and the leading ten components together captured 64.2%, indicating that the essential dynamics of the complex are confined to a small number of collective modes. Projection of the trajectory onto the first two components produced a compact, bounded distribution rather than dispersal across widely separated basins, confirming that VP40 did not undergo a discrete conformational switch. The projection was, however, ordered in time: early frames occupied the positive-PC1 region and late frames the negative-PC1 region, describing a gradual, largely monotonic relaxation of the complex along the dominant mode rather than reversible hopping between distinct states. This slow drift within a bounded subspace mirrors the gradual contraction of the radius of gyration and solvent-accessible surface area described above and, given the flat backbone RMSD, reflects progressive settling of surface loops rather than any change in the protein fold or in the binding mode of rutin. Consistent with this interpretation, the cosine content of the two leading components is high (0.79 for PC1 and 0.57 for PC2, versus 0.01 for PC3; Supplementary Figure S3f.). A cosine content approaching unity is the signature of a diffusion-like, random-walk mode rather than of a well-sampled transition between free-energy minima. The slowest collective motions of VP40 are therefore not exhaustively sampled on the 500 ns timescale, and PC1 and PC2 should be read as descriptors of a gradual relaxation rather than as evidence for a functionally distinct conformational state. Importantly, this limitation concerns the essential subspace only: the scalar structural and energetic observables that underpin the conclusions of this study, backbone RMSD, radius of gyration, contact occupancy, and binding free energy, are all stationary over the production trajectory (Supplementary Figure S3a-e).

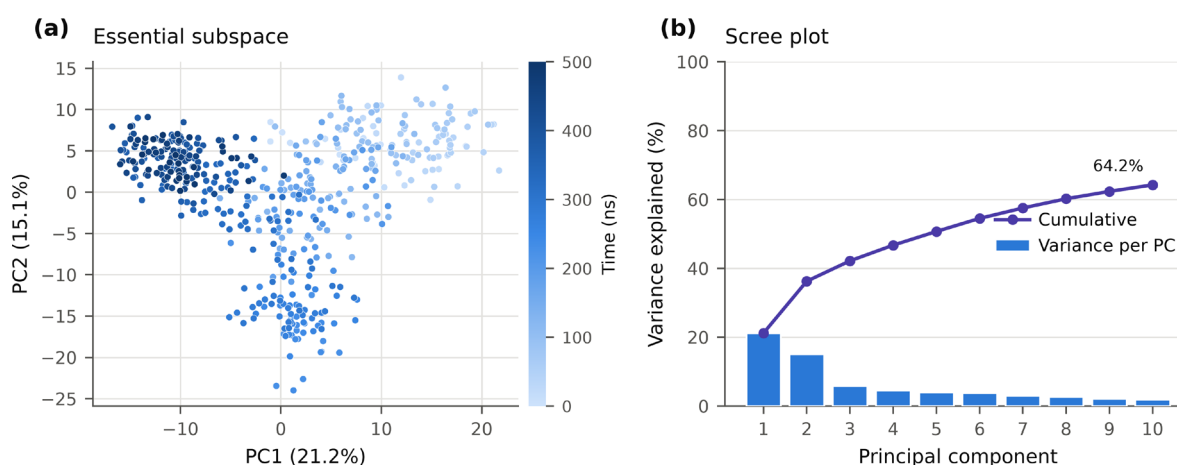


Fig. 12 Essential dynamics of the VP40-rutin complex. **(a)** Projection of the trajectory onto the first two principal components of the C α coordinates, coloured by simulation time. **(b)** Scree plot of the variance explained by the first ten principal components together with the cumulative total.

Correlated motions (DCCM)

The dynamic cross-correlation matrix of C α displacements (Fig. 13) revealed the expected band of strong positive correlation along the diagonal, corresponding to concerted motion within individual secondary-structure elements, together with well-defined off-diagonal blocks of both positive and negative correlation. Positively correlated off-diagonal blocks link the loop regions flanking the binding pocket, whereas anti-correlated blocks identify segments that move in opposition, a signature of coordinated, hinge-like breathing between the N- and C-terminal regions of the protein. The binding-pocket residues that anchor rutin lie within the positively correlated core rather than in the anti-correlated hinge regions, consistent with the low RMSF values recorded for those residues and with the persistence of the interaction network described above.

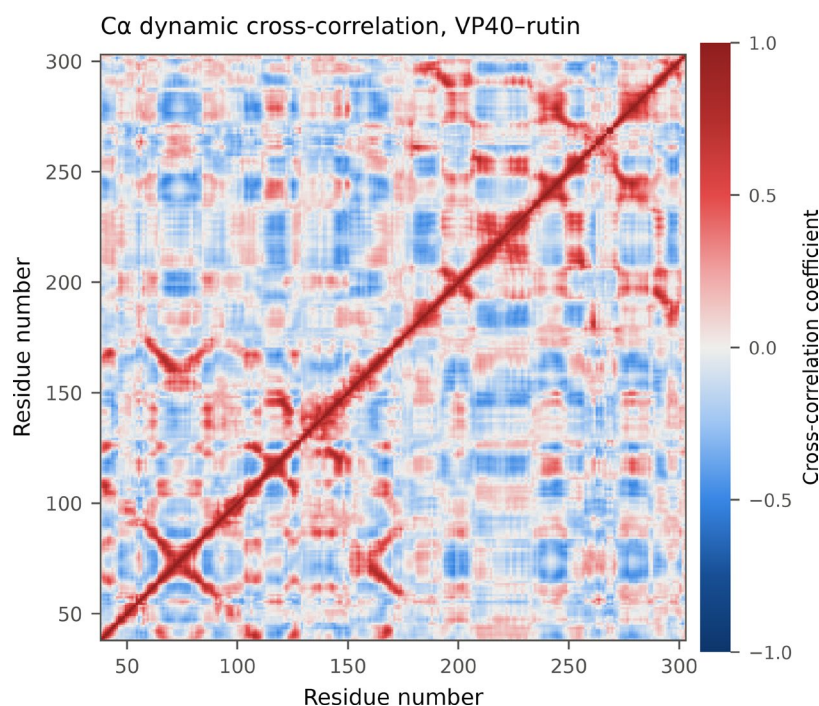
Per-residue energy decomposition

To identify which residues drive the favourable binding free energy, the MM/GBSA energy was decomposed per residue over 26 snapshots spaced 20 ns apart, with the printed output restricted to residues lying within 5 Å of rutin (Fig. 14, Table 12). The 34 residues identified in this way, together with the ligand's own contribution, sum to -46.17 kcal/mol. This reproduces the total MM/GBSA binding free energy of -46.08 ± 5.96 kcal/mol obtained independently from the 51-snapshot calculation to within 0.09 kcal/mol, confirming the internal consistency of the decomposition.

The interface is dominated by van der Waals contacts rather than by hydrogen bonding. The largest individual contributions come from PRO202 (-2.96 ± 1.01 kcal/mol), GLN276 (-2.15 ± 1.38), GLY87 (-1.95 ± 0.71), ILE88 (-1.91 ± 1.15), PRO298 (-1.68 ± 1.14), PHE145 (-1.55 ± 1.23) and ALA296 (-1.53 ± 1.00), every one of which lies within 4 Å of the ligand in 100% of trajectory frames. For PRO202, PRO298 and PHE145 the stabilisation is almost entirely dispersive, their electrostatic terms being smaller than 0.5 kcal/mol in magnitude.

ARG180, the residue with the highest hydrogen-bond occupancy at 1.21 bonds per frame, makes a net destabilizing contribution of +1.18 kcal/mol. Its strongly favourable electrostatic interaction with rutin (-5.21 kcal/mol) is more than cancelled by the polar desolvation penalty incurred on burying the charged guanidinium group

Fig. 13 Residue-residue dynamic cross-correlation matrix of C α displacements for the VP40-rutin complex over 500 ns. Red indicates positively correlated motion, blue anti-correlated motion, and white uncorrelated motion.



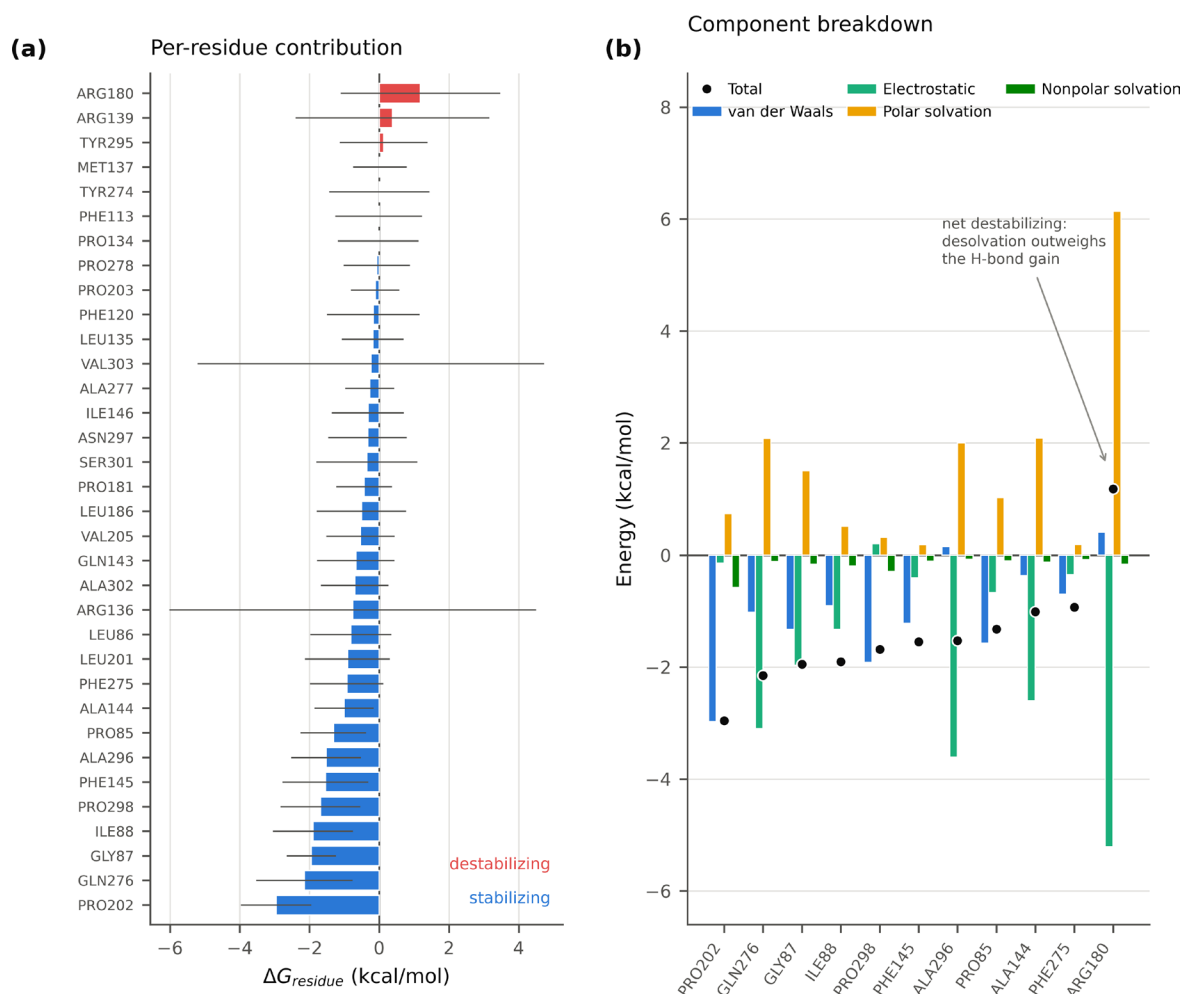


Fig. 14 Per-residue MM/GBSA energy decomposition of the VP40-rutin complex. **(a)** Net contribution of each of the 34 residues lying within 5 Å of rutin; blue denotes stabilizing and red destabilizing contributions, and error bars show the standard error of the mean. **(b)** Component breakdown for the ten most stabilizing residues together with ARG180. Although ARG180 forms the most persistent hydrogen bonds of any residue, it is net destabilizing because its polar desolvation penalty exceeds its favourable electrostatic interaction.

(+6.14 kcal/mol). ARG139 behaves the same way (+0.38 kcal/mol), and ARG136 is only marginally stabilizing (-0.76 kcal/mol) with a large frame-to-frame variance. This reconciles two observations that would otherwise appear contradictory: arginine residues form the most persistent hydrogen bonds in the simulation, yet the total electrostatic term (-53.15 kcal/mol) is almost entirely offset by the polar solvation penalty (+64.89 kcal/mol) in the global energy budget. The persistent arginine hydrogen bonds therefore appear to orient and anchor rutin within the pocket, whereas the net thermodynamic driving force for binding is hydrophobic and dispersive.

Per-residue decomposition partitions an effective free energy and is model-dependent; the values are informative for comparison with one another rather than as absolute residue affinities, and the standard deviations reported in Table 12 are large relative to the means for the charged residues.

Mechanistic interpretation

The combined RMSD, RMSF, contact-occupancy, interaction-energy, and binding free-energy data suggest that rutin remains favorably accommodated within the VP40 binding region throughout half a microsecond of

Table 12 Per-residue MM/GBSA decomposition of the VP40-rutin binding free energy over 26 snapshots (residues within 5 Å of the ligand).

The ten most stabilizing residues are listed, together with the arginines and the ligand's own contribution. Values in kcal/mol. Sum over all 34 printed residues plus the ligand = -46.17 kcal/mol, against -46.08 kcal/mol for the independent 51-snapshot total. Positive ΔG_{total} denotes a net destabilizing contribution.

Residue	ΔE_{vdw}	ΔE_{elec}	$\Delta G_{\text{solv, polar}}$	$\Delta G_{\text{solv, nonpolar}}$	$\Delta G_{\text{total}} \pm \text{SD}$
PRO202	-2.97	-0.14	+0.74	-0.58	-2.96 $\pm$ 5.13
GLN276	-1.02	-3.10	+2.08	-0.11	-2.15 $\pm$ 7.04
GLY87	-1.32	-1.97	+1.50	-0.16	-1.95 $\pm$ 3.60
ILE88	-0.90	-1.32	+0.52	-0.19	-1.91 $\pm$ 5.84
PRO298	-1.91	0.20	+0.32	-0.29	-1.68 $\pm$ 5.83
PHE145	-1.22	-0.40	+0.18	-0.11	-1.55 $\pm$ 6.26
ALA296	0.15	-3.61	+2.00	-0.07	-1.53 $\pm$ 5.10
PRO85	-1.57	-0.67	+1.02	-0.10	-1.32 $\pm$ 4.78
ALA144	-0.37	-2.60	+2.08	-0.13	-1.01 $\pm$ 4.31
PHE275	-0.70	-0.35	+0.19	-0.08	-0.93 $\pm$ 5.36
ARG136	-0.75	-4.50	+4.78	-0.29	-0.76 $\pm$ 26.82
ARG139	-0.21	-1.87	+2.46	-0.01	0.38 $\pm$ 14.16
ARG180	0.41	-5.21	+6.14	-0.16	1.18 $\pm$ 11.66
Rutin (ligand)	-25.06	-25.79	+32.54	-5.08	-23.39 $\pm$ 17.20

simulation. Local rigidification of binding-site residues together with persistent intermolecular contacts may interfere with conformational transitions required for VP40 functional activity, including membrane association or oligomerisation-related behaviour.

Although experimental validation is still required, the MD findings support the docking results and strengthen the hypothesis that rutin may serve as a promising lead scaffold against Marburg virus VP40.

DFT analysis

A value of -0.2217 Hartree was obtained for the HOMO energy level of Rutin. This corresponds to a vertical ionisation potential of 6.03 eV and indicates a capacity to donate electron density to electron-deficient residues of the pocket, which may reinforce hydrogen bonding. The energy level of the LUMO indicates the ability of the molecule to accept an electron, and the value -0.0701 Hartree suggests that, aside from being a donor, Rutin is also likely to accept an electron from the amino acid residues that are nucleophilic.

The HOMO–LUMO gap (ΔE) of 0.1516 Hartree, equivalent to 4.13 eV, is consistent with appreciable chemical reactivity and polarisability. A lower HOMO–LUMO energy gap may facilitate electronic adaptability and favorable intermolecular interactions within the protein-binding environment.

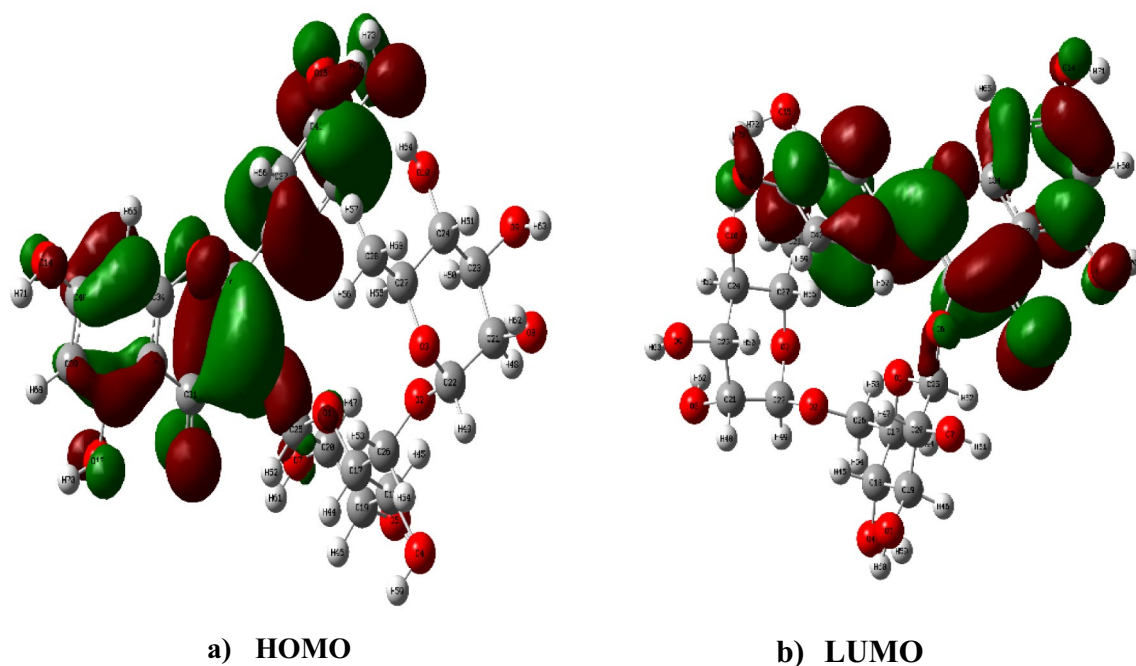
The chemical hardness of 0.0758 Hartree, or 2.06 eV, is modest, and the corresponding softness of 13.199 reciprocal Hartree, or 0.485 reciprocal electron volts, indicates that the electron distribution of rutin is readily deformed on approach to the protein. The additional descriptors derived from the same frontier orbital energies place this behaviour on a more quantitative footing. Applying Koopmans' theorem gives a vertical ionisation potential of 6.03 eV and a vertical electron affinity of 1.91 eV, from which the electronegativity is 3.97 eV and the chemical potential is -3.97 eV. The electrophilicity index is 3.82 eV, which falls in the moderate range on the Domingo scale and marks rutin as neither a strong electrophile nor a strong nucleophile but a molecule able to act in either capacity depending on the partner. The maximum charge transfer, ΔN_{max} , is 1.92 electrons, indicating substantial capacity to accept electron density from an electron-rich environment such as the carboxylate and carbonyl groups lining the predicted pocket. All values are collected in Table 13 and the frontier orbitals are shown in Fig. 15.

The DFT descriptors provide supporting electronic context for the docking and simulation results. They do not explain the binding affinity, which is a free energy and is addressed by the MM/GBSA and MM/PBSA calculations, but they are consistent with it. The high softness and the moderate gap indicate a molecule whose electron distribution can adjust to the electrostatic field of the pocket, which is compatible with the low ligand RMSD and the reduced fluctuation of the binding-site residues recorded in the simulation. The electronic adaptability implied by these descriptors is likewise consistent with rutin being accommodated in the pocket without destabilising the

Table 13 Frontier orbital energies and global reactivity descriptors of rutin, calculated at the B3LYP/6-31G(d,p) level in the gas phase.

Descriptor	Symbol	Hartree	eV
Highest occupied molecular orbital	ϵ_{HOMO}	-0.2217	-6.03
Lowest unoccupied molecular orbital	ϵ_{LUMO}	-0.0701	-1.91
Energy gap	ΔE	0.1516	4.13
Ionisation potential	IP	0.2217	6.03
Electron affinity	EA	0.0701	1.91
Electronegativity	χ	0.1459	3.97
Chemical potential	μ	-0.1459	-3.97
Chemical hardness	η	0.0758	2.06
Softness	$S = 1/\eta$	13.199	0.485
Electrophilicity index	$\omega = \mu^2/2\eta$	0.1404	3.82
Maximum charge transfer	$\Delta N_{\text{max}} = -\mu/\eta$	1.92	n/a (dimensionless)

Values are given in Hartree and in electron volts (1 Hartree = 27.2114 eV). Softness is reported as the reciprocal of the hardness, so its units are the reciprocal of the energy unit of the column. The maximum charge transfer is dimensionless and is expressed in electrons.

**Fig. 15** Frontier molecular orbital analysis of Rutin: (a) highest occupied molecular orbital (HOMO) and (b) lowest unoccupied molecular orbital (LUMO) obtained from DFT calculations.

protein fold. These descriptors were calculated in the gas phase and are interpreted comparatively rather than as predictors of affinity.

Comparison with previously reported filoviral antiviral compounds

A dedicated comparative analysis of representative computational studies targeting filoviral VP40 proteins is presented in Table 14, which places the present work side by side with earlier reports in terms of the target protein, the lead compound, the best docking score, and whether molecular dynamics, end-state free-energy and ADMET analyses were performed. Although the reported docking scores are not directly comparable, because different studies use different protein structures, docking protocols and scoring functions, the comparison shows that the

Table 14 Comparative analysis of the present study with previously reported computational studies targeting filoviral VP40 inhibitors.

Study	Virus	Target	Lead compound(s)	Best docking score (kcal/mol)	MD simulation	MM/GBSA / MM/PBSA analysis	ADMET evaluation	Reference
Present study	MARV	VP40	Rutin	-10.4	✓	✓	✓	
Das, R., et al.,	MARV	VP40	Procyanidin Estradiol Benzoate	Procyanidin = -11.3 Estradiol Benzoate = -8.9	Yes	Yes	Yes	45
Alanzi AR et al.,	MARV	VP40	Gliotoxin	-9.5	Yes	Yes	Yes	46
Broni, E, et al.,	EBOV	VP40	ZINC000034518176 NANPDB2933 ZINC000085531689	ZINC000034518176 = -8.1 NANPDB2933 = -8.5 ZINC000085531690 = -8.9	Yes	Yes	Yes	47
Urata, S., et al	EBOV	VP40	NUSU #1 NUSU #2	The compactness of the dimer interface made pocket selection improbable for molecular docking	Yes	No	No	48

present study couples docking to a 500 ns trajectory, to both MM/GBSA and MM/PBSA free energies with per-residue decomposition, and to ADMET and DFT analyses, a combination not present in the earlier filoviral VP40 studies surveyed. Direct comparison of docking scores across studies should nevertheless be interpreted with caution.

Discussion

The search for potent antiviral compounds effective against the Marburg virus (MARV), a deadly filovirus causing severe hemorrhagic fever, remains a major public health issue because of the lack of approved therapeutic agents for the treatment of the infection caused by this lethal virus^{49,50}. In the current study, a holistic computational approach has been used to evaluate the potential of the phytoconstituents of the plant *Ocimum basilicum*, especially rutin, as promising inhibitors of the matrix protein VP40 of the Marburg virus.

Marburg virus VP40 is a multifunctional matrix protein that plays crucial roles for viral assembly, membrane binding, oligomerization, and budding of the virus from the plasma membrane of the host cell^{51,52}. After the entry of the virus into the host cell through macropinocytosis and endosomal trafficking, the viral ribonucleoprotein is transported to the cytoplasm, where VP40 mediates the assembly of the virions by binding to the phospholipids of the host cell membrane and recruiting viral components to the budding sites⁵³. VP40 oligomerization and its interaction with the inner leaflet of the host cell membrane are critical for the egress of the virions from the host cell, and it has been demonstrated that the disruption of this process can lead to a substantial impairment of viral replication and release⁵⁴. Therefore, VP40 is a very attractive target for antiviral therapy, as its inhibition can effectively block the late stage of the viral life cycle without directly targeting the replication of the viral genome.

The homology model of Marburg virus VP40 used in this study met the accepted thresholds for stereochemical quality and for agreement with its template, which makes it a sound basis for the docking and simulation work that follows. The binding pocket under investigation is rich in both polar and hydrophobic residues, thus capable of binding diverse ligands. Phytochemicals from *Ocimum basilicum* were chosen in consideration of the traditional medicinal use of the plant and its broad antiviral activities, in addition to its flavonoid content, particularly rutin, which has been recognized for its pleiotropic activities including antioxidant and antiinflammatory activities⁵⁵.

Among the screened compounds, rutin ranked first on docking score and was selected on that basis for the later stages of the workflow. The docking stage is treated here as a ranking device and nothing more. The residues that recur in the analysis, chiefly the arginine, isoleucine and glutamine residues of the predicted pocket, are identified from the molecular dynamics contact occupancies and the per-residue free-energy decomposition rather than

from the docked pose, and residues of this kind have been reported to contribute to oligomer stabilisation and membrane association in filoviral VP40 proteins⁵⁶.

Benchmarking against the reference ligand Laurifolin placed rutin ahead of it in the screen, with a more favourable score and a larger contact list in the docked pose. This is a comparison of screening output and carries no implication about relative potency, since the scoring function has been shown here not to rank pose geometry reliably. The evidence that rutin is stably accommodated in the pocket comes instead from the simulation, in which the ligand remained bound for the whole 500 ns and maintained an average of eight direct hydrogen bonds and eighteen hydrophobic contacts per frame. On that evidence, and on the favourable end-state free energies, rutin may have the potential to interfere with VP40-associated functions, a proposition that remains to be tested experimentally.

Several residues surrounding the predicted binding pocket are located in regions reported to contribute to VP40 intermolecular assembly and membrane-associated oligomerization in filoviral VP40 proteins. Therefore, ligand occupation of this site may potentially perturb conformational transitions or protein–protein interactions required for efficient oligomer formation. However, this hypothesis requires direct experimental validation⁵⁷. It should be noted that the proposed VP40 binding pocket was identified through computational pocket mapping and structural comparison with previously characterized functional regions of filoviral VP40 proteins. Although the predicted site is structurally plausible, no experimental validation through site-directed mutagenesis, biochemical binding assays, or structural techniques has yet confirmed its direct involvement in ligand recognition. Therefore, the proposed binding interface should be regarded as a computationally supported hypothesis requiring future experimental verification. Accordingly, the term active site in this manuscript refers to a computationally predicted ligand-binding pocket rather than an experimentally confirmed catalytic or druggable site; individual pocket residues should not be interpreted as validated functional determinants until mutagenesis, binding, or structural assays are performed. A binding-site specificity control provides additional computational support for the choice of pocket: rutin scored 4.4 kcal/mol more favourably in the CASTp-predicted site than in five decoy boxes placed elsewhere on the protein surface (Supplementary Figure S4d). This strengthens the case that the pocket is the preferred small-molecule binding region of the model, but it is computational evidence and does not constitute the experimental validation, by mutagenesis, biochemical assay or structural determination, that the caveat above requires.

Structural studies of Marburg virus VP40 and homologous Ebola virus VP40 have shown that oligomerization involves coordinated domain rearrangements and membrane-associated interfaces enriched in charged and hydrophobic residues. Several rutin-interacting residues identified in this study are positioned adjacent to these functionally important regions, providing a structural rationale for the observed binding pattern⁵⁸.

Docking served only to select rutin for further study. The dynamic stability and the solvated behaviour of the VP40-rutin complex were established by molecular dynamics simulation. The RMSD values of the protein backbone were found to be within acceptable limits throughout the simulation, thus ensuring that the VP40 protein retains its integrity even after binding with the ligand. At the same time, the stabilization of the ligand RMSD values after the initial equilibration phase ensured that the rutin molecule was firmly bound to the binding pocket without any dissociation⁵⁹.

Analysis of the persistence of interactions at the residue level indicated that rutin is involved in long-lived hydrogen bonds with critical residues such as ARG180, ILE88, ARG136 and GLN276, which serve as anchoring residues within the binding site; ARG180 alone contributed more than one hydrogen bond per frame on average across the 500 ns trajectory. This is a critical point because it is well known that the arginine-rich region of VP40 is involved in electrostatic interactions with host membranes and in oligomerization^{60,61}. Thus interaction with these residues could potentially interfere with the ability of VP40 to adopt a membrane-interacting conformation. Per-residue MM/GBSA decomposition, however, shows that the net thermodynamic driving force for binding is dispersive rather than electrostatic: the largest stabilizing contributions come from PRO202, GLN276, GLY87, ILE88 and PRO298, whereas ARG180 is net destabilizing (+1.18 kcal/mol) because the penalty for desolvating its guanidinium group outweighs the energetic gain from its hydrogen bonds. The arginine contacts therefore

appear to orient and anchor the ligand rather than to pay for binding, a distinction that would be missed by an interaction-occupancy analysis alone. Moderate and consistent interaction with residues such as ALA144 and GLN143, along with water-mediated hydrogen bonds averaging 6.3 per frame, indicates a cooperative binding mechanism, and hydrophobic interactions with LEU86, ALA296, PRO202 and PHE275 further anchor the ligand through a synergistic interaction network.

From a mechanistic point of view, the inhibitory potential of rutin on the Marburg virus can be explained by several complementary mechanisms. Firstly, by stabilizing VP40 in a non-productive conformation, rutin could inhibit the proper oligomerization necessary for the formation of viral particles⁶². Secondly, the inhibition of VP40 membrane interactions could inhibit viral budding, thereby effectively trapping the viral components inside the host cell⁶³. Thirdly, the high hydrogen-bonding potential of rutin could inhibit the interaction of VP40 with host cellular factors necessary for virion trafficking and release. The antiviral potential of rutin identified in this study can also be viewed within the broader landscape of filoviral drug discovery. Several computational investigations have reported natural and synthetic compounds targeting filoviral proteins, including VP40, VP35, nucleoprotein, and glycoprotein complexes. Reported docking scores for promising filoviral inhibitors typically range between approximately -7 and -11 kcal/mol, depending on the target and docking protocol employed. The binding affinity observed for rutin toward MARV VP40 therefore falls within the upper range reported for several candidate filoviral inhibitors. Furthermore, the persistence of key intermolecular interactions during molecular dynamics simulations and the favorable MM/GBSA profile support the potential of rutin as a competitive lead scaffold for future anti-filoviral drug development.

The novelty of the present study lies in the systematic integration of molecular docking, molecular dynamics simulations, ADMET profiling, and density functional theory (DFT) analysis to evaluate phytochemicals from *Ocimum basilicum* against the Marburg virus VP40 protein. Unlike previous reports that primarily focus on docking-based screening, this work provides a multi-level mechanistic insight by correlating electronic properties of rutin with its dynamic binding stability and residue-level interaction profile. Moreover, to the best of our knowledge, this is one of the first studies exploring *Ocimum basilicum*-derived compounds specifically against MARV VP40, a critical target involved in viral assembly and budding. This integrated computational framework strengthens the reliability of rutin as a lead scaffold and provides a more comprehensive understanding of its interaction mechanism.

Although the docking, molecular dynamics, ADMET and DFT workflow employed here is widely adopted in structure-based drug discovery, the novelty of the present work lies in its target-specific application and in the depth of the validation rather than in the workflow itself. We agree with the assessment that the pipeline as such is not novel. Four elements distinguish this study from the typical docking-plus-short-simulation report. First, the matrix protein VP40, a multifunctional regulator of viral assembly and budding, is targeted instead of the polymerases and proteases that dominate antiviral in silico work. Second, the binding mode is probed with a 500 ns trajectory, five times the duration commonly reported for filoviral VP40 studies, and its convergence is quantified explicitly through block averaging, correlation-corrected standard errors and the cosine content of the principal components, which reveals that the slowest collective modes remain unconverged even at this timescale. Third, binding free energies are computed under two independent solvation models and decomposed per residue, which shows the interface to be dispersion-driven and identifies the most persistent hydrogen-bonding residue, ARG180, as net destabilizing once desolvation is accounted for. Fourth, the docking protocol is itself validated against independent molecular dynamics reference poses, an analysis that reveals its pose-ranking limitations and that, to our knowledge, is rarely reported in this literature. The novelty is therefore best interpreted as target- and evidence-specific rather than methodological, and the principal contribution of the work is a critically validated characterisation of the VP40-rutin interaction rather than a new computational technique.

Although rutin exhibited the highest binding affinity and favorable dynamic stability toward MARV VP40, its predicted pharmacokinetic profile represents an important limitation for direct therapeutic development. Specifically, rutin possesses a high molecular weight, large topological polar surface area (TPSA), multiple violations of Lipinski's rule of five, and poor predicted intestinal absorption, all of which may adversely affect its oral

bioavailability. These findings indicate that, rutin is unlikely to serve as a drug candidate in its native form. Nevertheless, such pharmacokinetic limitations are common among glycosylated flavonoids and do not necessarily preclude their value as lead scaffolds in early-stage drug discovery. Structural optimization strategies, including deglycosylation, prodrug design, or derivatization, together with advanced formulation approaches such as nano-encapsulation or liposomal delivery, may substantially improve its pharmacokinetic properties while preserving its favorable interaction with VP40. Therefore, rutin should be regarded as a promising lead compound for further optimization rather than a ready-to-use therapeutic agent. In the present ADMET results, these concerns are reflected by a molecular weight of 610.52 g/mol, TPSA of 269.43 Å², 10 hydrogen-bond donors, 16 hydrogen-bond acceptors, three Lipinski violations, and predicted intestinal absorption of only 23.45%. These values indicate that rutin's strong *in silico* affinity must be weighed against a substantial bioavailability liability⁶⁴.

To further rationalize the high binding affinity and dynamical stability of rutin specifically with the Marburg virus VP40 protein, density functional theory (DFT) computations were carried out to study the electronic properties of the compound. Frontier molecular orbital (FMO) theory analysis resulted in the determination of the spatial features of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO), specifically in defining the reactivity and propensity of a compound^{65,66}. The calculated eigenvalues of the highest occupied molecular orbital and lowest unoccupied molecular orbital of rutin imply a moderate reactivity and reasonable stability in the compound, a property of a compound that could be favorable for any biological activity⁶⁷.

In addition, the relatively small HOMO–LUMO gap may contribute to favorable electronic adaptability and potential non-covalent interaction capability with protein functional groups through hydrogen bonding and electrostatic interactions⁶⁸. From the HOMO of rutin, it is clear the compound has the ability to act as an electron donor. From the LUMO, it is also clear that the compound can act as an electron acceptor, which may support stable interactions within the VP40 binding pocket.

The chemical hardness (η) and softness (S) are global reactivity descriptors obtained from DFT calculations that provide supportive insight into the interaction behavior of rutin with the VP40 binding pocket^{69,70}. However, it should be emphasized that HOMO–LUMO gap, hardness, and softness are theoretical electronic descriptors and do not directly correspond to biological activity or binding affinity in complex biological systems. In the present study, these parameters were used as complementary tools to rationalize molecular reactivity, charge-transfer potential, and the ability of rutin to adapt electronically within the protein environment. Lower hardness and higher softness values suggest greater polarizability and flexibility, which may facilitate favorable non-covalent interactions such as hydrogen bonding and electrostatic contacts with key residues including ARG180, ILE88, and GLN276⁷¹. These interpretations are consistent with the stable RMSD profiles, persistent hydrogen bond occupancy, and residue-level interactions observed during molecular dynamics simulations. Therefore, the DFT findings should be considered supportive mechanistic evidence that complements the structural and dynamic results obtained from docking and MD simulations, rather than standalone predictors of antiviral efficacy. Similar integrative computational approaches have previously been reported in several recent antiviral and phytochemical-based drug discovery studies; doing so lends strength to structure–activity relationships and enhances the reliability of *in silico* predictions⁷².

Although the present study is entirely computational, the findings provide a strong rationale for future experimental validation. The predicted binding affinity of rutin toward VP40 can be verified through *in vitro* assays such as recombinant VP40 protein binding studies, enzyme inhibition assays, or virus-like particle (VLP) budding assays. Furthermore, cell-based antiviral assays using MARV pseudotyped systems could be employed to evaluate the inhibitory effect on viral replication and egress. Biophysical techniques such as surface plasmon resonance (SPR) or isothermal titration calorimetry (ITC) may also be used to confirm direct ligand–protein interactions. These experimental approaches would help to validate the *in silico* predictions and strengthen the biological relevance of the proposed inhibitory mechanism.

Despite the advantages of this study, some limitations need to be considered. The results of this study are purely computational and therefore need to be verified experimentally to ensure antiviral activity. The predicted

interactions can be further validated through in vitro assays such as SPR (Surface Plasmon Resonance) and ITC (Isothermal Titration Calorimetry) for binding confirmation, as well as cell-based antiviral assays and VP40-specific virus-like particle (VLP) budding assays to evaluate functional inhibition. These approaches will provide experimental support to the in silico findings and improve the translational relevance of the study. The homology model, despite being rigorously validated, may not be able to represent all the conformational states of VP40 that exist in vivo. Moreover, the molecular dynamics simulations carried out in this study are limited by time and may not be able to capture long-range conformational changes or rare binding events. The study also targeted VP40 as a monomer, whereas the replication of Marburg virus requires the concerted action of multiple viral and host proteins. While our computational results support stable ligand occupancy near functionally relevant VP40 surfaces, dedicated protein–protein docking, oligomeric-state simulations, or biochemical oligomerization assays would be required to directly confirm disruption of VP40 multimerization. Additionally, although apo/holo comparative simulations were not included in the current study, such analysis would be valuable in future work to quantify ligand-dependent modulation of VP40 dynamics. Consistent with this, the cosine content of the two leading principal components (0.79 and 0.57; Supplementary Figure S3) indicates that the slowest collective motions of VP40 retain diffusion-like character even after 500 ns, so the essential subspace is not exhaustively sampled and the principal components should not be interpreted as resolved conformational substates; the scalar structural and energetic observables on which our conclusions rest are nevertheless stationary across the production trajectory. A further limitation concerns the docking stage itself. Validation against independent molecular dynamics reference poses showed that, although the predicted binding site and affinity are robust and reproducible, the top-ranked Vina pose does not reach the 2 Å accuracy threshold, and raising the exhaustiveness does not correct this (Supplementary Figure S4). The reproducibility of the docking scores therefore reflects the precision of the search and not the accuracy of the pose. The docking results are consequently used only to prioritise compounds, and the binding mode analysed throughout this work is the molecular dynamics ensemble rather than the static docked pose. A further limitation applies to the electronic structure calculations. These were performed on the isolated molecule in the gas phase, so the reported descriptors do not include the effect of the aqueous environment or of the protein field. A continuum solvation treatment, together with molecular electrostatic potential mapping and Fukui function analysis, would be required before the descriptors could be used quantitatively, and they are therefore presented here in a supporting and comparative role only. Redocking against a crystallographic reference and retrospective ROC/AUC enrichment remain impossible for MARV VP40 until an experimentally resolved protein–ligand complex, or a set of experimentally confirmed inhibitors, becomes available; both should be revisited when such data emerge.

Conclusion

The present study employed an integrated computational approach to explore the antiviral potential of phytochemicals from *Ocimum basilicum* against the Marburg virus VP40 protein. Among the screened compounds, rutin obtained the most favourable docking score and was selected on that basis for detailed study. Molecular dynamics simulations further supported the structural stability of the VP40-rutin complex, indicating persistent non-covalent interactions with key functional residues involved in viral assembly and membrane association. In addition, gas-phase DFT analysis provided complementary evidence of favorable electronic properties, including charge transfer potential and molecular reactivity, which may contribute to ligand–protein interaction stability. These descriptors were obtained for the isolated molecule without a continuum solvation treatment, so they are presented in a supporting and comparative role rather than as quantitative predictors of affinity. Set against these findings, rutin carries a substantial pharmacokinetic liability. It violates three of Lipinski's criteria, has a topological polar surface area of 269.43 Å² and a predicted human intestinal absorption of only 23.45%, so it should be regarded as a lead scaffold requiring structural or formulation optimisation rather than a ready-to-use agent, in the terms already set out in the Abstract, the Results and the Discussion. Moreover, the study is limited by its

purely computational nature, absence of experimental validation, and simplified modeling of the viral protein system. Therefore, these results should be interpreted as preliminary mechanistic insights rather than definitive antiviral evidence. Future studies should focus on in vitro and in vivo validation, as well as structural optimization or formulation strategies to improve the pharmacokinetic profile of rutin and related flavonoid scaffolds for potential antiviral development against Marburg virus infection. It should be emphasised that the docking stage, although reproducible and site-specific, does not rank the pose geometry of rutin to within 2 Å of the simulated binding mode, and that the predicted binding pocket remains computationally rather than experimentally defined. The conclusions drawn here therefore rest on the dynamic and thermodynamic analyses, and require experimental confirmation.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-026-71976-3>.

Acknowledgements The authors sincerely acknowledge the support of BioPC—A Bioinformatics Lab of Research and Training for their valuable assistance in the design of this study. We also extend our gratitude to Aishee Devi, Samsad Shamima Ridy, Md Sadman Yasir, Muhareb Hussain, Tanjuma Tasnim Hira, Ifthesum, and Tasnin Neha for their contributions and support throughout the research process.

Author contributions Conceptualization : M. F. R., M.M.K, M.H.A; Methodology : all authors; Validation : M. F. R., M. H. A.; Formal analysis : all authors; Investigation : M. F. R., M. M. K., M. H. A.; Resources : all authors; Data curation : all authors; Writing—original draft preparation : M. F. R.; Writing-review and editing : all authors; Visualization : M. F. R., M. H. A.; Supervision : M. H. A.; Project administration : M. M. K., M. H. A., M. F. R.

Data availability All data generated or analyzed during this study are included in this published article. Further information and datasets are available from the corresponding author upon reasonable request.

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. WHO. World Health Organisation. Marburg virus disease (fact sheet). Available online at: <https://www.who.int/news-room/fact-sheets/detail/marburg-virus-disease> (Accessed on 2023 Feb 11). (2021).
2. Uppala, P. K., Karanam, S. K., Kandra, N. V. & Edhi, S. Marburg virus disease: Emerging threat, pathogenesis, and global public health strategies. *World J Virol***14**, (2025).
3. Francis, D. L., Reddy, S. S. P., Logaranjani, A., Karthikeyan, S. S. S. & Rath, M. Marburg virus disease: Pathophysiology, diagnostic challenges, and global health preparedness strategies. *Ann. Glob. Health***91**, 24 (2025).
4. Amiar, S. et al. Lipid-specific oligomerization of the Marburg virus matrix protein VP40 is regulated by two distinct interfaces for virion assembly. *J. Biol. Chem.* <https://doi.org/10.1016/j.jbc.2021.100796> (2021).
5. Harty, R. N. No exit: Targeting the budding process to inhibit filovirus replication. *Antivir. Res.***81**, 189–197 (2009).
6. Dubey, A. K., Chaudhry, S. K., Singh, H. B., Gupta, V. K. & Kaushik, A. Perspectives on nano-nutraceuticals to manage pre and post COVID-19 infections. *Biotechnol. Rep.***33**, e00712 (2022).

7. Patel, C. N., Jani, S. P., Kumar, S. P., Modi, K. M. & Kumar, Y. Computational investigation of natural compounds as potential main protease (Mpro) inhibitors for SARS-CoV-2 virus. *Comput. Biol. Med.* **151**, 106318 (2022).
8. Patel, C. N. et al. Excavating phytochemicals from plants possessing antiviral activities for identifying SARS-CoV hemagglutinin-esterase inhibitors by diligent computational workflow. *J. Biomol. Struct. Dyn.* **41**, 2382–2397 (2023).
9. Guan, H., Sun, H. & Zhao, X. Application of density functional theory to molecular engineering of pharmaceutical formulations. *Int. J. Mol. Sci.* **26**, 3262 (2025).
10. Bais, H. P., Walker, T. S., Schweizer, H. P. & Vivanco, J. M. Root specific elicitation and antimicrobial activity of rosmarinic acid in hairy root cultures of *Ocimum basilicum*. *Plant Physiol. Biochem.* **40**, 983–995 (2002).
11. Bisaria, V. & Panda, A. Large-scale plant cell culture: methods, applications and products. *Curr. Opin. Biotechnol.* **2**, 370–374 (1991).
12. Bilal, A. et al. Phytochemical and pharmacological studies on *Ocimum basilicum* Linn-A review. *Int. J. Curr. Res. Rev.* **4**, 73–83 (2012).
13. Azizah, N. S. et al. Sweet basil (*Ocimum basilicum* L.) -A review of its botany, phytochemistry, pharmacological activities, and biotechnological development. *Plants* **12**, 4148 (2023).
14. Mulugeta, S. M., Gosztola, B. & Radácsi, P. Diversity in morphology and bioactive compounds among selected *Ocimum* species. *Biochem. Syst. Ecol.* **114**, 104826 (2024).
15. Kamelnia, E., Mohebbati, R., Kamelnia, R., El-Seedi, H. R. & Boskabady, M. H. Anti-inflammatory, immunomodulatory and anti-oxidant effects of *Ocimum basilicum* L. and its main constituents: A review. *Iran. J. Basic Med. Sci.* **26**, 617 (2023).
16. Adigüzel, A. et al. Antimicrobial effects of *Ocimum basilicum* (Labiatae) extract. *Turk. J. Biol.* **29**, 155–160 (2005).
17. Cadet, E. et al. Antibacterial activity and antibiotic-potentiating effects of methanol extracts from *Ocimum basilicum* and *Sarcocephalus latifolius* against multidrug-resistant Gram-negative bacteria overexpressing efflux pumps. *Pain* **16**, 17 (2024).
18. Amrani, S. et al. Vasorelaxant and anti-platelet aggregation effects of aqueous *Ocimum basilicum* extract. *J. Ethnopharmacol.* **125**, 157–162 (2009).
19. Freire, C. M. M., Marques, M. O. M. & Costa, M. Effects of seasonal variation on the central nervous system activity of *Ocimum gratissimum* L. essential oil. *J. Ethnopharmacol.* **105**, 161–166 (2006).
20. Al-Harbi, N. A. et al. Evaluation of insecticidal effects of plants essential oils extracted from basil, black seeds and lavender against *Sitophilus oryzae*. *Plants* **10**, 829 (2021).
21. Meera, R., Devi, P., Kameswari, B., Madhumitha, B. & Merlin, N. Antioxidant and hepatoprotective activities of *Ocimum basilicum* Linn. and *Trigonella foenum-graecum* Linn. against H₂O₂ and CCL₄ induced hepatotoxicity in goat liver. *Indian J. Exp. Biol.* **47**, 584 (2009).
22. Naithani, R., Mehta, R. G., Shukla, D., Chandrasekera, S. N. & Moriarty, R. M. Antiviral activity of phytochemicals: A current perspective. In *Dietary Components and Immune Function* 421–468 (Springer, 2010).
23. Oda, S. et al. Crystal structure of Marburg virus VP40 reveals a broad, basic patch for matrix assembly and a requirement of the N-terminal domain for immunosuppression. *J. Virol.* **90**, 1839–1848 (2016).
24. Arman, M., Alam, S., Maruf, R. A., Shams, Z. & Islam, M. N. Molecular modeling of some commercially available antiviral drugs and their derivatives against SARS-CoV-2 infection. *Narra J.* **4**, e319 (2024).
25. Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723 (1997).
26. Siddiquee, N. H. et al. In-silico identification of novel natural drug leads against the Ebola virus VP40 protein: A promising approach for developing new antiviral therapeutics. *Informatics Med. Unlocked* **45**, 101458 (2024).
27. Eberhardt, J., Santos-Martins, D., Tillack, A. F. & Forli, S. AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings. *J. Chem. Inform. Model.* **61**, 3891 (2021).
28. Morris, G. M. et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).
29. Bronowska, A. K. Thermodynamics of ligand-protein interactions: Implications for molecular design. In *Thermodynamics-Interaction Studies-Solids, Liquids and Gases* (IntechOpen, 2011).
30. Lagorce, D., Douguet, D., Miteva, M. A. & Villoutreix, B. O. Computational analysis of calculated physicochemical and ADMET properties of protein-protein interaction inhibitors. *Sci. Rep.* **7**, 46277 (2017).
31. Shrake, A. & Rupley, J. A. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *J. Mol. Biol.* **79**, 351–371 (1973).
32. Baker, E. N. & Hubbard, R. E. Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* **44**, 97–179 (1984).
33. Amadei, A., Linssen, A. B. & Berendsen, H. J. Essential dynamics of proteins. *Proteins Struct. Funct. Bioinform.* **17**, 412–425 (1993).
34. Ichiye, T. & Karplus, M. Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins: Struct. Function, Bioinform.* **11**, 205–217 (1991).
35. Parr, R. G. Density functional theory of atoms and molecules. In *Horizons of Quantum Chemistry* 5–15 (Springer, 1989).

36. Pearson, R. G. Absolute electronegativity and hardness correlated with molecular orbital theory. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 8440–8441 (1986).
37. Patel, K. & Patel, D. K. The beneficial role of rutin, a naturally occurring flavonoid in health promotion and disease prevention: A systematic review and update. In *Bioactive Food as Dietary Interventions for Arthritis and Related Inflammatory Diseases* 457–479 (2019).
38. Prasad, R. & Prasad, S. B. A review on the chemistry and biological properties of Rutin, a promising nutraceutical agent. *Asian J. Pharm. Pharmacol.* **5**, 1–20 (2019).
39. Wang, Y., He, Z. & Deng, S. Ursolic acid reduces the metalloprotease/anti-metalloprotease imbalance in cerebral ischemia and reperfusion injury. *Drug design, development and therapy* 1663–1674 (2016).
40. Yu, S.-G. et al. Ursolic acid derivative ameliorates streptozotocin-induced diabetic bone deleterious effects in mice. *Int. J. Clin. Exp. Pathol.* **8**, 3681 (2015).
41. Sultana, N. Clinically useful anticancer, antitumor, and antiwrinkle agent, ursolic acid and related derivatives as medicinally important natural product. *J. Enzyme Inhib. Med. Chem.* **26**, 616–642 (2011).
42. Mlala, S., Oyediji, A. O., Gondwe, M. & Oyediji, O. O. Ursolic acid and its derivatives as bioactive agents. *Molecules* **24**, 2751 (2019).
43. Shanmugam, M. K. et al. Ursolic acid in cancer prevention and treatment: Molecular targets, pharmacokinetics and clinical studies. *Biochem. Pharmacol.* **85**, 1579–1587 (2013).
44. Kashyap, D., Sharma, A., Tuli, H., Punia, S. & Sharma, A. Ursolic acid and oleanolic acid: pentacyclic terpenoids with promising anti-inflammatory activities. *Recent Patents Inflamm. Aller. Drug Discov.* **10**, 21–33 (2016).
45. Das, R., Bhattarai, A., Tamang, B. & Thakur, N. Identification of potential VP40 inhibitor of Marburg virus through molecular docking, pharmacokinetic analysis and molecular dynamics simulation. *Sci. Rep.* **15**, 27129 (2025).
46. Alanzi, A. R., Alajmi, M. F., Al-Dosari, M. S., Parvez, M. K. & Alqahtani, M. J. In silico exploration of deep-sea fungal metabolites as inhibitor of Ebola and Marburg VP35 and VP40. *PLoS ONE* **19**, e0307579 (2024).
47. Broni, E. et al. Cheminformatics-based study identifies potential Ebola VP40 inhibitors. *Int. J. Mol. Sci.* **24**, 6298 (2023).
48. Urata, S. et al. Identification of novel chemical compounds targeting filovirus VP40-mediated particle production. *Antivir. Res.* **199**, 105267 (2022).
49. Sharma, G., Sharma, A. R. & Kim, J.-C. Recent advancements in the therapeutic development for Marburg virus: Updates on clinical trials. *Curr. Infect. Dis. Rep.* **26**, 57–67 (2024).
50. Srivastava, S. et al. Novel antiviral approaches for Marburg: A promising therapeutics in the pipeline. *Front. Microbiol.* **15**, 1387628 (2024).
51. Koehler, A., Pfeiffer, S., Kolesnikova, L. & Becker, S. Analysis of the multifunctionality of Marburg virus VP40. *J. Gen. Virol.* **99**, 1614–1620 (2018).
52. Wijesinghe, K. J. & Stahelin, R. V. Investigation of the lipid binding properties of the Marburg virus matrix protein VP40. *J. Virol.* **90**, 3074–3085 (2016).
53. Mercer, J. & Helenius, A. Virus entry by macropinocytosis. *Nat. Cell Biol.* **11**, 510–520 (2009).
54. Motsa, B. B. et al. Variants of the Ebola virus matrix protein VP40 have differential effects on oligomerization and plasma membrane interactions. *Journal of Biological Chemistry* 110489 (2025).
55. Choi, S.-S., Park, H.-R. & Lee, K.-A. A comparative study of rutin and rutin glycoside: Antioxidant activity, anti-inflammatory effect, effect on platelet aggregation and blood coagulation. *Antioxidants* **10**, 1696 (2021).
56. Dessen, A., Volchkov, V., Dolnik, O., Klenk, H. & Weissenhorn, W. Crystal structure of the matrix protein VP40 from Ebola virus. *EMBO J.* **19**, 4228–4236 (2000).
57. Stahelin, R. V. Membrane binding and bending in Ebola VP40 assembly and egress. *Front. Microbiol.* **5**, 300 (2014).
58. Wan, W. et al. Ebola and Marburg virus matrix layers are locally ordered assemblies of VP40 dimers. *Elife* **9**, e59225 (2020).
59. Liu, K., Watanabe, E. & Kokubo, H. Exploring the stability of ligand binding modes to proteins by molecular dynamics simulations. *J. Comput. Aided Mol. Des.* **31**, 201–211 (2017).
60. Motsa, B. B., Sharma, T., Cioffi, M. D., Chapagain, P. P. & Stahelin, R. V. Minor electrostatic changes robustly increase VP40 membrane binding, assembly, and budding of Ebola virus matrix protein derived virus-like particles. *J. Biol. Chem.* <https://doi.org/10.1016/j.jbc.2024.107213> (2024).
61. Del Vecchio, K. et al. A cationic, C-terminal patch and structural rearrangements in Ebola virus matrix VP40 protein control its interactions with phosphatidylserine. *J. Biol. Chem.* **293**, 3335–3349 (2018).
62. Scianimanico, S. et al. Membrane association induces a conformational change in the Ebola virus matrix protein. *EMBO J.* <https://doi.org/10.1093/emboj/19.24.6732> (2000).
63. Johnson, K. A., Taghon, G. J., Scott, J. L. & Stahelin, R. V. The Ebola virus matrix protein, VP40, requires phosphatidylinositol 4, 5-bisphosphate (PI (4, 5) P2) for extensive oligomerization at the plasma membrane and viral egress. *Sci. Rep.* **6**, 19125 (2016).
64. Mahmud, A.-S., Rabby, M. F., Rahman, M. F. & Begum, A. GC–MS–identified phytochemicals from *Dimocarpus longan* leaves as promising HER2 inhibitors for breast cancer: Molecular docking, ADMET, and toxicity assessment. *Nat. Prod. Anal.* **2**, 100002 (2026).

65. Huang, Y., Rong, C., Zhang, R. & Liu, S. Evaluating frontier orbital energy and HOMO/LUMO gap with descriptors from density functional reactivity theory. *J. Mol. Model.* **23**, 3 (2017).
66. Yu, J., Su, N. Q. & Yang, W. Describing chemical reactivity with frontier molecular orbitals. *JACS Au* **2**, 1383–1394 (2022).
67. Chakraborty, D. & Chattaraj, P. K. Conceptual density functional theory based electronic structure principles. *Chem. Sci.* **12**, 6264–6279 (2021).
68. Fukui, K. Role of frontier orbitals in chemical reactions. *Science* **218**, 747–754 (1982).
69. Chandrakumar, K. & Pal, S. The concept of density functional theory based descriptors and its relation with the reactivity of molecular systems: A semi-quantitative study. *Int. J. Mol. Sci.* **3**, 324–337 (2002).
70. Parr, R. G. & Yang, W. Density functional approach to the frontier-electron theory of chemical reactivity. *J. Am. Chem. Soc.* **106**, 4049–4050 (1984).
71. Berisha, A. Computational discovery of glioblastoma multiforme inhibitors from medicinal fungi: An integrated approach combining virtual screening, DFT, and toxicity profiling. *Silico Res. Biomed.* <https://doi.org/10.1016/j.insci.2025.100023> (2025).
72. Babalola, O. O. et al. Integrating phytochemicals and in silico methods for modern drug discovery: A comprehensive review. *Discov. Chem.* **2**, 297 (2025).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Md. Foyzur Rahman¹ · Md. Mustak Khan² · Md. Hridoy Ahmed^{3,10} · Shishir Dutta⁴ · Farah Naima Tapti⁵ · Umme Marzia Tasnim⁶ · Marwah Talukder⁷ · Md. Mahfuzur Rahman⁸ · Md. Foisal Imtiaz⁵ · Md. Taleb Hossain Roman⁹ · Laila Khaleda Anghi³ · Mohammad Al-Forkan³

✉ Md. Hridoy Ahmed
hridoy.geb.cu@gmail.com

¹ Department of Pharmacy, Dhaka International University, Dhaka, Bangladesh

² Department of Biochemistry and Molecular Biology, University of Chittagong, Chittagong 4331, Bangladesh

³ Department of Genetic Engineering and Biotechnology, University of Chittagong, Chittagong 4331, Bangladesh

⁴ Department of Soil Science, University of Chittagong, Chittagong, Bangladesh

⁵ Department of Pharmacy, Jahangirnagar University, Savar, Dhaka, Bangladesh

⁶ Department of Biochemistry and Molecular Biology, University of Dhaka, Dhaka, Bangladesh

⁷ Department of Pharmaceutical Technology, Faculty of Pharmacy, Universiti Malaya, 50603 Kuala Lumpur, Malaysia

⁸ Department of Microbiology, Jagannath University, Dhaka, Bangladesh

⁹ Department of Pharmacy, International Islamic University Chittagong, Chittagong, Bangladesh

¹⁰ BioPC - A Bioinformatics Lab of Research and Training, Chittagong, Bangladesh