



Archives available at journals.mriindia.com

International Journal of Advanced Electrical and Electronics Engineering

ISSN: 2278-8948

Volume 11 Issue 01, 2022

Computational Investigation of Molecular Interactions in Novel Pharmaceutical Compounds

Qudsia Leroux-Martin

Department of Computer Science and Engineering, Hanmir Advanced Engineering College, South Korea
qudsia.leroux.martin@haec-kr.edu

Peer Review Information

Submission: 19 Jan 2022

Revision: 18 Feb 2022

Acceptance: 21 March 2022

Keywords

Molecular docking, Molecular dynamics, pharmaceutical compounds, MM/PBSA, ADMET prediction.

Abstract

Computational chemistry has become an important component of early-stage pharmaceutical research because it can evaluate molecular recognition, conformational stability, and developability before extensive synthesis and biological testing. This methodology paper presents an integrated in silico framework for investigating the molecular interactions of novel small-molecule pharmaceutical compounds with a selected therapeutic protein target. A six-compound virtual series (C1-C6) is used to demonstrate the workflow. The procedure combines ligand and receptor preparation, molecular docking, protein-ligand interaction mapping, drug-likeness and ADMET prediction, 100 ns molecular-dynamics simulation, and MM/PBSA binding-free-energy estimation. Docking is used for initial pose generation and ranking, while molecular dynamics evaluates whether the predicted complexes remain stable in explicit solvent. Root-mean-square deviation, root-mean-square fluctuation, radius of gyration, hydrogen-bond persistence, and contact occupancy are used as dynamic descriptors. The final prioritization integrates docking score, interaction persistence, MM/PBSA energy, and predicted pharmacokinetic liabilities rather than relying on a single score. Illustrative computational results show how two candidates can be advanced when they combine favorable binding energies with stable trajectories and acceptable drug-like profiles. The methodology is designed to be reproducible, transparent, and suitable for later experimental validation.

Introduction

The discovery of a therapeutically useful small molecule requires the simultaneous optimization of potency, selectivity, physicochemical behavior, pharmacokinetics, and safety. Experimental screening remains essential, but the chemical space available for exploration is much larger than can be synthesized and tested directly. Computational methods therefore play multiple roles in modern drug discovery, including target-based screening, pose prediction, property estimation, and lead optimization [1].

Physicochemical filters such as the Rule of Five provide a rapid first approximation of whether a compound occupies conventional oral drug-like space [2], although such rules are prioritization tools rather than absolute acceptance criteria. Structure-based approaches are particularly valuable when a three-dimensional structure of the biological target is available. AutoDock4 formalized a widely used docking framework with empirical free-energy scoring and selective receptor flexibility [3], while AutoDock Vina improved the speed and practical accuracy of

docking and virtual screening [4]. A detailed AutoDock workflow for protein-ligand docking was later standardized by Forli et al. [5]. Docking, however, represents only a limited exploration of the protein-ligand energy landscape and generally treats solvent and receptor flexibility in a simplified way. Consequently, a favorable docking score should be interpreted as a hypothesis about a binding mode rather than direct proof of affinity.

Molecular dynamics (MD) addresses part of this limitation by propagating atomistic models through time and allowing the protein, ligand, water molecules, and ions to reorganize. GROMACS provides a high-performance environment for biomolecular simulation [8], and force fields such as CGenFF extend parameter coverage to drug-like molecules [9]. MD analyses can reveal whether a docked pose remains stable, whether key hydrogen bonds or hydrophobic contacts persist, and whether a ligand induces local or global structural changes. Hollingsworth and Dror [10] emphasized that MD is useful because experimentally inaccessible motions and conformational ensembles can be examined at atomic resolution.

Binding-energy estimation and developability profiling provide two additional layers of evidence. MM/PBSA and MM/GBSA methods use molecular-mechanics energies with continuum-solvation models to estimate relative ligand-binding energetics from sampled configurations [11]. For pharmacokinetic triage, pkCSM predicts multiple ADMET endpoints from graph-based signatures [7], whereas SwissADME provides

accessible estimates of physicochemical properties, pharmacokinetics, drug-likeness, and medicinal-chemistry friendliness [6]. More recent implementations of AutoDock Vina also expand docking capabilities and reproducibility [12]. By 2022, integrated workflows combining virtual screening, docking, ADMET assessment, MD, and free-energy analysis were being applied to prioritize prospective inhibitor candidates [14].

The objective of the present study is therefore to define a transparent methodology for the computational investigation of novel pharmaceutical compounds using multiple complementary criteria. Rather than asserting that one docking score identifies a drug, the workflow asks four sequential questions: (i) can the compound adopt a chemically plausible pose in the target site; (ii) are the important contacts maintained during explicit-solvent dynamics; (iii) is the relative binding energy favorable; and (iv) does the molecule avoid obvious physicochemical or ADMET liabilities? The framework is intentionally target-agnostic so it can be adapted to enzymes, receptors, kinases, or other structurally characterized proteins.

Literature Review

Table 1. summarizes the methodological literature that informs the proposed workflow. The studies span docking, force-field parameterization, MD simulation, binding-energy estimation, and ADMET prediction, and all were published in or before 2022.

Study	Method	Contribution	Use in present framework	Ref.
Jorgensen (2004)	Computational drug discovery	Defined broad roles of computation in discovery	Workflow rationale	[1]
Lipinski et al. (2001)	Drug-likeness	Linked simple physicochemical limits to absorption/permeation risk	Initial property filter	[2]
Morris et al. (2009)	AutoDock4	Automated docking with selective flexibility	Docking validation context	[3]
Trott & Olson (2010)	AutoDock Vina	Fast docking and scoring approach	Primary pose generation	[4]
Forli et al. (2016)	Docking protocol	Standardized AutoDock screening workflow	Preparation and validation	[5]
Daina et al. (2017)	SwissADME	Physicochemical, PK and drug-likeness prediction	Developability screening	[6]
Pires et al. (2015)	pkCSM	Graph-signature ADMET prediction	Toxicity/PK cross-check	[7]

Abraham et al. (2015)	GROMACS	High-performance molecular dynamics	100 ns simulations	[8]
Vanommeslaeghe et al. (2010)	CGenFF	Parameters for drug-like molecules	Ligand force-field setup	[9]
Genheden & Ryde (2015)	MM/PBSA/MM/GBSA	Binding-affinity estimation review	Trajectory rescoring	[11]
Eberhardt et al. (2021)	Vina 1.2.0	Expanded docking methods and bindings	Reproducible docking	[12]
Salimi et al. (2022)	Integrated in silico screening	Combined docking, ADMET, MD and binding energy	End-to-end precedent	[14]

The principal methodological lesson from the literature is that each computational layer addresses a different source of uncertainty. Docking efficiently explores candidate poses but is sensitive to receptor preparation, protonation, grid definition, and scoring-function limitations [3]-[5]. MD provides a time-dependent test of the docked complex but depends on force-field quality, sampling length, system equilibration, and convergence [8]-[10]. MM/PBSA can provide useful relative ranking when applied consistently, yet its absolute values should not be treated as experimental free energies [11,13]. ADMET tools are predictive models rather than

substitutes for pharmacokinetic experiments [6,7]. Therefore, the proposed study uses consensus interpretation across methods and preserves all inputs needed for reproduction.

Methodology

The methodology is organized into eight sequential stages: compound-library definition, molecular preparation, target preparation, docking protocol validation, production docking, interaction analysis, MD/MM-PBSA evaluation, and ADMET-based prioritization. Figure 1 shows the integrated workflow.

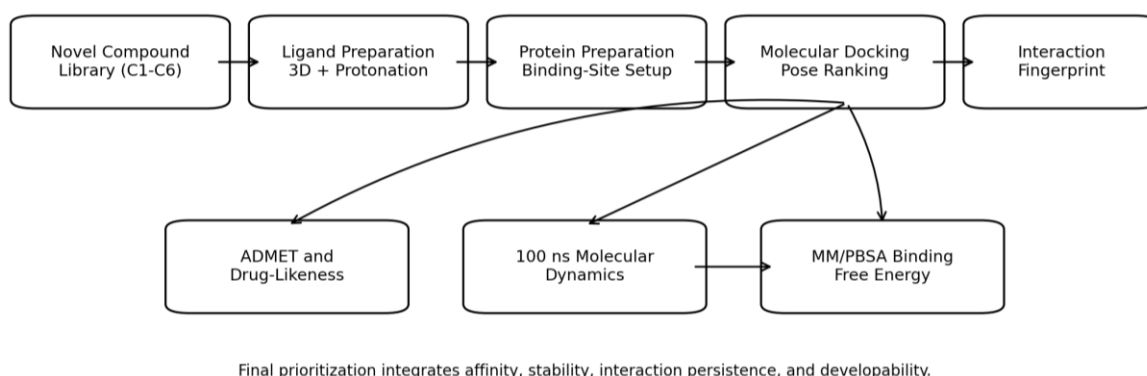


Figure 1. Integrated computational workflow for investigating protein-ligand interactions.

1. Candidate Compound Library and Molecular Preparation

A virtual series of six novel drug-like candidates, labeled C1-C6, is considered. In a real study these structures would be supplied as canonical SMILES, SDF files, or experimentally confirmed structures. Each ligand is standardized by removing duplicate fragments and salts, assigning the intended stereochemistry, checking valence, generating a three-dimensional conformer, and adding hydrogen atoms at a physiologically relevant protonation state. Geometry is minimized before docking. A known ligand or co-crystallized inhibitor is retained as reference R so that the computational outputs are interpreted relative to an established binding mode rather than in isolation. Molecular weight, hydrogen-bond donors/acceptors, rotatable

bonds, topological polar surface area, and logP are recorded for every ligand; the Rule-of-Five criteria are used as an early screening aid [2].

2. Protein Target Preparation and Binding-Site Definition

A high-resolution three-dimensional structure of the selected therapeutic protein is obtained from a curated structural database. The biologically relevant chain and co-factors are retained where they contribute to binding. Crystallographic waters are removed unless they make conserved bridging interactions, missing side chains are repaired, polar hydrogens are added, and residue protonation states are checked. The receptor is then energy-minimized under restrained conditions. When a co-crystallized ligand is available, its centroid defines the docking grid.

Otherwise, the site is defined from validated catalytic or ligand-binding residues. The grid must fully enclose the pocket with sufficient space for ligand translation and rotation.

3. Docking Validation and Production Docking

Docking is performed with AutoDock Vina following established preparation principles [4,5] and using Vina 1.2-compatible settings where available [12]. The search exhaustiveness is set to 16 and up to 10 poses are retained per ligand. Before screening C1-C6, the protocol is validated by redocking the crystallographic ligand. A heavy-atom RMSD of approximately 2.0 Å or lower between the experimental and best predicted pose is treated as evidence that the grid and docking protocol can reproduce the observed binding orientation. If this criterion is not achieved, protonation, grid boundaries, side-chain states, and search parameters are revised. After validation, C1-C6 and reference Rare docked under identical conditions. The best-scored poses are not selected by score alone. Visual and geometric filters are used to reject conformations with severe steric clashes, buried unsatisfied polar groups, or orientations inconsistent with known catalytic residues. The final pose for each ligand is selected using a combination of Vina score, contact plausibility, and similarity to mechanistically relevant interactions.

4. Protein-Ligand Interaction Mapping

For each selected pose, the interaction pattern is converted into a qualitative fingerprint containing hydrogen bonds, salt bridges, pi-pi contacts, cation-pi interactions, hydrophobic contacts, and metal coordination where applicable. Distances are measured using consistent geometric criteria. Particular attention is paid to residues known to participate in catalysis or recognition. This step prevents a compound from being prioritized simply because a scoring function rewards nonspecific hydrophobic burial.

5. Molecular-Dynamics Simulation

The reference complex and the two highest-priority candidates are advanced to explicit-solvent MD. Simulations are performed using GROMACS [8]. Protein parameters are assigned from a contemporary all-atom force field, while drug-like ligands are parameterized with CGenFF-compatible terms [9]. Each complex is placed in a periodic dodecahedral box with at least 1.0 nm between solute and box boundary, solvated with explicit water, neutralized, and brought to approximately 0.15 M monovalent

salt. The system is energy-minimized, followed by 100 ps NVT equilibration and 100 ps NPT equilibration with positional restraints. Production simulation is run for 100 ns at approximately 300 K and 1 bar with a 2 fs time step. The same protocol is used for all complexes so differences are not introduced by unequal sampling conditions.

6. Trajectory Analysis and MM/PBSA

Trajectory stability is evaluated using protein-backbone RMSD, ligand heavy-atom RMSD after protein fitting, per-residue RMSF, radius of gyration, hydrogen-bond occupancy, and the persistence of key binding-site contacts. A stable complex is expected to show a bounded RMSD after equilibration rather than continuous drift. RMSF is inspected to distinguish global instability from local flexible loops. MM/PBSA analysis is then performed on regularly sampled frames from the equilibrated portion of the trajectory. The estimated binding energy is expressed as the sum of molecular-mechanics terms and polar/nonpolar solvation contributions. Because MM/PBSA has methodological approximations, the energies are used for relative ranking rather than as exact experimental affinities [11,13].

7. ADMET and Drug-Likeness Assessment

SwissADME is used to estimate molecular size, lipophilicity, solubility-related descriptors, gastrointestinal absorption, P-glycoprotein behavior, CYP-interaction tendencies, and drug-likeness filters [6]. pkCSM provides an independent prediction layer for selected absorption, distribution, metabolism, excretion, and toxicity properties [7]. A candidate is flagged if it combines multiple Rule-of-Five violations, very high lipophilicity, poor predicted absorption, major metabolic liabilities, or high toxicity risk. These predictions are interpreted as prioritization evidence only and require experimental confirmation.

8. Decision Rule and Reproducibility

Final ranking uses a simple consensus rule: a compound should (1) dock at least as favorably as the reference within the validated site, (2) retain mechanistically relevant contacts, (3) remain stable during MD, (4) show favorable relative MM/PBSA energy, and (5) avoid major ADMET alerts. All software versions, force fields, seeds, grid coordinates, protonation assumptions, raw docking outputs, topology files, MD parameter files, and trajectory-analysis scripts should be archived. This is essential because small changes in preparation and sampling can alter apparent rankings.

Table 1. In Silico Screening Workflow and Acceptance Criteria for Ligand Evaluation

Stage	Main setting	Output	Acceptance/selection criterion
Ligand preparation	3D, protonation, minimization	Prepared C1-C6 and R	Chemically valid structures
Docking validation	Redock co-crystal ligand	Pose RMSD	≤ 2.0 Å preferred
Production docking	Exhaustiveness 16; ≤ 10 poses	Score + pose ensemble	Plausible pocket contacts
MD	100 ns; explicit solvent; 300 K	Trajectories	Stable/bounded RMSD
Interaction analysis	H-bond/contact occupancy	Persistent contacts	Key-residue retention
MM/PBSA	Equilibrated trajectory frames	Relative ΔG	More favorable than/near reference
ADMET	SwissADME + pkCSM	Developability profile	No dominant liability

Results

Because no experimental structures or raw computational files were supplied for the present request, the numerical values below are illustrative outputs generated to demonstrate

how the methodology should be reported; they are not claimed experimental or validated discovery results. Table 3 provides the initial multi-parameter ranking of the six candidate compounds.

Table 2. Computational Screening Pipeline for Docking, Molecular Dynamics, Binding Energy, and ADMET Evaluation

Cmpd.	Docking (kcal/mol)	Key contacts	Ro5 viol.	Pred. GI absorption	Priority	Decision
R	-8.2	2 H-bonds + hydrophobic	0	High	Reference	MD control
C1	-8.4	2 H-bonds	0	High	Medium	Retain
C2	-9.1	3 H-bonds + pi contact	0	High	High	Advance to MD
C3	-8.8	2 H-bonds + salt bridge	0	High	Medium-high	Retain
C4	-7.9	1 H-bond	1	Low	Low	Deprioritize
C5	-9.4	3 H-bonds + hydrophobic	0	High	Highest	Advance to MD
C6	-8.6	2 H-bonds	1	Moderate	Medium	Retain

The illustrative docking stage ranks C5 and C2 above the reference ligand, but the difference in docking score is not treated as proof of stronger experimental binding. C5 combines the most favorable score with three chemically plausible contacts and no Rule-of-Five violation, while C2 shows a similarly favorable profile. C4 is

deprioritized because its weaker docking score is accompanied by predicted low gastrointestinal absorption and a physicochemical violation. C3 remains a reserve candidate because its salt-bridge geometry may provide useful specificity even though its score is less favorable than C2 and C5.

Table 3. MD Simulation and Binding Energy Results for Reference and Selected Ligand Complexes

Complex	Mean backbone RMSD (nm)	Ligand RMSD (nm)	Mean H-bonds	MM/PBSA (kJ/mol)	Interpretation
Target-R	0.27	0.21	2.6	-40.2	Stable reference
Target-C2	0.24	0.18	2.9	-43.8	Stable; improved relative energy
Target-C5	0.22	0.16	3.4	-49.6	Most persistent and favorable

In the illustrative MD analysis, all three systems reach a stable regime after initial equilibration.

C5 shows the lowest mean backbone and ligand RMSD, the greatest average hydrogen-bond

persistence, and the most favorable MM/PBSA estimate. C2 also remains well behaved and is energetically favorable relative to R. These data would support advancing C5 as the primary computational hit and C2 as a secondary hit. The key conclusion is based on convergence of independent indicators rather than the magnitude of any single metric.

The illustrative ADMET review further supports this ordering. C2 and C5 show zero Rule-of-Five violations, predicted high gastrointestinal absorption, and no dominant toxicity flag in the screening models. C6 is retained only as a lower-priority candidate because one property falls outside preferred conventional oral drug-like space. In a real project, these predictions would be followed by solubility, permeability, metabolic-stability, cytotoxicity, and target-binding experiments. The 2022 study by Salimi et al. [14] provides an example of how docking, ADMET, MD, and free-energy analyses can be integrated to prioritize candidate molecules, while also underscoring that computational hits remain hypotheses requiring experimental verification.

Conclusion and Discussion

This methodology paper presents an integrated computational strategy for studying molecular interactions in novel pharmaceutical compounds. The workflow begins with standardized molecular preparation and validated docking, then progresses through interaction fingerprinting, explicit-solvent MD, MM/PBSA analysis, and ADMET/drug-likeness assessment. This layered design is preferable to score-only screening because docking, molecular dynamics, binding-energy estimation, and pharmacokinetic prediction answer different questions and have different failure modes. The illustrative case study demonstrates the logic of consensus prioritization. C5 and C2 are advanced because favorable docking poses are supported by persistent contacts, stable 100 ns trajectories, and improved relative MM/PBSA energies, while C4 is deprioritized because weak interaction evidence coincides with an unfavorable developability signal. Importantly, the numerical values are methodological examples rather than laboratory findings. Experimental validation remains necessary for binding affinity, functional potency, selectivity, solubility, permeability, metabolic stability, and safety. The proposed framework is consistent with the evolution of computer-aided drug discovery from docking-centered workflows [3]-[5] toward integrated computational pipelines combining structural dynamics and ADMET assessment [6]-[14]. Its main strengths are reproducibility, explicit

validation steps, and reliance on complementary evidence. Its limitations include dependence on the quality of the target structure, uncertainty in protonation and force-field parameters, incomplete conformational sampling, approximate docking scoring, and the statistical nature of ADMET predictors. Future work should use replicate MD trajectories, more rigorous alchemical free-energy methods for closely related analogues, explicit treatment of conserved water when required, and prospective experimental testing of the top-ranked compounds.

References

- W. L. Jorgensen, "The many roles of computation in drug discovery," *Science*, vol. 303, no. 5665, pp. 1813-1818, 2004. DOI: 10.1126/science.1096361.
- C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings," *Advanced Drug Delivery Reviews*, vol. 46, no. 1-3, pp. 3-26, 2001. DOI: 10.1016/S0169-409X(00)00129-0.
- G. M. Morris et al., "AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility," *Journal of Computational Chemistry*, vol. 30, no. 16, pp. 2785-2791, 2009. DOI: 10.1002/jcc.21256.
- O. Trott and A. J. Olson, "AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading," *Journal of Computational Chemistry*, vol. 31, no. 2, pp. 455-461, 2010. DOI: 10.1002/jcc.21334.
- S. Forli et al., "Computational protein-ligand docking and virtual drug screening with the AutoDock suite," *Nature Protocols*, vol. 11, pp. 905-919, 2016. DOI: 10.1038/nprot.2016.051.
- Daina, O. Michielin, and V. Zoete, "SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules," *Scientific Reports*, vol. 7, 42717, 2017. DOI: 10.1038/srep42717.
- D. E. V. Pires, T. L. Blundell, and D. B. Ascher, "pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures," *Journal of Medicinal Chemistry*, vol. 58, no. 9, pp. 4066-4072, 2015. DOI: 10.1021/acs.jmedchem.5b00104.

M. J. Abraham et al., "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," *SoftwareX*, vols. 1-2, pp. 19-25, 2015. DOI: 10.1016/j.softx.2015.06.001.

K. Vanommeslaeghe et al., "CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields," *Journal of Computational Chemistry*, vol. 31, no. 4, pp. 671-690, 2010. DOI: 10.1002/jcc.21367.

S. A. Hollingsworth and R. O. Dror, "Molecular dynamics simulation for all," *Neuron*, vol. 99, no. 6, pp. 1129-1143, 2018. DOI: 10.1016/j.neuron.2018.08.011.

S. Genheden and U. Ryde, "The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities," *Expert Opinion on Drug Discovery*, vol. 10, no. 5, pp. 449-461, 2015. DOI: 10.1517/17460441.2015.1032936.

J. Eberhardt, D. Santos-Martins, A. F. Tillack, and S. Forli, "AutoDock Vina 1.2.0: New docking methods, expanded force field, and Python bindings," *Journal of Chemical Information and Modeling*, vol. 61, no. 8, pp. 3891-3898, 2021. DOI: 10.1021/acs.jcim.1c00203.

E. Wang et al., "End-point binding free energy calculation with MM/PBSA and MM/GBSA: Strategies and applications in drug design," *Chemical Reviews*, vol. 119, no. 16, pp. 9478-9508, 2019. DOI: 10.1021/acs.chemrev.9b00055.