AJPR
Avicenna Journal of Pharmaceutical Research
Avicenna J Pharm Res 2023;4(2)62-73. doi:10.34172/ajpr.1129
http://ajpr.umsha.ac.ir

Original Article

Molecular Docking using MK-2048 and its Structural Analogs Against the HIV-1 Integrase Catalytic Core (PDB ID: 6NUJ)

Roohallah Yousefi[1,2*]

[1]Behbahan Faculty of Medical Sciences, Behbahan, Iran
[2]Department of Biochemistry, Faculty of Biology, Tarbit Modares University, Tehran, Iran

Article history:
Received: June 8, 2023
Accepted: July 30, 2023
ePublished: September 23, 2023

***Corresponding author:**
Roohallah Yousefi,
Email: ry@behums.ac.ir

Abstract

Background: This study investigated the molecular docking of MK-2048, a second-generation integrase inhibitor, and compound CID 76212788 against the HIV-1 integrase catalytic core (PDB ID: 6NUJ). The main objective of the study was to comprehend the structural basis of integrase inhibition and pinpoint potential new antiviral agents.

Methods: The model of the HIV-1 integrase catalytic core was retrieved from the RCSB PDB database, and the ligands were prepared using Molegro Virtual Docker. The SwissADME tool was utilized to forecast physicochemical properties and pharmacokinetics, while molecular docking was executed using Molegro Virtual Docker.

Results: Compound CID 76212788 displayed the strongest binding affinity, establishing hydrogen bonds with essential amino acids of the HIV-1 integrase catalytic core model. These interactions are crucial for designing potent selective inhibitors. The study also emphasized the importance of specific amino acid residues in drug resistance and natural variations among HIV-1 subtypes. All compounds exhibited high gastrointestinal absorption but were unlikely to penetrate the blood-brain barrier. Compound CID 136340365 was identified as a CYP2C19 inhibitor, while CID 54698642, 56841816, 76419031, 68116543, and 136340365 inhibited CYP2C9. Additionally, CID 136340365 and others inhibited CYP2D6. All compounds had low human skin permeability coefficients, suggesting limited dermal absorption.

Conclusion: The obtained data underscore the intricacy of HIV genetic diversity and the necessity of developing effective broad-spectrum integrase inhibitors. Further in vitro and in vivo studies are necessary to validate these findings and enhance the compounds for clinical use in the treatment and prevention of HIV infection.

Keywords: MK-2048, HIV-1 integrase, 6NUJ, Molecular docking

Please cite this article as follows: Yousefi R. molecular docking using MK-2048 and its structural analogs against the HIV-1 integrase catalytic core (PDB ID: 6NUJ). Avicenna J Pharm Res 2023;4(2)62-73. doi:10.34172/ajpr.1129

Introduction

The HIV-1 integrase is a crucial enzyme for the life cycle of the virus, especially in integrating viral DNA into the host DNA. Approximately 62.5% of the integrase residues are conserved, which is essential for the stability and function of the enzyme. The highly conserved regions of the HIV-1 integrase are identified by specific amino acid sequences. Six main sequences are highlighted: (W61-K71), (V75-Y83), (K127-A133), (Q137-V150), (L241-Q252), and (I257-K264). These conserved regions can aid in developing new inhibitors with a lower risk of resistance (1).

The catalytic triad (D64, D116, E152) and zinc-binding motif (H12, H16, C40, C43) are especially conserved and important for the activity and stability of the enzyme. Limited variability in these areas reflects strong selective pressures to preserve their essential roles (1,2).

Region I spans from W61 to K71 and includes the first amino acid of the catalytic triad, D64. Region II, from V75 to Y83, contains D83, which plays a crucial role in the key chemical reactions for integration. Region III (K127-A133) involves important lysine residues for DNA binding, while Region IV (Q137-V150) is located near E152, another vital component of the catalytic triad (3,4).

Lastly, the non-canonical nuclear localization signal (I161-K173) is crucial for importing the pre-integration complex into the host nucleus, necessary for effective viral infection. Variants in nearby residues have been linked to resistance to integrase inhibitors and can impact viral replication. The conservation of these regions across various HIV-1 subtypes also highlights their functional importance (4).

The study conducted by Drelich et al showed that mutations in certain key residues can impact the function of the HIV-1 integrase, which is vital for the life cycle of the virus. Important regions of the HIV-1 integrase include the C-terminal domain and conserved areas V and VI, which are critical for integrating viral DNA into the host genome. Area V and area VI are almost entirely conserved among retroviruses, underscoring their significance in viral integration (4,5).

The C region of the integrase is fully conserved and contains a key motif essential for HIV-1 replication. Specific residues within area VI are crucial for both integrase multimerization and DNA binding. For example, V260 is essential for multimerization, while K258, R262, R263, and K264 are involved in binding DNA. Mutations in these residues, including INV260E and a triple mutant, can significantly reduce viral replication, causing undetectable viral growth in some studies (1,5).

In contrast, the sequence Q, while still maintaining a basic positive charge, shows more variability among lentiviruses. Despite this variability, it retains essential properties for viral function. Additionally, region IV includes a flexible loop that plays a critical role after DNA binding, where mutations can affect integrase activity without altering DNA binding. Overall, the C-terminal domain of the virus has an SH3 fold that is vital for DNA binding, with specific conserved residues contributing to the structural integrity and function of the integrase enzyme (1,5).

Key conserved residues that help integrase connect with the cellular cofactor LEDGF/p75 include H12, L102, A128, C130, W131, and E170. Any mutations that affect these interactions could harm the virus, which is why such changes are rare. For example, mutations at H171 are infrequent, and the H171Y variant appears in less than 1% of analyzed sequences, indicating a strong preference to maintain this residue due to its role in binding to LEDGF/p75 (3). Mutations such as H12N alter integrase activity, affecting viral integration and reducing infectivity. The K34A mutant can act catalytically but struggles with effective replication. Conversely, certain amino acids like S17, I72, T124, T125, V201, S119, and L101 show a mutation frequency of over 25% (1,6).

Integrase inhibitors such as raltegravir and elvitegravir have significantly improved antiretroviral treatment. Variants like L74M, T97A, E157Q, and others are more common in integrase of specific HIV subtypes, which may impact how resistance develops (7). The main mutations associated with resistance to raltegravir (Y143R, Q148H/K/R, N155H) and elvitegravir (T66I, E92Q, S147G, Q148H/K/R, N155H) are rare or not found in patients who have not received integrase inhibitors, indicating that these mutations do not occur naturally (8). Secondary mutations that support resistance alongside the main ones are also uncommon, while others like L74M, T97A, and E157Q appear as natural variations in different HIV-1 subtypes, highlighting the importance of considering specific genetic backgrounds in drug resistance evaluations (9).

MK-2048

MK-2048 is a molecule created by Merck & Co. as a second-generation integrase inhibitor for treating and preventing HIV infection. It was developed after raltegravir, the first integrase inhibitor approved in 2007. MK-2048 has the chemical formula C21H21ClFN5O4 and a half-life of 32 hours, which may lead to longer effects and increased resistance against HIV. It targets HIV strains that have become resistant to earlier drugs (10,11).

The molecule works by disrupting the interaction between viral integrase and DNA during integration, thus preventing HIV from replicating. In laboratory tests, MK-2048 showed strong activity against both wild-type HIV and drug-resistant variants, with a low IC50 value of 42 nMol indicating effectiveness. It has been tested on different cell types, such as human macrophages and T cells, showing similar or improved activity in macrophages (10-12).

Additionally, MK-2048 has been included in intravaginal rings (IVRs) for HIV prevention, allowing drug release for up to 28 days, providing a discreet option for at-risk individuals. This development is expected to enhance HIV therapy, especially against resistant strains (12).

MK-2048 is a second-generation integrase inhibitor that significantly improves HIV treatment due to its effectiveness against drug-resistant virus types. In contrast to first-generation inhibitors, MK-2048 can handle drug resistance better and presents different resistance patterns. This includes resistance mutations such as G118R and E138K, which require careful monitoring and may need combination therapies for better results (13).

Allosteric integrase inhibitors (LEDGINs) show promise for combination therapies by targeting the integrase and related host proteins. Early studies suggest that using LEDGINs with first-generation inhibitors may enhance treatment effectiveness. Despite the success of MK-2048 in pre-clinical studies, a human trial from 2015 to 2016 did not find a clear link between its levels and its ability to inhibit HIV, raising doubts about its potential for pre-exposure prophylaxis (PrEP) (11-13).

Macrocyclic Integrase Inhibitors

A new series of macrocyclic pyrazinopyrrolopyridazine dione derivatives includes compound CID 76212788, which has the ability to inhibit HIV replication, particularly against drug-resistant strains. These compounds differ from known integrase inhibitors and show effective antiviral activity against mutant HIV strains with specific mutations that confer resistance to raltegravir. Additionally, they have favorable pharmacokinetic and pharmacodynamic properties, suggesting their potential use in combination therapies for treating HIV infection (14).

The compounds follow a general formula that includes various substitution patterns on the pyrazinopyrrolopyridazine core, which may enhance their biological activity. There are a variety of compounds with different substituents that allow for optimization of their

action and pharmacological properties. These compounds can exist in different forms, such as isomers, salts, and solvates. Their antiviral activity was tested in cell-based assays using infected cells with various HIV-1 strains, and the effectiveness of the compounds was measured by EC50 values (14, 15).

In this study, the molecular docking of the drug compound MK-2048 and other compounds with similar structures, including compound CID 76212788, was investigated. Then, the physicochemical and pharmacological properties of these compounds were examined (Figure 1).

Materials and Methods

Preparation of the HIV Integrase Catalytic Core Model [PDB ID: 6NUJ]

The molecular model of the HIV-1 integrase catalytic core [PDB ID: 6NUJ] was obtained from the RCSB PDB database. The 6NUJ structure represents the HIV integrase catalytic core complexed with an allosteric inhibitor, BI-224436, crystallized using X-ray diffraction. The structure was determined using X-ray diffraction with a resolution of 2.10 Å (13).

Ligand Model Preparation

The 3D model of compound MK-2048 and other compounds with similar structures, including compound CID 76212788, downloaded from the PubChem database, were prepared using Molegro Virtual Docker version 5.0. The manufacturer's proprietary molecule preparation algorithm was used to prepare the ligand and protein (16,17).

Predicting Physicochemical Properties and Pharmacokinetics

The SwissADME web tool is utilized for predicting the physicochemical properties and pharmacokinetics of compounds. It offers a range of predictive models, such as BOILED-Egg and iLOGP, for evaluating drug development parameters. The platform is user-friendly and accessible at http://www.swissadme.ch. This tool is renowned for its fast and accurate predictions (18-20).

Molecular Docking

Molegro Virtual Docker (MVD), a versatile and reliable software for molecular docking, was used in this study. This powerful tool predicts the binding affinity and orientation of small molecules to target proteins using advanced computational methods. MVD stands out for its flexible docking algorithm, which considers multiple conformations of both ligand and receptor molecules, ensuring accurate results in both rigid and flexible docking scenarios. Some notable features of MVD include grid-based scoring functions, support for multiple docking engines, and comprehensive output analysis tools, aiding in further exploration of docking results (21).

Within the Docking Wizard of the Molegro software, we had the option to choose from various scoring algorithms for docking. We opted for the MolDock Scoring Grid algorithm with a grid resolution of 0.3 Å. Additionally, we selected the MolDock SE algorithm for ligand energy optimization and minimization, conducting 10 iterations of approximately 1500 runs each. In defining the grid for the binding site on the protein target, the Docking Wizard of the Molegro software allowed us to set grid dimensions. The search space is defined by a position (x,y,z) and a radius. We established search space dimensions along the X axis at -21.38, the Y axis at 27.90, and the Z axis at -15.72 with a radius of 26. We decided to use the entire protein molecule as the binding site, as we found it to be the most effective binding site within the protein target (21-23).

The GUI of Molegro Visual Docker version 5.0 allows users to visualize how ligands and receptors interact by displaying energy contributions from different atom pairs. A key feature enables users to switch between displaying a table and syncing atom selections, and between the table and the 3D view, which enhances comprehension

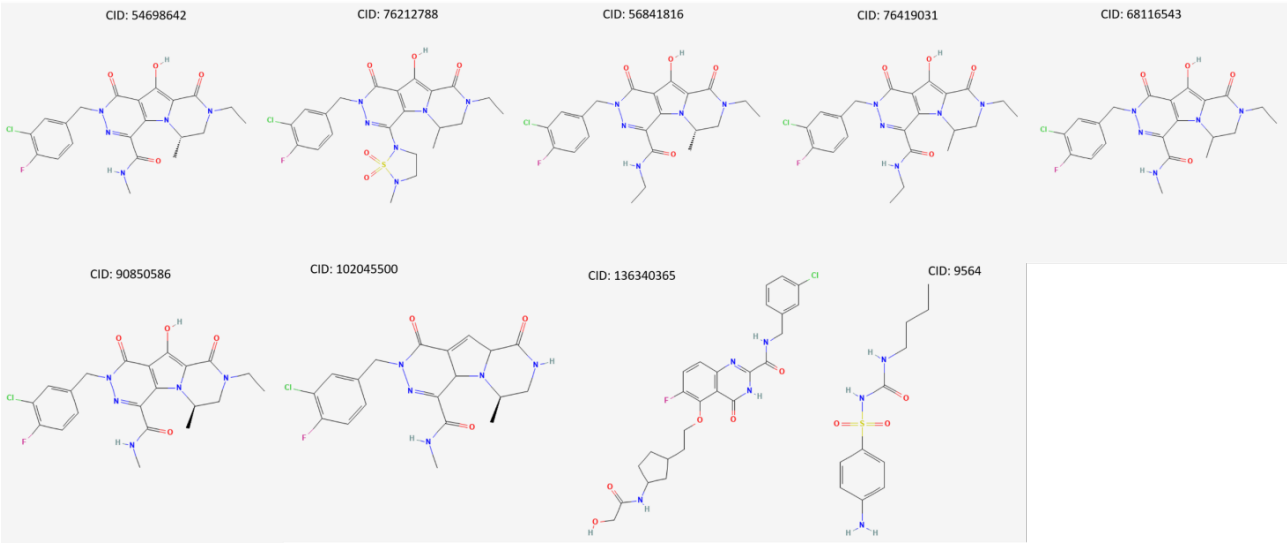

Figure 1. 2D Structures of the Compounds under Study

of complex molecular interactions. There are three main types of energy contributions. EPair represents the energy between a ligand atom and a receptor atom, including steric interactions (repulsions when atoms are too close) and hydrogen bonding (an attractive interaction between a hydrogen atom and nearby electronegative atoms). EIntra represents the internal energy within the ligand, encompassing bonded interactions among ligand atoms and non-bonded interactions between non-bonded ligand atoms. This is crucial for understanding ligand stability and flexibility when bound. EElec includes electrostatic interactions, referring to attractions and repulsions between charged particles. The GUI distinguishes between short-range (less than 4.5 Å) and long-range (greater than 4.5 Å) interactions. Knowledge of these interactions is essential for evaluating the stability of the ligand-receptor complex (21-23).

Results

The Molecular Docking Results

In this study, the entire catalytic core of integrase was examined. Additionally, we explored the regions of the enzyme to which each ligand showed the highest specificity. The results indicated that the studied compounds often target the hydrophobic pocket in the enzyme (Figures 2, 3, 4, and 5).

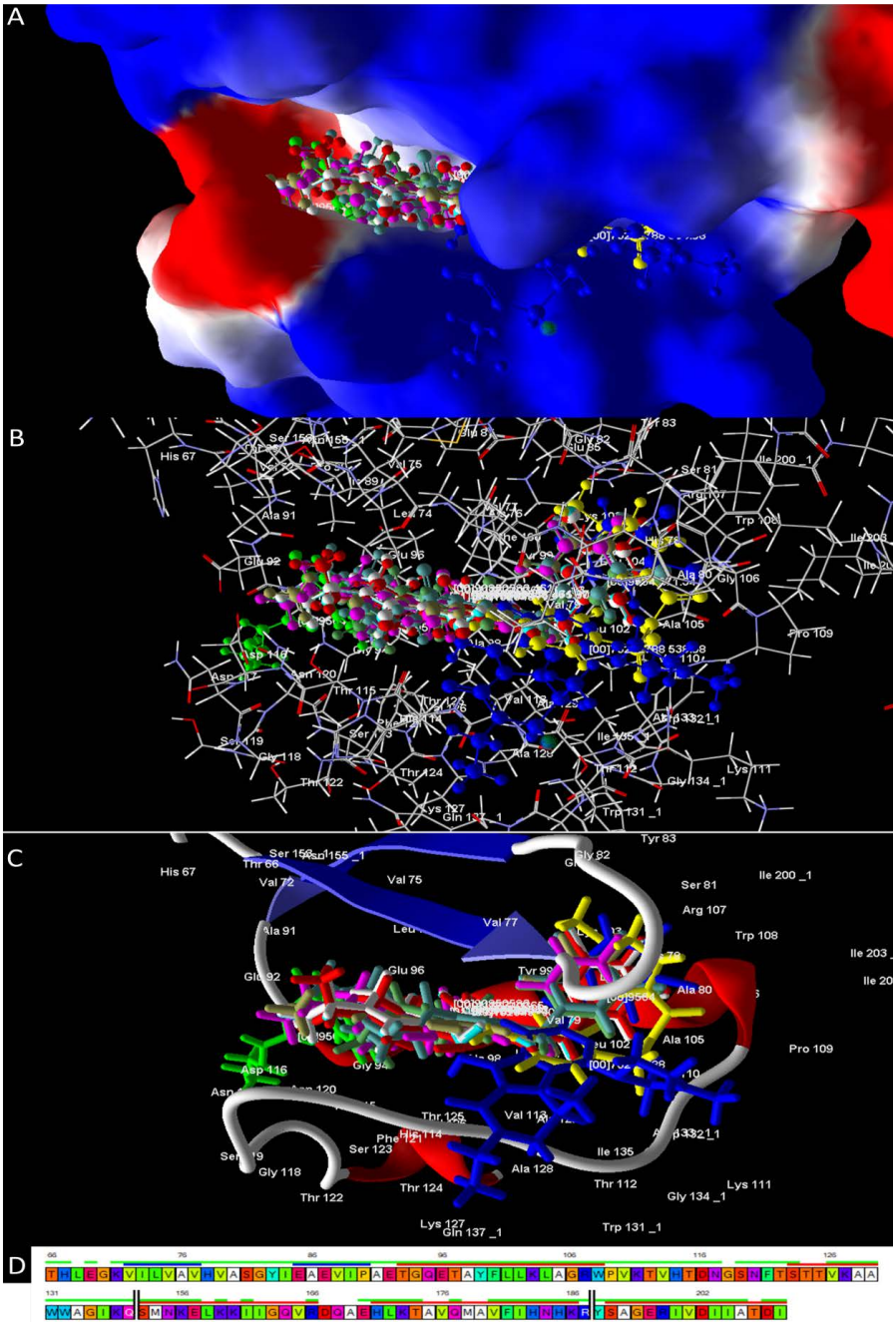

Figure 2. A: The ligand binding site on the integrase molecular model displays the electrostatic surfaces of the integrase catalytic core binding to the studied ligands. B: In the same location, the integrase molecular model is depicted as a wireframe, and the ligand model is shown as a ball-and-stick. C: At the same location, the integrase molecular model is shown as a secondary structure, while the ligand model is represented as a stick. D: The sequence of the catalytic core of the integrase enzyme model is presented.

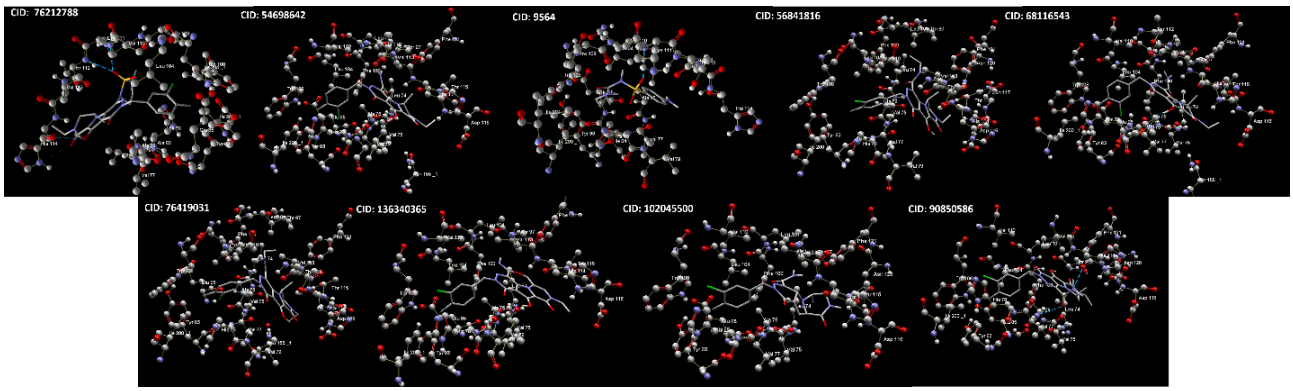

Figure 3. The Illustration of the Binding Site of the Studied HIV-1 Integrase Inhibitor Compounds on the Enzyme. Hydrogen bonds are indicated by blue dotted lines. The enzyme model is displayed as a ball-and-stick model.

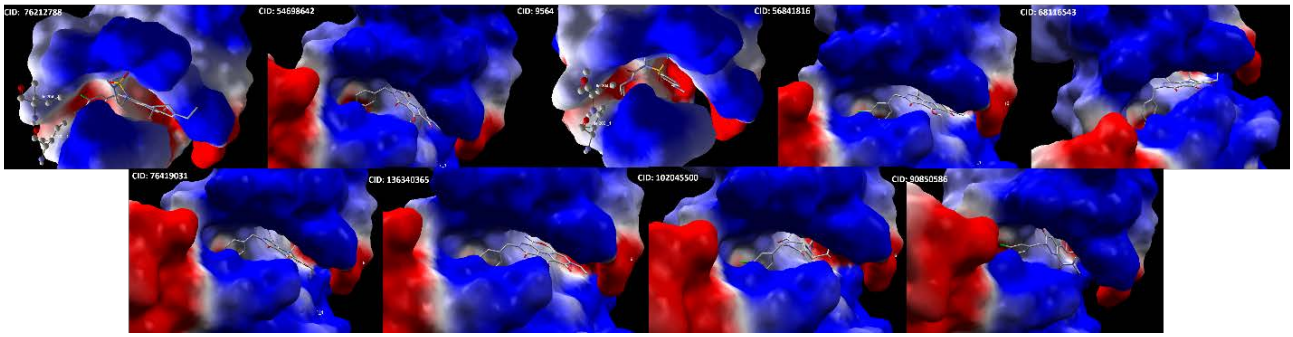

Figure 4. The Illustration of the Binding Site of the Studied HIV-1 Integrase Inhibitor Compounds on the Enzyme. The enzyme model is shown as electrostatic surfaces.

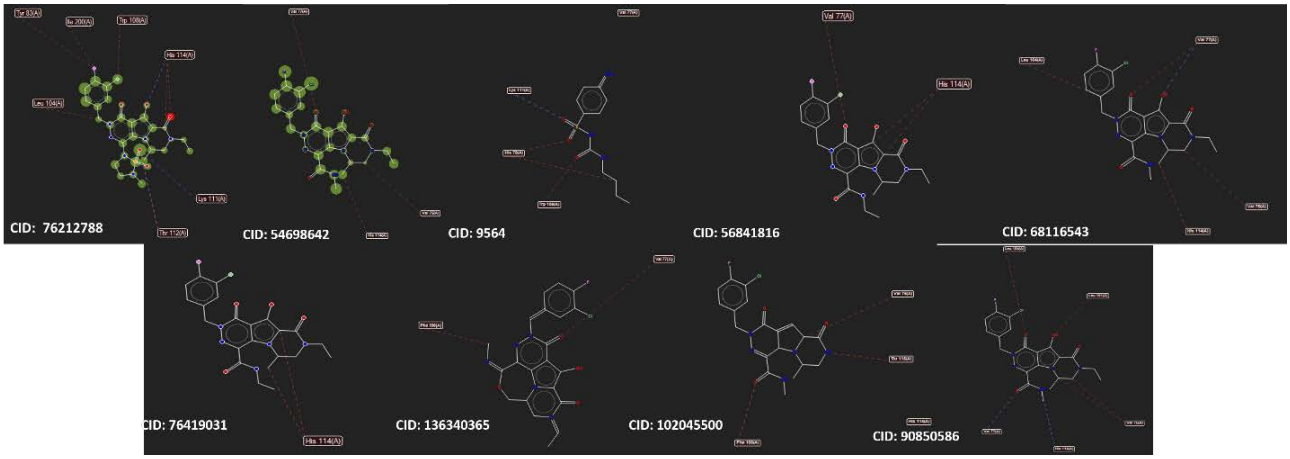

Figure 5. A 2D Map Illustrating the Interactions between Binding Site Residues and the Studied Ligands

The binding site comprises Leu74, Val75, Ala76, Val77, His78, Val79, Ala80, Tyr83, Ile84, Glu85, Phe100, Trp108, Pro109, Val110, Lys111, Thr112, Val113, His114, Thr115, Asp116, Phe121, Asn155, Ile200, and Ile204 for all compounds under study (Table 1).

Among the compounds analyzed, compound CID 76212788 exhibited the highest binding affinity for the integrase binding site. This compound binds to the amino acids in the following order based on the most negative binding energy: Ala76, Val77, His78, Val79, Ala80, Tyr83, Glu85, Leu104, Trp108, Pro109, Val110, Lys111, Thr112, Val113, His114, Ile200, and Ile204 (Table 2).

Physicochemical Properties of the Studied Compounds

Among the compounds studied, compound CID 76212788 has the highest Fraction Csp3 value at 0.41. This compound also has the most hydrogen bond acceptors, with a total of 8. Additionally, it exhibits the highest molar refractivity (142.92). Compound CID 136340365 has the highest topological polar surface area (133.41 Å2), followed by compound CID 76212788 with a topological polar surface area of 129.36 Å2. The highest Ilogp value is found in the combination of compounds CID 56841816 and CID 76419031 (3.17). Most of the studied compounds, including CID 54698642 and CID 76212788, demonstrate

Table 1. Molecular Docking Results of the Studied Compounds on the HIV Integrase Catalytic Core Model [PDB ID: 6NUJ]

Ligand	MolDock score (kcal/mol)	Rerank score (kcal/mol)	Interaction (kcal/mol)	Protein (kcal/mol)	Internal (kcal/mol)	Torsions	HBond	Heavy atoms	MW	(Ligand efficiency) LE1 (MolDock score/heavy atoms)	Docking score (kcal/mol)
9564	-106.834	-82.2502	-117.747	-117.747	10.9123	7	-2.43554	18	271.336	-5.93524	-106.853
54698642	-155.995	-122.318	-167.687	-167.687	11.6923	4	-0.53109	32	461.874	-4.87483	-153.556
56841816	-162.677	-122.63	-170.861	-170.861	8.18422	5	0	33	475.901	-4.92961	-159.85
68116543	-156.409	-121.689	-167.086	-167.086	10.6773	4	-0.82587	32	461.874	-4.88777	-154.062
76212788	-174.001	-117.487	-171.438	-171.438	-2.56272	4	-6.70997	36	538.98	-4.83335	-173.618
76419031	-161.076	-127.194	-174.057	-174.057	12.981	5	0	33	475.901	-4.88111	-158.161
90850586	-155.979	-112.618	-165.681	-165.681	9.70206	4	-4.23768	32	461.874	-4.87434	-153.859
102045500	-142.583	-112.176	-154.142	-154.142	11.5594	3	0	29	419.837	-4.91664	-144.389
136340365	-139.471	-113.228	-164.833	-164.833	25.3619	1	0	32	455.826	-4.35848	-137.843

moderate solubility in aqueous media (Table 3, Figure 6).

Pharmaceutical Properties of Studied Compounds

All studied compounds have high gastrointestinal absorption. None of the studied compounds are BBB permeant. All compounds except for compound CID 9564 are Pgp substrates. None of the studied compounds are CYP1A2 inhibitor. Only compound CID 136340365 is a CYP2C19 inhibitor. Compounds CID 54698642, CID 56841816, CID 76419031, CID 68116543, CID 90850586, and CID 136340365 are CYP2C9 inhibitors. Compound CID 136340365 is also a CYP2D6 inhibitor. Compounds CID 56841816, CID 76419031, and CID 136340365 are inhibitors of the CYP3A4 enzyme. Compound CID 76212788 has the lowest dermal absorption, with a log Kp (cm/s) of -8.24. This compound also has the lowest bioavailability score (0.17). The compound with the highest Synthetic Accessibility score in the study is CID 102045500, with a score of 4.88 (Table 4, Figure 7).

Discussion

The Molecular Docking Study

The high conservation of the integrase enzyme underscores its essential role in the viral life cycle, making it an attractive target for antiviral therapies. The identification of compound CID 76212788 as the ligand with the highest binding affinity to the integrase binding site is particularly noteworthy. The detailed interaction analysis reveals that this compound forms several hydrogen bonds with key amino acid residues such as His 78, Val 77, Val 79, Thr 112, Val 113, His 114, Lys 111, Val 110, Leu 104, Trp 108, Ala 76, Pro 109, Ile 200, Ile 204, Glu 85, and Ala 80. These interactions are crucial for the design of potent and selective inhibitors.

The utilization of molecular docking (Figures 2-5) and the calculation of binding energies (Tables 1 and 2) offer a robust framework for understanding the structural basis of integrase inhibition. The ligand efficiency analysis assesses how well the ligands interact with the enzyme relative to their size, which is essential for minimizing off-target effects and enhancing drug potency.

Moreover, the emphasis on the significance of specific amino acid residues in binding sites and their roles in drug resistance and natural variations among HIV-1 subtypes underscores the complexity of HIV genetic diversity and its implications for antiretroviral therapy. Knowledge of these variations can guide the development of more effective and broad-spectrum integrase inhibitors (1-9).

Physicochemical Properties of the Studied Compounds

Fraction Csp3 measures the sp3 hybridization of carbon atoms in a molecule, indicating lipophilicity and hydrophobic interaction potential. A higher fraction Csp3 suggests greater lipophilicity. In the study, the molecule CID: 76212788 has the highest fraction Csp3 (0.41), indicating better cell membrane permeability and adaptability for binding to the active site of integrase enzyme (19,22).

The moderate Log P values of the studied compounds, ranging from 1.05 to 3.09, suggest a suitable balance between hydrophobicity and hydrophilicity, which is essential for good absorption and distribution characteristics. The octanol-water partition coefficient (Log P) is another essential parameter for predicting the distribution of a drug between lipid and aqueous phases. The Log P values calculated by different algorithms (iLOGP, XLOGP3, and Consensus Log P) provide an estimate of the hydrophobicity of a molecule. A higher Log P value signifies higher hydrophobicity and a greater tendency to accumulate in fatty tissues. Compounds CID 136340365, CID 56841816, and CID 76419031 exhibit the highest Log P values, which might imply better absorption and distribution in organisms due to their affinity for fatty tissues (22,23).

The topological polar surface area (TPSA) is a measure of the hydrophilic character of the molecule, which is important for determining the ability of the compound to interact with polar environments, such as water. Compounds with higher TPSA values are typically more soluble in water and may have greater bioavailability due to their capacity to form hydrogen bonds with water molecules. Here, compound CID 136340365 has the highest TPSA value, indicating that it may be more water-soluble and potentially more bioavailable compared to

Table 2. Molecular Docking Results of the Ligand CID 76212788 on the HIV Integrase Catalytic Core Model (PDB ID: 6NUJ)

	Calculation of the binding energy for each atom of amino acid in the binding site of ligand with CID number 76212788					Calculation of the docking energy for each atom of ligand in the target protein with CID Number 76212788				
Target Atoms: Molecule	**Residue**	**ID**	**Atom**	**Total**	**E_{Pair}**	**Atom ID**	**Name**	**Total**	**E_{Pair}**	**E_{Intra}**
6NUJ [A]	Ala 76	78	CA	-1.16841	-1.16841	0	CL	-7.50222	-6.73774	-0.76449
6NUJ [A]	Ala 76	79	C	-1.53301	-1.53301	1	S	-7.68008	-6.6325	-1.04758
6NUJ [A]	Ala 76	80	O	-0.80538	-0.80538	2	F	-0.30061	-0.66854	0.367932
6NUJ [A]	Ala 76	81	CB	-2.15267	-2.15267	3	O	-3.58937	-5.23744	1.64808
6NUJ [A]	Ala 80	106	N	-0.54928	-0.54928	4	O	-3.94944	-4.7323	0.782867
6NUJ [A]	Glu 85	141	N	-0.31087	-0.31087	5	O	-11.412	-9.86324	-1.54873
6NUJ [A]	Glu 85	146	CG	-0.77547	-0.77547	6	O	4.59051	2.52333	2.06718
6NUJ [A]	Glu 85	147	CD	-0.53565	-0.53565	7	O	-3.04147	-4.44335	1.40189
6NUJ [A]	Glu 85	149	OE2	-0.76032	-0.76032	8	N	-3.9079	-3.48419	-0.42371
6NUJ [A]	His 114	368	N	-3.1015	-3.1015	9	N	-4.45569	-2.83168	-1.62401
6NUJ [A]	His 114	369	CA	-3.41688	-3.41688	10	N	-2.18899	-3.19418	1.00519
6NUJ [A]	His 114	370	C	-1.87093	-1.87093	11	N	-4.41015	-3.25253	-1.15762
6NUJ [A]	His 114	371	O	-3.44609	-3.44609	12	N	-4.82798	-4.59984	-0.22814
6NUJ [A]	His 114	372	CB	4.66463	4.66463	13	N	-0.30314	-4.45489	4.15175
6NUJ [A]	His 114	373	CG	-1.81248	-1.81248	14	C	-4.98609	-3.8712	-1.11489
6NUJ [A]	His 114	374	ND1	-1.57599	-1.57599	15	C	-5.64323	-4.69134	-0.95189
6NUJ [A]	His 114	375	CD2	-1.46619	-1.46619	16	C	-3.1417	-4.14119	0.999492
6NUJ [A]	His 114	376	CE1	-0.79504	-0.79504	17	C	-5.56225	-5.49543	-0.06683
6NUJ [A]	His 114	377	NE2	-0.55771	-0.55771	18	C	-5.17793	-4.39584	-0.78209
6NUJ [A]	His 78	89	N	-3.24665	-3.24665	19	C	-6.55575	-5.71372	-0.84203
6NUJ [A]	His 78	90	CA	-5.45581	-5.45581	20	C	-5.78886	-3.87213	-1.91673
6NUJ [A]	His 78	91	C	-1.41983	-1.41983	21	C	-6.03929	-6.0302	-0.00909
6NUJ [A]	His 78	93	CB	-3.67468	-3.67468	22	C	-3.974	-2.30019	-1.67381
6NUJ [A]	His 78	94	CG	-4.51847	-4.51847	23	C	-1.8052	-2.60379	0.798591
6NUJ [A]	His 78	95	ND1	-1.87031	-1.87031	24	C	-4.07923	-2.28393	-1.7953
6NUJ [A]	His 78	96	CD2	-4.74513	-4.74513	25	C	-5.3788	-5.23962	-0.13918
6NUJ [A]	His 78	97	CE1	-1.75572	-1.75572	26	C	-2.67657	-3.02319	0.346619
6NUJ [A]	His 78	98	NE2	-4.07207	-4.07207	27	C	-5.02236	-4.45112	-0.57124
6NUJ [A]	Leu 104	293	CB	-0.41797	-0.41797	28	C	-3.29734	-2.45925	-0.83809
6NUJ [A]	Leu 104	294	CG	-2.4974	-2.4974	29	C	-3.67627	-3.60237	-0.07389
6NUJ [A]	Leu 104	295	CD1	-2.85602	-2.85602	30	C	-5.09961	-6.18293	1.08332
6NUJ [A]	Leu 104	296	CD2	-0.55184	-0.55184	31	C	-4.52606	-6.7706	2.24454
6NUJ [A]	Lys 111	345	N	-3.34683	-3.34683	32	C	-8.49959	-8.66476	0.165164
6NUJ [A]	Lys 111	346	CA	-1.059	-1.059	33	C	-7.59253	-7.75732	0.164792
6NUJ [A]	Lys 111	347	C	-0.71342	-0.71342	34	C	-10.5527	-11.2694	0.716714
6NUJ [A]	Lys 111	348	O	-0.3097	-0.3097	35	C	-8.14969	-9.11449	0.964801
6NUJ [A]	Lys 111	349	CB	-1.44016	-1.44016					
6NUJ [A]	Lys 111	350	CG	-1.15187	-1.15187					
6NUJ [A]	Lys 111	351	CD	-0.84611	-0.84611					
6NUJ [A]	Pro 109	331	N	-0.61958	-0.61958					
6NUJ [A]	Pro 109	332	CA	-0.47237	-0.47237					
6NUJ [A]	Pro 109	333	C	-2.0128	-2.0128					
6NUJ [A]	Pro 109	334	O	-1.97859	-1.97859					
6NUJ [A]	Pro 109	336	CG	-0.49785	-0.49785					
6NUJ [A]	Pro 109	337	CD	-0.53143	-0.53143					
6NUJ [A]	Thr 112	354	N	-3.42045	-3.42045					
6NUJ [A]	Thr 112	355	CA	-1.53994	-1.53994					
6NUJ [A]	Thr 112	356	C	-4.55706	-4.55706					
6NUJ [A]	Thr 112	357	O	-4.94931	-4.94931					
6NUJ [A]	Thr 112	358	CB	-1.5176	-1.5176					

Table 2. Continued.

	Calculation of the binding energy for each atom of amino acid in the binding site of ligand with CID number 76212788					Calculation of the docking energy for each atom of ligand in the target protein with CID Number 76212788				
Target Atoms: Molecule	**Residue**	**ID**	**Atom**	**Total**	**E_{Pair}**	**Atom ID**	**Name**	**Total**	**E_{Pair}**	**E_{Intra}**
6NUJ [A]	Thr 112	359	OG1	-2.26321	-2.26321					
6NUJ [A]	Trp 108	318	CA	-0.84976	-0.84976					
6NUJ [A]	Trp 108	319	C	-0.5682	-0.5682					
6NUJ [A]	Trp 108	321	CB	-0.49603	-0.49603					
6NUJ [A]	Trp 108	322	CG	-2.31122	-2.31122					
6NUJ [A]	Trp 108	323	CD1	-2.84588	-2.84588					
6NUJ [A]	Trp 108	324	CD2	-1.71027	-1.71027					
6NUJ [A]	Trp 108	325	NE1	-2.10189	-2.10189					
6NUJ [A]	Trp 108	326	CE2	-1.45093	-1.45093					
6NUJ [A]	Trp 108	327	CE3	-1.36917	-1.36917					
6NUJ [A]	Trp 108	328	CZ2	-0.86367	-0.86367					
6NUJ [A]	Trp 108	329	CZ3	-0.73771	-0.73771					
6NUJ [A]	Trp 108	330	CH2	-0.52356	-0.52356					
6NUJ [A]	Tyr 83	123	C	-0.45551	-0.45551					
6NUJ [A]	Tyr 83	124	O	-1.07684	-1.07684					
6NUJ [A]	Tyr 83	125	CB	-0.76248	-0.76248					
6NUJ [A]	Tyr 83	126	CG	-1.02707	-1.02707					
6NUJ [A]	Tyr 83	127	CD1	0.305094	0.305094					
6NUJ [A]	Tyr 83	128	CD2	-0.30977	-0.30977					
6NUJ [A]	Tyr 83	129	CE1	-1.26324	-1.26324					
6NUJ [A]	Tyr 83	131	CZ	-0.85709	-0.85709					
6NUJ [A]	Tyr 83	132	OH	-0.33164	-0.33164					
6NUJ [A]	Val 110	338	N	-0.73907	-0.73907					
6NUJ [A]	Val 110	339	CA	-2.10825	-2.10825					
6NUJ [A]	Val 110	340	C	-1.50641	-1.50641					
6NUJ [A]	Val 110	341	O	-0.62195	-0.62195					
6NUJ [A]	Val 110	342	CB	-2.12557	-2.12557					
6NUJ [A]	Val 110	343	CG1	-3.12316	-3.12316					
6NUJ [A]	Val 110	344	CG2	-3.19958	-3.19958					
6NUJ [A]	Val 113	361	N	-1.9077	-1.9077					
6NUJ [A]	Val 113	362	CA	-4.5017	-4.5017					
6NUJ [A]	Val 113	363	C	-3.79706	-3.79706					
6NUJ [A]	Val 113	364	O	-0.84827	-0.84827					
6NUJ [A]	Val 113	365	CB	-1.81827	-1.81827					
6NUJ [A]	Val 113	366	CG1	-2.03948	-2.03948					
6NUJ [A]	Val 113	367	CG2	-2.39565	-2.39565					
6NUJ [A]	Val 77	82	N	-1.58128	-1.58128					
6NUJ [A]	Val 77	83	CA	-1.57818	-1.57818					
6NUJ [A]	Val 77	84	C	-4.33809	-4.33809					
6NUJ [A]	Val 77	85	O	-4.98418	-4.98418					
6NUJ [A]	Val 79	99	N	-2.48439	-2.48439					
6NUJ [A]	Val 79	100	CA	-1.18984	-1.18984					
6NUJ [A]	Val 79	103	CB	-3.55865	-3.55865					
6NUJ [A]	Val 79	104	CG1	-1.45217	-1.45217					
6NUJ [A]	Val 79	105	CG2	-4.96966	-4.96966					
6NUJ [A]_1	Ile 200	400	CB	-0.4225	-0.4225					
6NUJ [A]_1	Ile 200	401	CG1	-1.4014	-1.4014					
6NUJ [A]_1	Ile 200	403	CD1	5.76477	5.76477					
6NUJ [A]_1	Ile 204	432	CG1	-0.31433	-0.31433					
6NUJ [A]_1	Ile 204	434	CD1	-0.7942	-0.7942					

Table 3. Physicochemical Properties of the Studied Compounds

Molecule	Formula	MW	Heavy atoms	Aromatic heavy atoms	Fraction Csp3	Rotatable bonds	H-bond acceptors	H-bond donors	MR (molar refractivity)	TPSA (topological polar surface area)	iLOGP	XLOGP3	Consensus Log P	ESOL Log S	ESOL solubility (mg/mL)	ESOL solubility (mol/L)	ESOL Class
54698642	C21H21ClFN5O4	461.87	32	15	0.33	5	6	2	120.3	109.46	3.02	2.35	2.33	-4.2	0.029100	0.000063	Moderately soluble
76212788	C22H24ClFN6O5S	538.98	36	15	0.41	4	8	1	142.92	129.36	2.98	1.9	2.06	-4.42	0.020400	0.000038	Moderately soluble
56841816	C22H23ClFN5O4	475.9	33	15	0.36	6	6	2	125.11	109.46	3.17	2.71	2.63	-4.44	0.017300	0.000037	Moderately soluble
76419031	C22H23ClFN5O4	475.9	33	15	0.36	6	6	2	125.11	109.46	3.17	2.71	2.63	-4.44	0.017300	0.000037	Moderately soluble
68116543	C21H21ClFN5O4	461.87	32	15	0.33	5	6	2	120.3	109.46	3.02	2.35	2.33	-4.2	0.029100	0.000063	Moderately soluble
90850586	C21H21ClFN5O4	461.87	32	15	0.33	5	6	2	120.3	109.46	3.03	2.35	2.33	-4.2	0.029100	0.000063	Moderately soluble
102045500	C19H19ClFN5O3	419.84	29	6	0.37	4	6	2	118	94.11	2.45	1.12	1.05	-3.04	0.385000	0.000917	Soluble
136340365	C25H26ClFN4O5	516.95	36	16	0.36	11	7	4	131.81	133.41	3.12	2.78	3.09	-4.4	0.020600	0.000040	Moderately soluble
9564	C11H17N3O3S	271.34	18	6	0.36	7	3	3	69.36	109.67	1.06	1.01	0.98	-1.94	3.090000	0.011400	Very soluble

Figure 6. The Radar Scale of Physicochemical Properties and Bioavailability of the Compounds Studied

others in the study.

The molecule CID 76212788 has the highest number of hydrogen bond acceptors (8), which suggests a strong potential for hydrogen bond formation with the enzyme. This is a key factor in inhibitor potency. On the other hand, CID 10204500, soluble in physiological liquids but with fewer rotatable bonds, could still exhibit favorable bioavailability due to its rigid structure and ability to interact with biological targets. The number of rotatable bonds affects the flexibility of the molecule, which can influence its ability to interact with biological targets and pass through biological barriers. Fewer rotatable bonds generally lead to higher molecular rigidity and potentially better bioavailability. In this dataset, compound CID 136340365 has the highest number of rotatable bonds (4), making it very flexible (19,22,23).

The solubility values, shown as the Esol Solubility (mg/mL), indicate how well a compound dissolves in water. Higher solubility means better bioavailability, allowing compounds to reach biological targets more easily. CID 9564 is a very soluble compound with a solubility of 3.09 mg/mL, indicating excellent water solubility and bioavailability. The Esol bioavailability classification helps assess how easily a compound can dissolve in gastrointestinal fluids, affecting drug absorption (19,22,23).

Pharmaceutical Properties of Studied Compounds

All the compounds exhibit high gastrointestinal (GI) absorption, suggesting they are likely to be efficiently

Table 4. Pharmaceutical Properties of Studied Compounds

Molecule	GI absorption	BBB permeant	Pgp substrate	CYP1A2 inhibitor	CYP2C19 inhibitor	CYP2C9 inhibitor	CYP2D6 inhibitor	CYP3A4 inhibitor	log Kp (cm/s)	Bioavailability score	Synthetic accessibility
54698642	High	No	Yes	No	No	Yes	No	No	-7.45	0.55	3.93
76212788	High	No	Yes	No	No	No	No	No	-8.24	0.17	4.57
56841816	High	No	Yes	No	No	Yes	No	Yes	-7.28	0.55	4.04
76419031	High	No	Yes	No	No	Yes	No	Yes	-7.28	0.55	4.04
68116543	High	No	Yes	No	No	Yes	No	No	-7.45	0.55	3.93
90850586	High	No	Yes	No	No	Yes	No	No	-7.45	0.55	3.93
102045500	High	No	Yes	No	No	No	No	No	-8.07	0.55	4.88
136340365	High	No	Yes	No	Yes	Yes	Yes	Yes	-7.48	0.55	4.12
9564	High	No	No	No	No	No	No	No	-7.24	0.55	2.38

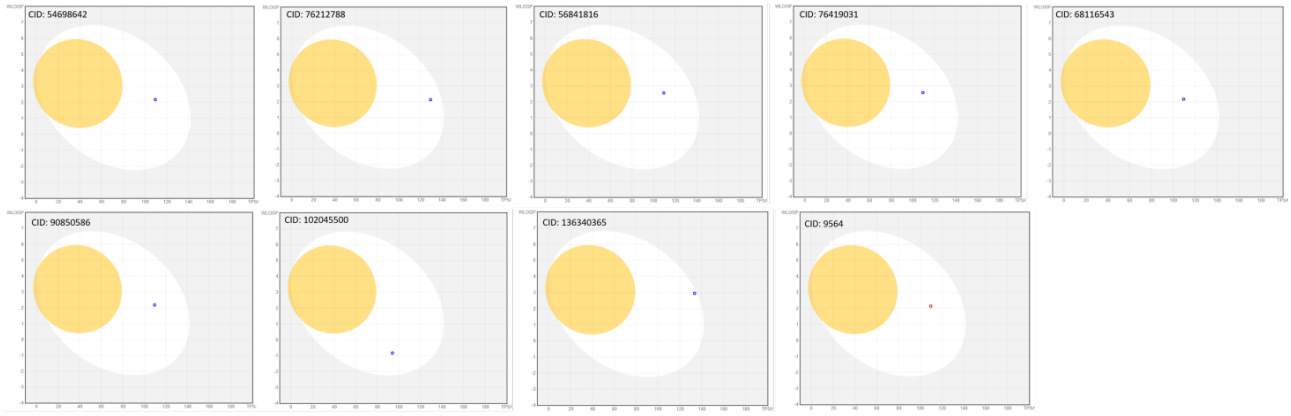

Figure 7. Illustration of the Interaction of the Compounds with Biological Barriers like the Blood-brain Barrier (BBB) and the Digestive System. The yellow area indicates the ability of the compound' to cross the BBB, while the white area represents its potential for absorption in the digestive system. Blue dots on the plot signify the compound entering the central nervous system (CNS) through P-glycoproteins, while red dots indicate its exit from the CNS

absorbed into the systemic circulation after oral administration. This is crucial for drugs intended for oral use. None of the studied compounds are able to cross the blood-brain barrier (BBB), which may limit their potential for treating central nervous system (CNS) disorders as they may not reach the target site in the brain effectively (20).

In terms of their interaction with P-glycoprotein (Pgp), all compounds except for CID 9564 are substrates, meaning they are likely to be actively transported out of cells, including those in the brain and intestines. This property can affect their bioavailability and distribution, potentially leading to drug-drug interactions, reduced therapeutic efficacy, and limiting drug concentration in target cells (20).

Compound CID 136340365 is identified as a CYP2C19 inhibitor. CYP2C19 is an enzyme responsible for metabolizing several drugs, including the antidepressants citalopram and escitalopram, and the antiplatelet agent clopidogrel. Inhibiting CYP2C19 can lead to higher plasma concentrations of these drugs, potentially increasing therapeutic effects or the risk of adverse effects. It is crucial to consider the patient's genotype for CYP2C19, as poor metabolizers may experience significantly higher drug levels when co-administered with CYP2C19 inhibitors (18,24).

Compounds CID 54698642, CID 56841816, CID 76419031, CID 68116543, and CID 136340365 are known as CYP2C9 inhibitors. CYP2C9 plays a significant role in metabolizing various medications, such as warfarin (an anticoagulant), phenytoin (an anticonvulsant), and many non-steroidal anti-inflammatory drugs (NSAIDs). Inhibiting this enzyme can increase drug bioavailability and prolong half-life, which can have positive or negative effects depending on the therapeutic index of the drug and the patient's response. Monitoring drug levels and adjusting doses is important when co-administering CYP2C9 inhibitors with CYP2C9 substrates (18,24).

Compound CID 136340365 is also reported to inhibit CYP2D6, which metabolizes approximately 25% of all drugs, including many antidepressants (e.g., fluoxetine, paroxetine) and beta-blockers (e.g., metoprolol). Inhibiting CYP2D6 can elevate plasma concentrations of these drugs, potentially enhancing efficacy or toxicity. This is especially critical for drugs with a narrow therapeutic window, such as certain opioids and antidepressants, where precise dosing is crucial (18,24).

Compounds CID 56841816, CID 76419031, and CID 136340365 are classified as CYP3A4 inhibitors. CYP3A4 is the most abundant CYP enzyme in the liver and is involved in metabolizing a wide range of drugs, including statins, calcium channel blockers, and immunosuppressants. Inhibiting CYP3A4 can increase systemic exposure to

CYP3A4 substrates, leading to therapeutic or toxic effects. For instance, using a CYP3A4 inhibitor with a statin can significantly raise the risk of myopathy, a severe muscle condition (18,24).

The human skin permeability coefficient (log Kp) measures how easily a compound can be absorbed through the skin. All studied compounds have low log Kp values. Compound CID 76212788 has a log Kp value of -8.24 cm/s, indicating it is unlikely to be absorbed dermally. This low value is consistent with its low bioavailability score of 0.17, indicating limited overall absorption. While poor skin absorption is not a concern for oral or intravenous use, the low bioavailability score suggests that careful dosing and formulation are necessary to achieve effective therapeutic levels in the bloodstream (18,20).

The synthetic accessibility score of compound 102045500 is particularly higher than that of the other studied compounds, indicating that the synthesis of all studied compounds is readily possible. This is an essential factor in the practicality and cost-effectiveness of producing drugs for clinical use (18,20).

Conclusion

The molecular docking study on the HIV-1 integrase model [PDB ID: 6NUJ] highlights the potential of MK-2048 and CID 76212788 as inhibitors. The integrase enzyme is crucial for the life cycle of the virus, making it a key target for antiviral therapies. CID 76212788 shows the strongest binding affinity to the integrase, forming hydrogen bonds with important amino acids, which indicates a solid basis for designing effective inhibitors. Research on integrase inhibitors, including MK-2048 and new macrocyclic derivatives, is vital for the treatment of HIV, especially as drug resistance grows. Understanding how these molecules interact with the integrase and their pharmacokinetic properties is essential for improving antiretroviral therapy. This study contributes to knowledge about integrase inhibition and supports the development of more effective inhibitors for various HIV-1 subtypes.

Acknowledgements

The authors would like to thank Behbahan Faculty of Medical Sciences for their support and cooperation in conducting this research.

Competing Interests

The author declared no conflict of interests in this study.

References

1. Ceccherini-Silberstein F, Malet I, D'Arrigo R, Antinori A, Marcelin AG, Perno CF. Characterization and structural analysis of HIV-1 integrase conservation. AIDS Rev. 2009;11(1):17-29.
2. Quashie PK, Han YS, Hassounah S, Mesplède T, Wainberg MA. Structural studies of the HIV-1 integrase protein: compound screening and characterization of a DNA-binding inhibitor. PLoS One. 2015;10(6):e0128310. doi: 10.1371/journal.pone.0128310.
3. Busschots K, Voet A, De Maeyer M, Rain JC, Emiliani S, Benarous R, et al. Identification of the LEDGF/p75 binding site in HIV-1 integrase. J Mol Biol. 2007;365(5):1480-92. doi: 10.1016/j.jmb.2006.10.094.
4. Esposito D, Craigie R. HIV integrase structure and function. Adv Virus Res. 1999;52:319-33. doi: 10.1016/s0065-3527(08)60304-8.
5. Maertens GN, Engelman AN, Cherepanov P. Structure and function of retroviral integrase. Nat Rev Microbiol. 2022;20(1):20-34. doi: 10.1038/s41579-021-00586-9.
6. Mouscadet JF, Delelis O, Marcelin AG, Tchertanov L. Resistance to HIV-1 integrase inhibitors: a structural perspective. Drug Resist Updat. 2010;13(4-5):139-50. doi: 10.1016/j.drup.2010.05.001.
7. Goethals O, Clayton R, Van Ginderen M, Vereycken I, Wagemans E, Geluykens P, et al. Resistance mutations in human immunodeficiency virus type 1 integrase selected with elvitegravir confer reduced susceptibility to a wide range of integrase inhibitors. J Virol. 2008;82(21):10366-74. doi: 10.1128/jvi.00470-08.
8. Ndashimye E, Avino M, Olabode AS, Poon AF, Gibson RM, Li Y, et al. Accumulation of integrase strand transfer inhibitor resistance mutations confers high-level resistance to dolutegravir in non-B subtype HIV-1 strains from patients failing raltegravir in Uganda. J Antimicrob Chemother. 2020;75(12):3525-33. doi: 10.1093/jac/dkaa355.
9. Rhee SY, Liu TF, Kiuchi M, Zioni R, Gifford RJ, Holmes SP, et al. Natural variation of HIV-1 group M integrase: implications for a new class of antiretroviral inhibitors. Retrovirology. 2008;5:74. doi: 10.1186/1742-4690-5-74.
10. Trivedi J, Mahajan D, Jaffe RJ, Acharya A, Mitra D, Byrareddy SN. Recent advances in the development of integrase inhibitors for HIV treatment. Curr HIV/AIDS Rep. 2020;17(1):63-75. doi: 10.1007/s11904-019-00480-3.
11. Pandey KK, Bera S, Vora AC, Grandgenett DP. Physical trapping of HIV-1 synaptic complex by different structural classes of integrase strand transfer inhibitors. Biochemistry. 2010;49(38):8376-87. doi: 10.1021/bi100514s.
12. Shacklett BL, Blanco J, Hightow-Weidman L, Mgodi N, Alcamí J, Buchbinder S, et al. HIV research for prevention 2018: from research to impact conference summary and highlights. AIDS Res Hum Retroviruses. 2019;35(7):598-607. doi: 10.1089/aid.2019.0074.
13. Koneru PC, Francis AC, Deng N, Rebensburg SV, Hoyte AC, Lindenberger J, et al. HIV-1 integrase tetramers are the antiviral target of pyridine-based allosteric integrase inhibitors. Elife. 2019;8. doi: 10.7554/eLife.46344.
14. National Center for Biotechnology Information. PubChem Compound Summary for CID 76212788. 2025. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/76212788. Retrieved May 24, 2025.
15. Sun L, Gao P, Zhan P, Liu X. Pyrazolo[1,5-a]pyrimidine-based macrocycles as novel HIV-1 inhibitors: a patent evaluation of WO2015123182. Expert Opin Ther Pat. 2016;26(9):979-86. doi: 10.1080/13543776.2016.1210127.
16. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, et al. PubChem substance and compound databases. Nucleic Acids Res. 2016;44(D1):D1202-13. doi: 10.1093/nar/gkv951.
17. Bitencourt-Ferreira G, de Azevedo WF Jr. Molegro virtual docker for docking. Methods Mol Biol. 2019;2053:149-67. doi: 10.1007/978-1-4939-9752-7_10 .
18. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. 2017;7:42717. doi: 10.1038/srep42717.
19. Daina A, Michielin O, Zoete V. iLOGP: a simple, robust, and efficient description of n-octanol/water partition coefficient for drug design using the GB/SA approach. J Chem Inf Model. 2014;54(12):3284-301. doi: 10.1021/ci500467k.
20. Daina A, Zoete V. A boiled-egg to predict gastrointestinal absorption and brain penetration of small molecules.

ChemMedChem. 2016;11(11):1117-21. doi: 10.1002/cmdc.201600182.
21. Bitencourt-Ferreira G, de Azevedo WF Jr. Molegro virtual docker for docking. Methods Mol Biol. 2019;2053:149-67. doi: 10.1007/978-1-4939-9752-7_10.
22. Yousefi R, Mokarmian R, Jamshidi A. Molecular docking of pepstatin A and compounds with structural similarity to the molecular model of human BACE-1 enzyme. J Adv Pharm Res. 2023;7(4):181-98. doi: 10.21608/aprh.2023.219909.1223.
23. Yousefi R, Mokarmian S, Jamshidi A. Efficacy of beta-secretase-1 enzyme inhibitors in Alzheimer's disease. J Adv Pharm Res. 2023;7(4):243-50. doi: 10.21608/aprh.2023.230890.1234.
24. Hasler JA, Estabrook R, Murray M, Pikuleva I, Waterman M, Capdevila J, et al. Human cytochromes P450. Mol Aspects Med. 1999;20(1):1-137. doi: 10.1016/s0098-2997(99)00005-9.