e-ISSN: 2687-6698 **Research Article**

Turkish Journal of Analytical Chemistry
www.turkjac.org
2019
TurkJAC

Designing potential HIV-1 integrase inhibitors: An *in silico* approach

Arif Mermer*

*University of Health Sciences-Turkey, Experimental Medicine Research and Application Center, Uskudar, 34662, Istanbul, Turkey

Abstract

Human immunodeficiency virus is a spectrum of conditions caused by infection with the human immunodeficiency virus. In 2019, about 38 million people worldwide were living with HIV and 690,000 deaths had occurred in that year. To date, for the treatment of HIV-1 disease, many compounds have been synthesized and some of them were approved by FDA. However, the use of these drugs has been limited due to reasons such as resistance caused by the misuse of drugs and bad side effects. We describe herein designing 48 novel compounds as a potential inhibitor of HIV-1 integrase through *in silico* studies such as molecular docking, target analysis, toxicity prediction, and ADME prediction. The online web-based platform, SwissADME, also predicts these molecules' solubility, pharmacodynamics property, and target accuracy.

Keywords: Swissadme, HIV-1 integrase, molecular docking, *in silico* study, toxicity

1. Introduction

HIV-1 integrase (IN) represents a charming aim in anti-HIV drug design fundamentally because of its specificity. Hence, HIV-1 IN does not possess a functional equivalent in humans and exhibits a crucial role in forming irreversible and productive viral infections [1]. This enzyme catalyzes the insertion of proviral DNA, obtained from reverse transcription of HIV-1 RNA, into the genome of the host-infected cells. The insertion is carried through a two-step enzymatic phase which is endonucleolytic cleavage of a terminal dinucleotide (GT) and "strand transfer" (ST). Consequently, both reactions are finalized by the catalytic core domain of HIV-1 IN which includes two divalent metal ion cofactors (Mg^{2+}). These metal ions are coordinated by three catalytic carboxylate residues: Asp64, Asp116, and Glu152 (DDE triad) within the enzyme active site [2]. To aim the metal cofactors within the active site of a viral metal-activated enzyme-like HIV-1 IN has shown up as an appealing and confirmed strategy for the improvement of novel anti-HIV agents. For this purpose, a metal-binding pharmacophore model has been utilized to design novel HIV-1 IN inhibitors as given in Fig. 1A. Two distinctive features are particularly taken into account in this model: the first is a planar metal binding site that can interact with metals present in the enzyme's active site, while the second is the presence of an aromatic or heteroaromatic hydrophobic functional group located close to the metal-binding site [3,4]. Constant workings in using this pharmacophore model have ended up the design and following FDA approval of three INIs for clinical use as potent anti-HIV drugs: Raltegravir (RLT), Elvitegravir (EVG), and Dolutegravir (DTG) in 2007, 2012, and 2013, respectively (Fig. 1B) [5-7].

Various MBGs have been widely researched to design progressive and potent INIs [8,9]. Isonaphthylic acid (INA) derivatives can potentially be studied as novel HIV-1 IN inhibitors due to the metal-binding sites in their structure. Thus, INA derivatives represent an effective class of aromatic ligands with powerful bidentate chelating capacity toward metal ions. Moreover, the presence of an amide group as a linker plays an important role in the inhibitory activity of these types of compounds. And an aromatic functional group linked to the amide is considered for ensuring the basic interactions with the hydrophobic pocket of the enzyme [10].

In drug discovery studies, many parameters such as which reactants to be used in which amount, substituent selection, and whether they are biologically active or not are evaluated in order to avoid time and resource

Citation: A. Mermer, Designing Potential HIV-1 Integrase Inhibitors: An *in Silico* Approach, Turk J Anal Chem, 3(2), 2021, 45-53.

https://doi.org/10.51435/turkjac.950248

***Author of correspondence:** arfmermer@hotmail.com
Tel: +90 (216) 418 9616
Fax +90 (216) 418 9620

Received: 09 June, 2021
Accepted: 14 October, 2021

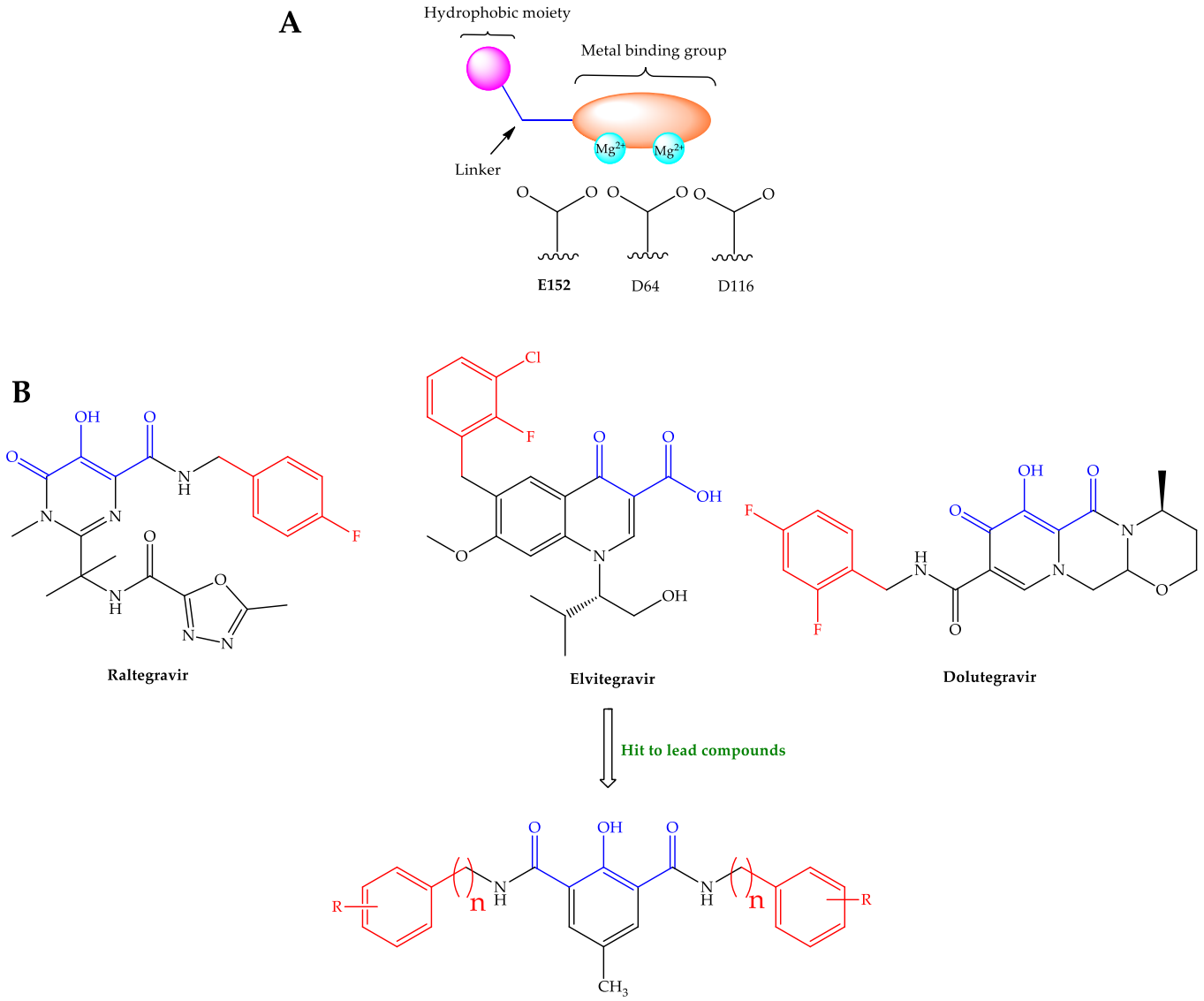


Figure 1. Demonstration of the pharmacophore model for HIV-1 INs. (**B**) FDA-approved drugs against HIV-1 INIs. Atoms marked with blue of MBG of the drugs can chelate the Mg^{2+} ions. The hydrophobic part of the compounds is marked with red

consumption. The molecule to be synthesized must have high activity and low toxicity at the same time. Equally significant is the access to and concentration at the therapeutic target in the organism. The conventional step to consider pharmacokinetics is to separate the diverse impacts that influence the access to the target into individual parameters. In this context, these ADMET parameters (for absorption, distribution, metabolism, excretion, and toxicity) can be evaluated separately by dedicated methods. In drug design and discovery studies, it has been dedicated that the early estimation of ADMET properties greatly reduces the potential disadvantages in clinical and phase stages. Computer models have been encouraged as an effective alternative to experimental procedures for the estimation of ADME, particularly at early steps, when researched chemical structures are countless but the availability of compounds is limited [11,12].

The SwissADME is a web-based platform that is freely accessible at http://www.swissadme.ch and it can be simply used even by non-experts in computer-aided drug design (CADD) studies and the results can be easily analyzed. Compared to other web-based tools for determining ADME and pharmacokinetic properties, SwissADME key points are, partially: various input ways, computation for multiple compounds, and the opportunity to show, save and share outcomes per individual compounds or through global intuitive and interactive graphs. Consequently, SwissADME is combined with the SwissDrugDesign field. One-click interoperability gains access to different CADD tools improved by the Molecular Modeling Group of the SIB Swiss Institute of Bioinformatics, e.g., ligand-based virtual screening, bio target prediction, molecular docking, bioisosteric design, or molecular mechanics [13-21].

In the light of the above considerations, this paper includes *in silico* analysis of 48 novel HIV-1 IN inhibitors, containing different substituents and linkers such as fluorine, chlorine, bromide, methoxy, which is provided

with the target protein's inhibition site (PDB ID: 1QS4), prediction of ADME, Target Prediction were done by using SwissBioinformatics online Tools. The prediction of toxicity of designing compounds was screened via the pkCSM online web tool. Furthermore, the molecular docking studies of the most potent compound were performed.

2. Materials and Methods

2.1. Molecular Docking

The co-crystallized structure of HIV-1 integrase enzyme [PDB: 1QS4] was taken from Protein Data Bank (PDB) and prepared by utilizing the Protein Preparation Wizard Module of Schrödinger Software Suite. Then it was optimized by removing the water molecules, heteroatoms, and co-factors. The hydrogens, missing atoms, bonds, and charges were computed through Maestro [26, 27]. Compound **24** exhibited the highest drug-likeness score was selected for docking study. The ligand preparation and optimization contain forming different tautomers, assigning bond orders, ring conformations that were minimized using the OPLS2005 force field before the docking study, and stereochemistries were performed via the LigPrep module of Schrödinger Software Suite. Besides, a receptor grid was formed around the co-crystallized ligand of the enzyme. The grid box size was set to 20 Å Radius, using the receptor Grid Generation applied in Glide. Extra Precision (XP) mode and Glide programs were used for the docking calculations [28-30].

2.2. ADME Prediction

ADME (adsorption, distribution, metabolism, and excretion) is significant to search the pharmacodynamics of the designed compounds which could be a target agent in drug design and discovery studies. SwissADME is a web-based platform that lets the user upload or draw their hit compounds with structure or SMILES code. This tool supplies many parameters such as lipophilicity (iLOGP, XLOGP3, WLOGP, MLOGP, SILICOS-IT, Log $P_{o/w}$), water solubility- Log S (ESOL, Ali, SILICOS-IT), drug-likeness rules (Lipinski, Ghose, Veber, Egan, and Muegge) and Medicinal Chemistry (PAINS, Brenk, Leadlikeness, Synthetic accessibility) methods [22]. The designed novel HIV-1 IN inhibitors were uploaded with SMILES codes and analyzed.

2.1. Target Prediction

Molecular Target investigations are crucial to determine the phenotypical side effects or potential cross-reactivity induced by the action of bioorganic compounds of which molecular weight is not bigger than 500 g/mol [23]. The designed compound's SMILES codes were uploaded to the Swiss Target Prediction website in order to analyze their target prediction (https://www.swisstargetprediction.ch).

2.2. Toxicity Prediction

Toxicology prediction of bioorganic compounds is substantial to estimate the amount of tolerability of the hit compounds before *in vitro*, *in vivo*, and clinical studies. pkCSM is also a web-based platform for analyzing physicochemical properties of small compounds, and this online website supplies many toxicology results such as LD50, hERG-I inhibitor, AMES Toxicity, hERG-II inhibitor, human maximum tolerated dose, LOAEL, Skin Toxicity, *T. pyriformis* toxicity, Hepatotoxicity, and Minnow toxicity. The designed compounds' SMILES codes were uploaded pkCSM website to analyze their target prediction (http://biosig.unimelb.edu.au/pkcsm/) [24].

3. Results and Discussion

3.1. Molecular Docking

Considering the ADME and toxicity outcomes, a deep docking study was implemented to regard the possible binding modes of the promising compound (24) inside the active site of HIV-1 integrase enzyme (PDB ID: 1QS4) by using Schrödinger Software. Initially, the co-crystalized ligand 1-(5-chloroindol-3-yl)-3-hydroxy-3-(2H-tetrazol-5-yl)-propenone (CCL) and the protein (PDB ID: 1QS4) were modeled to validate the Glide. Superimposition of the experimental bound (co-crystallized) conformation of 1-(5-chloroindol-3-yl)-3-hydroxy-3-(2H-tetrazol-5-yl)-propenone estimated by Glide. The experimental binding conformations of co-crystalized ligand in the binding pocket of 1QS4 were successfully generated with an acceptable root-mean-square deviation (RMSD) of 0.175 Å (< 2 Å) [31]. The interactions between co-crystallized ligand and the active site of HIV-1 IN was shown in Fig. 2A. Compound **24** and its interactions with the binding side of 1QS4 were demonstrated in Fig. 2B. Hereby, compound **24** was surrounded with LYS 159, LYS 156, ASN 155, GLU 152, ILE 151, GLY 149, and GLN 148 amino acid residues of the active site of HIV-1 IN enzyme. Further, the carbonyl group of amide generated a strong hydrogen-bonding interaction with ASN 155, and the hydroxy group between two amide bonds in the phenyl ring formed a strong hydrogen-bonding interaction with ASP 64. The docking score of the chosen compound and co-crystallized ligand were given in Table 1.

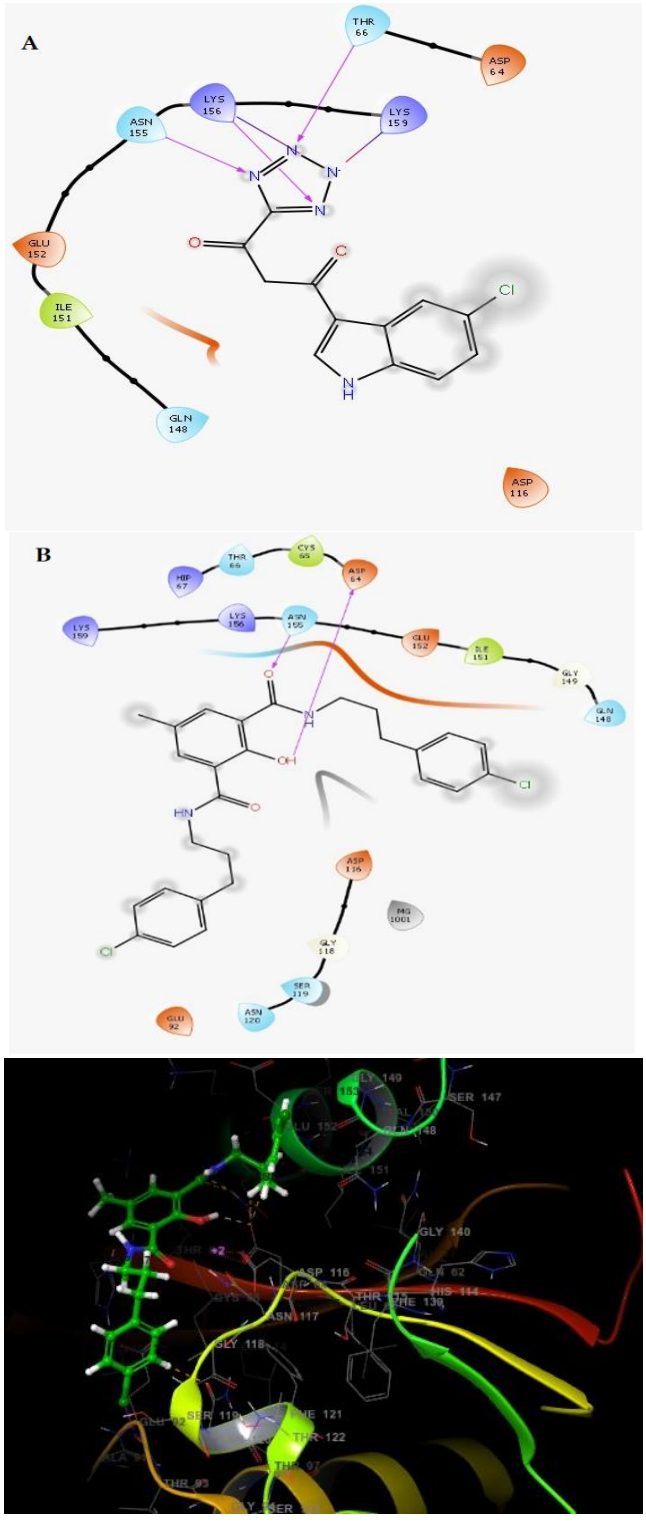


Figure 2. **A**: Co-crystallized ligand and its interaction of active site of HIV-1 IN, **B**: Compound **24** and its interaction of active site of HIV-1 IN

Table 1. The docking scores the chosen compound, co-crystallized ligand and its binding interactions with the active site of HIV-1 IN (PDB ID: 1QS4)

Ligand	Docking score	H-bond	Metal coordination	Pi-Pi stacking	Pi-cation	Salt bridge
CCL	-7.5	ASN 155 LYS 156 THR 66 LYS 159	x	x	x	x
24	-6.890	ASN 155 ASP 64	x	x	x	x

Table 2. The physiochemical property results of the designed compounds

Comp. No	TPSA (Å²)	*n*HD	*n*HA	Log $P_{o/w}$	Log *S*	Log *Kp* (cm/s)	Lipinski	Druglikeness Score
1	78.43	3	5	4.13	-5.03	-5.54	0	0.24
2	78.43	3	5	4.16	-4.97	-5.8	0	0.25
3	78.43	3	5	4.70	-5.56	-5.32	1	0.31
4	78.43	3	5	5.38	-6.02	-4.99	1	0.35
5	78.43	3	5	4.01	-5.03	-5.54	0	0.11
6	78.43	3	5	4.24	-4.97	5.8	0	0.49
7	78.43	3	5	4.80	-5.56	-5.32	1	0.58
8	78.43	3	5	5.35	-6.02	-4.99	1	0.63
9	78.43	3	5	3.99	-5.03	-5.54	0	0.60
10	78.43	3	5	4.17	-4.97	-5.8	0	0.54
11	78.43	3	5	4.72	-5.56	-5.32	1	0.66
12	78.43	3	5	5.40	-6.02	-4.99	1	0.72
13	78.43	3	3	4.49	-5.9	-4.99	0	0.42
14	78.43	3	3	4.68	-5.83	-5.26	0	0.46
15	78.43	3	3	5.17	-6.43	-4.77	1	0.50
16	78.43	3	3	5.76	-6.9	-4.43	1	0.53
17	78.43	3	3	4.46	-5.9	-4.99	0	0.18
18	78.43	3	3	4.55	-5.83	-5.26	0	0.58
19	78.43	3	3	5.12	-6.43	-4.77	1	0.66
20	78.43	3	3	5.68	-6.9	-4.43	1	0.71
21	78.43	3	3	4.45	-5.9	-4.99	0	0.77
22	78.43	3	3	4.68	-5.83	-5.26	0	0.71
23	78.43	3	3	5.20	-6.43	-4.77	1	0.81
24	78.43	3	3	5.85	-6.9	-4.43	1	0.86
25	78.43	3	3	4.69	-6.54	-5.44	2	0.16
26	78.43	3	3	4.81	-6.46	-5.71	2	0.19
27	78.43	3	3	5.34	-7.06	-5.22	2	0.25
28	78.43	3	3	5.90	-7.53	-4.88	2	0.25
29	78.43	3	3	4.61	-6.54	-5.44	2	-0.02
30	78.43	3	3	4.74	-6.46	-5.71	2	0.39
31	78.43	3	3	5.31	-7.06	-5.22	2	0.47
32	78.43	3	3	5.93	-7.53	-4.88	2	0.52
33	78.43	3	3	4.61	-6.54	-5.44	2	0.46
34	78.43	3	3	4.78	-6.46	-5.71	2	0.42
35	78.43	3	3	5.41	-7.06	-5.22	2	0.53
36	78.43	3	3	5.94	-7.53	-4.88	2	0.59
37	96.89	3	5	3.57	-5.21	-5.48	0	-0.18
38	96.89	3	5	3.50	-4.79	-6.13	0	0.12
39	96.89	3	5	4.16	-5.39	-5.65	0	0.20
40	96.89	3	5	4.70	-5.86	-5.31	0	0.25
41	96.89	3	5	3.38	-4.86	-5.87	0	-0.04
42	96.89	3	5	3.50	-4.79	-6.13	0	0.39
43	96.89	3	5	4.17	-5.39	-5.65	0	0.54
44	96.89	3	5	4.72	-5.86	-5.31	0	0.60
45	96.89	3	5	3.41	-4.86	-5.87	0	0.30
46	96.89	3	5	3.52	-4.79	-6.13	0	0.32
47	96.89	3	5	4.17	-5.39	-5.65	0	0.48
48	96.89	3	5	4.72	-5.86	-5.31	0	0.55

Comp.: Compound

3.2. ADME Prediction

The ADME predictions of the designed compounds were performed by SwissADME database, and

physiochemical properties; hydrogen bond acceptors (*n*HA), hydrogen bond donors (*n*HD), and topological polar surface area (TPSA). Besides, lipophilicity (iLOGP, XLOGP3, WLOGP, MLOGP, SILICOS-IT, and Consensus $P_{o/w}$), water-solubility properties (ESOL, Ali, SILICOS-IT), and drug-likeness factors (Lipinski's Rules, Ghose, Veber, Egan, Muegge) were also calculated. All designed compounds showed high topological polar surface area (TPSA) ranging from 78.43 Å^2 to 96.89 Å^2. The methoxy-containing compounds were found to be higher than the others. *n*HD of the compounds was found 3, whereas *n*HA of compounds was between 3 and 5. The calculated lipophilicity properties were given in the consensus model (*n*-octanol and water: Log $P_{o/w}$) which was ranging from 3.41 to 5.94. When the water solubility properties were examined, all compounds were found to be moderately or poorly soluble with the Log *S* values ranging from -4.79 to -7.53 mg/mL. Pharmacokinetic properties such as GI absorption, BBB permeant, P-gp substrate, and skin permeation (Log *K*p) were also calculated, and the designed compounds exhibited high Gastrointestinal absorption (GI), while none of them showed blood-brain barrier permeant. Skin permeation kinetics (Log *K*p) of them were found to be -4.43-6.13 cm/s (Table 2).

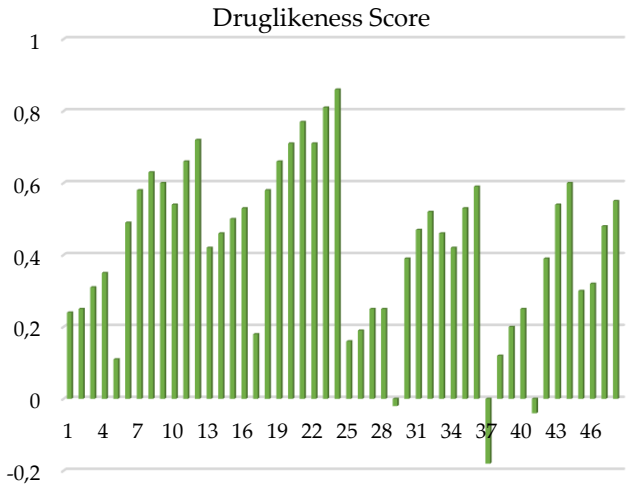


Figure 3. Drug-likeness score of the designed compounds

The Lipinski rule of five is an important parameter for drug-likeness factors of the compounds, and it has to be ≤ 1. Among the 48 designed compounds, compounds **25**-**36** containing bromophenyl as a hydrophobic moiety did not obey this rule, and for this reason, these compounds were eliminated. Drug-likeness scores was specified via the Molsoft database (www.molsoft.com). Although all compounds showed drug-likeness score with varying values, among them, compounds **12**, **21**-**24** displayed the highest values with 0.72, 0.77, 0.71, 0.81, and 0.86, respectively.

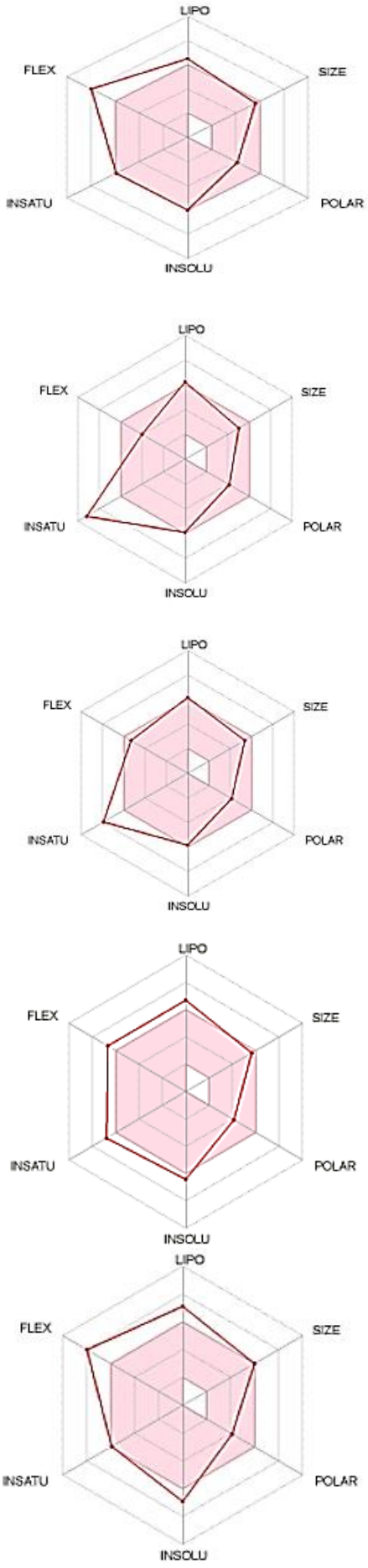


Figure 4. The Bioavailability Radar enables at first glance at the drug-likeness of the compounds

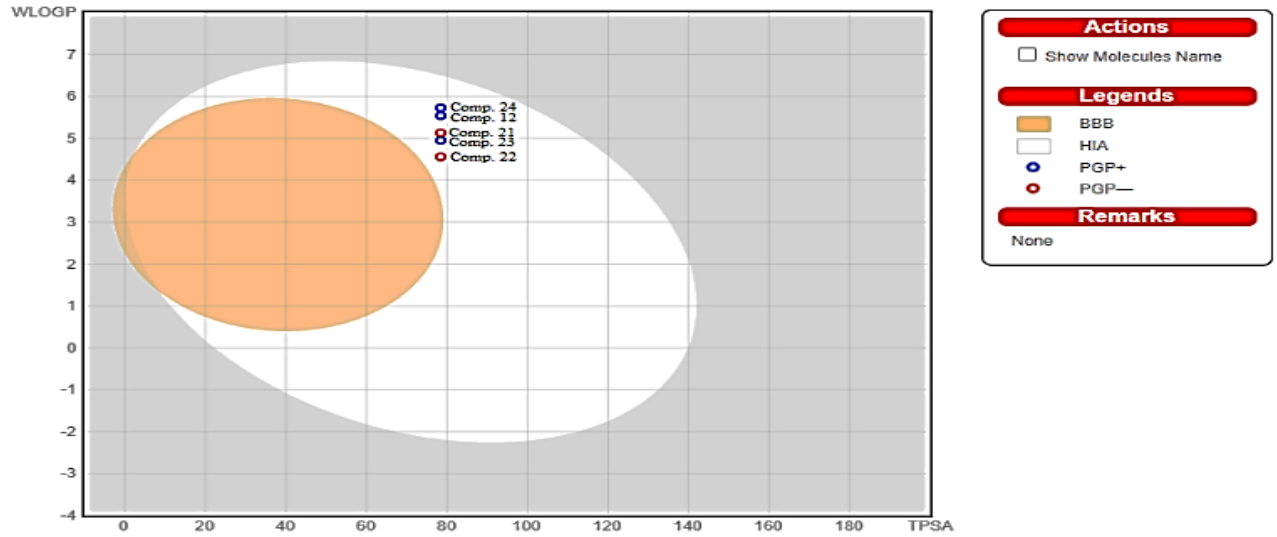


Figure 5. BOILED-Egg presentation of the compounds

Compared to methoxy-containing compounds which showed the highest TPSA results with 96.89, these five compounds exhibited a 1-6-fold better drug-likeness score than them (All SwissADME results were given in detail in Fig. 3). Moreover, the bioavailability score of these compounds was found to be 0.55 (Fig. 4). For further *in silico* analysis, the above compounds were chosen as hit compounds.

In Fig. 4, the pink area describes the optimal range for each property (lipophilicity: XLOGP3 between -0.7 and +5.0, size: MW between 150 and 500 g/mol, polarity: TPSA between 20 and 130 Å^2, solubility: log *S* not higher than 6, saturation: fraction of carbons in the sp^3 hybridization not less than 0.25, and flexibility: no more than 9 rotatable bonds.

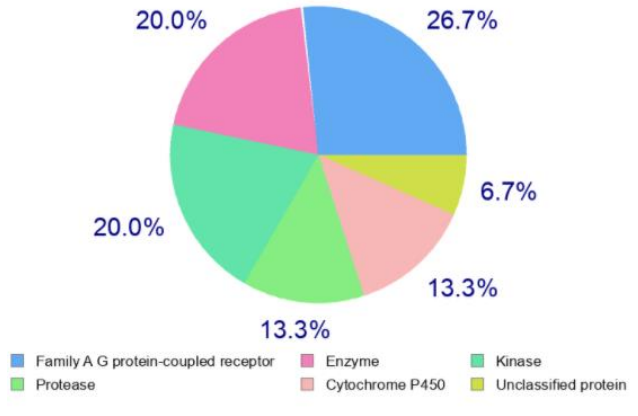
Furthermore, the BOILED-Egg profile which lets for intuitive consideration of passive gastrointestinal absorption (HIA) and brain penetration (BBB) in the function of the position of the molecules in the WLOGP-vs-TPSA referential was screened for the selected five compounds [25]. The white area is for the high probability of passive absorption by the gastrointestinal tract, while the yellow area is for the high probability of brain penetration. Also, the marks are colored in blue if predicted as actively effluxed by P-gp (PGP+) and in red if estimated as non-substrate of P-gp (PGP-). It was concluded that all our compounds were estimated well-absorbed but not accessing the brain, and compounds **12**, **23**, and **24** were subject to active efflux (blue dot), whereas compounds **21** and **22** were not subject to active efflux (red dot) (Fig. 5).

3.3. Target Prediction

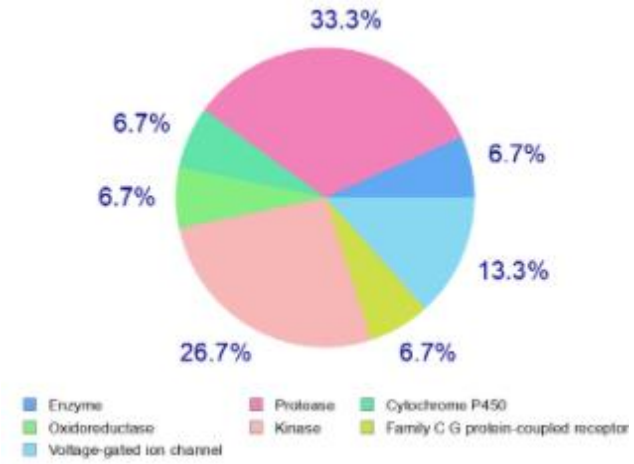
The target estimation of the chosen compounds was examined using the SwissTargetPrediction platform with the following investigations 15 of the outcomes depicted as a pie-chart (Fig. 6). The compound 12 containing *p*-fluorophenyl as a hydrophobic aromatic unit and n = 3 as a linker was predicted as 26.7% of Family AG protein-coupled receptor, 20% of both enzyme and kinase, while compound 21 comprising *p*-chlorophenyl as a hydrophobic aromatic unit and n = 0 was predicted as 33.3% of protease and 26.7% of the kinase. Compounds 22 and 23 were estimated as Family AG protein coupled-receptor with 40.0% and 46.7%, respectively, whereas compound 24 including *p*-chlorophenyl as a hydrophobic aromatic unit and n = 3 was predicted as 46.7% of Family AG protein coupled-receptor and 20% of the enzyme. One of the compound's result tables comprising of Common Name, Uniprot ID, Target, ChEMBL-ID, Probability, Target Class, and Known actives in 2D/3D are given in the Supporting Information file.

3.4. Toxicity Prediction

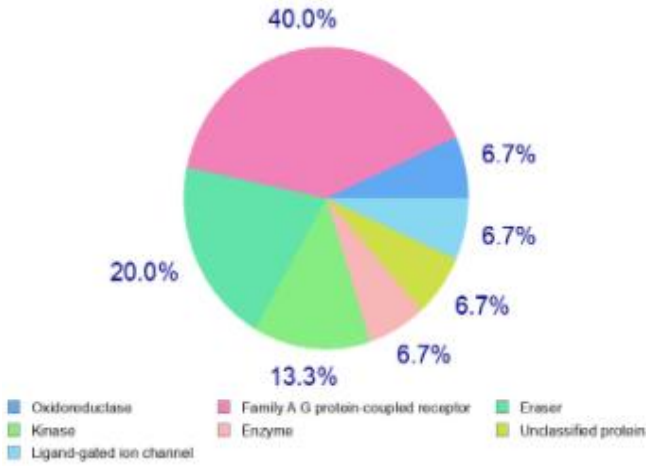
Toxicity predictions were screened and all compounds, except for **22**, do not have any AMES toxicity, while all compounds, except for **12**, produce hepatotoxicity. Moreover, all compounds were not found to be cause skin sensitivity. Although they were also predicted as hEGR II inhibitors, none of them showed any hEGR I inhibition results. It was estimated maximum tolerated dose for humans which was ranging from 0.426 log mg/kg/day to 0.79 log mg/kg/day. Compound **12** exhibited the lowest both oral rat acute toxicity with a LD_{50} value of 2.271 mol/kg and chronic oral rat toxicity with 1.497 log mg/kg_bw/day. 0.397 - 0.672 log mg/L induced *T. Pyriformis* toxicity and -0.765 - 0.027 log mM induced Minnow toxicity were found for the compounds (Table 3).



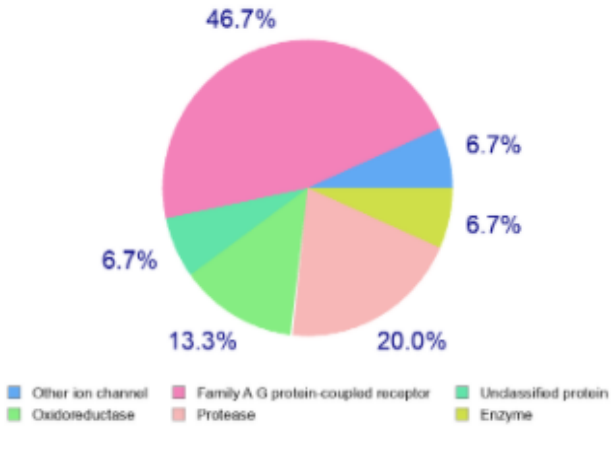
Compound 12



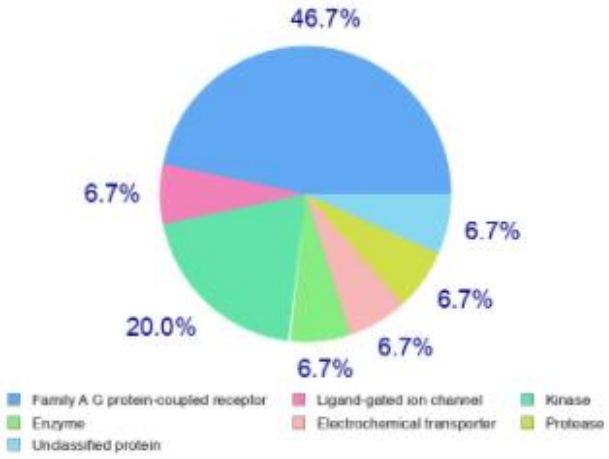
Compound 21



Compound 22



Compound 23



Compound 24

Figure 6. SwissTargetPrediction of the chosen compounds

Table 3. Toxicity prediction results of the selected compounds

Comp. No	AMES Toxicity	Maximum Tolerated Dose	hEGR I inhibitor	hEGR II inhibitor	Oral Rat Acute Toxicity (LD_{50})	Chronic Oral Rat Toxicity (LOAEL)	Hepatotoxicity	Skin Sensitisation	*T.Pyriformis* Toxicity	Minnow Toxicity
12	No	0.457	No	Yes	2.271	1.497	No	No	0.397	-0.954
21	No	0.426	No	Yes	2.462	1.757	Yes	No	0.464	-1.005
22	Yes	0.790	No	Yes	2.395	1.536	Yes	No	0.470	-0.765
23	No	0.581	No	Yes	2.455	1.631	Yes	No	0.672	0.027
24	No	0.520	No	Yes	2.414	1.740	Yes	No	0.555	-0.779

Comp.: Compound

4. Conclusion

In the present study, we have investigated the novel potential compounds against HIV-1 IN with the *in silico* studies. We designed 48 novel compounds with various functional groups such as -F, Cl, -Br, -OCH3 in the phenyl ring which is the hydrophobic site, and we added to the structure linker which is the alkyl chain between amide bond and hydrophobic part. It was observed that the presence of carbonyl and hydroxyl group in the compound was very significant for the binding of the active site of the enzyme. Also, it can be concluded from the SwissADME and toxicity results that the alkyl chain length was also important for both solubility and drug-likeness. It was determined that the solubility decreases as the alkyl chain length increases in the compounds, on the contrary, it was found that there is an increase in the drug-likeness score as the alkyl number increases. We also considered the position of the substituent in the phenyl ring for the activity when designing the compounds, and it was clearly seen that the substituent in the p-position was more important than the other positions. The bromine compounds were found to be more potential for activity, whereas methoxy-containing compounds had less potential. For the best compound, we performed molecular docking studies and hydrogen bonding interactions between ligands and ASN155, ASP 64 amino acids of the receptor exhibited significance for potent HIV-1 IN. Considering the overall results, the five compounds and especially compound **24** can be a potent inhibitor against HIV-1 IN.

Declaration of Competing Interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

I am very grateful to Efe Doğukan Dinçel from İstanbul University, Department of Pharmacy for molecular docking studies.

References

[1] O. Delelis, K. Carayon, A. Saib, E. Deprez, J.F. Mouscadet, Integrase and integration: biochemical activities of HIV-1 integrase, Retrovirology 5, 2008, 114-127.

[2] N. Neamati, Z. Lin, R.G. Karki, A. Orr, K. Cowansage, D. Strumberg, Metal-dependent inhibition of HIV-1 integrase, J Med Chem, 45, 2002, 5661-5670.

[3] T. Kawasuji, T. Yoshinaga, A. Sato, M. Yodo, T. Fujiwara, R. Kiyama, A platform for designing HIV integrase inhibitors. Part 1: 2-hydroxy-3-heteroaryl acrylic acid derivatives as novel HIV integrase inhibitor and modeling of hydrophilic and hydrophobic pharmacophores, Bioorgan Med Chem, 14, 2006, 8430-8445.

[4] B.A. Johns, A.C. Svolto, Advances in two-metal chelation inhibitors of HIV integrase, Expert Opin Ther Pat, 18, 2008, 1225-1237.

[5] M. Rowley, The discovery of raltegravir, an integrase inhibitor for the treatment of HIV infection, Prog Med Chem, 46, 2008, 1-28.

[6] M. Sato, H. Kawakami, T. Motomura, H. Aramaki, T. Matsuda, M. Yamashita, Quinolone carboxylic acids as a novel monoketo acid class of human immunodeficiency virus type 1 integrase inhibitors, J Med Chem, 52, 2009, 4869-4882.

[7] C. Katlama, R. Murphy, Dolutegravir for the treatment of HIV, Expert Opin Inv Drug, 21, 2008, 523-530.

[8] R. Di Santo, Inhibiting the HIV integration process: past, present, and the future, J Med Chem, 57, 2014, 539-566.

[9] C. Liao, C. Marchand, T.R. Burke, Y. Pommier, M.C. Nicklaus, Authentic HIV-1 integrase inhibitors, Future Med Chem, 2, 2010, 1107-1122.

[10] H. Sirous, G. Chemi, S. Gemma, S. Butini, Z. Debyser, F. Christ, L. Saghaie, S. Brogi, A. Fassihi, G. Campiani, M. Brindisi, Identification of novel 3-hydroxy-pyran-4-one derivatives as potent HIV-1 integrase inhibitors using *in silico* structure-based combinatorial library design approach, Front Chem, 7, 2019, 1-20.

[11] M. Hay, D.W. Tomas, J.L. Craighead, C. Economides, J. Rosenthal, Clinical development success rates for investigational drugs, Nat Biotechnol, 32, 2014, 40-51.

[12] J.L. Dahlin, J. Inglese, M.A. Walters, Mitigating risk in academic preclinical drug discovery, Nature Rev Drug Discov, 14, 2015, 279-294.

[13] D.E.V. Pires, T.L. Blundell, D.B. Ascher, pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties usin graph-based signatures, J Med Chem, 58, 2015, 4066-4072.

[14] F. Cheng, W. Li, Y. Zhou, J. Shen, Z. Wu, G. Liu, P.W. Lee, Y. Tang, admetSAR: a comprehensive source and free tool for assessment of chemical ADMET properties, J Chem Inf Model, 52, 2012, 3099-3105.

[15] A. Daina, O. Michielin, V. Zoete, iLOGP: A simple, robust, and efcient description of n-octanol/water partition coefcient for drug design using the GB/SA approach, J Chem Inf Model, 54, 2014, 3284-3301.

[16] L. Di, P. Artursson, A. Avdeef, G.F. Ecker, B. Faller, H. Fischer, J.B. Houston, M. Kansy, E.H. Kerns, S.D. Krämer, H. Lennernäs, K. Sugano, Evidence-based approach to assess passive diffusion and carrier-mediated drug transport, Drug Discov Today, 17, 2012, 905-912.

[17] V. Zoete, A. Daina, C. Bovigny, O. Michielin, SwissSimilarity: A web tool for low to ultra high troughput ligand-based virtual screening, J Chem Inf Model, 56, 2016, 1399-1404.

[18] D. Gfeller, SwissTargetPrediction: a web server for target prediction of bioactive small molecules, Nucleic Acids Res, 42, 2014, W32-W38.

[19] A. Grosdidier, V. Zoete, O. Michielin, SwissDock, a protein-small molecule docking web service based on EADock DSS, Nucleic Acids Res, 39, 2011, W270-W277.

[20] M. Wirth, V. Zoete, O. Michielin, W.H.B. Sauer, SwissBioisostere: a database of molecular replacements for ligand design, Nucleic Acids Res, 41, 2013, D1137-D1143.

[21] V. Zoete, M.A. Cuendet, A. Grosdidier, O. Michielin, SwissParam: a fast force feld generation tool for small organic molecules, J Comput Chem, 32, 2011, 2359-2368.

[22] A. Daina, O. Michielin, V. Zoete, SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules, Sci Rep, 7, 2017, 42717.

[23] M.J. Keiser, B.L. Roth, B.N. Armbruster, P. Ernsberger, J.J. Irwin, B.K. Shoichet, Relating protein pharmacology by ligand chemistry, Nat Biotechnol, 25, 2007, 197-206.

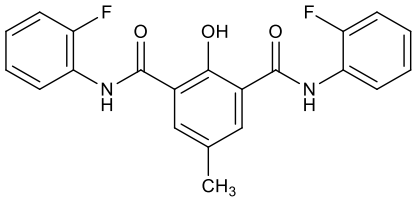
[24] D.E. Pires, T.L. Blundell, D.B. Ascher, PkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graphbased signatures, J Med Chem, 58, 2015, 4066-4072.
[25] A. Daina, V. Zoete, A BOILED-Egg to predict gastrointestinal absorption and brain penetration of small molecules, ChemMedChem, 11, 2016, 1117-1121.
[26] P. Selvam, M. Chandramohan, E. De Clercq, C. Pannecouque, M. Witrouw, Synthesis and anti-HIV activity of 4-[(1,2-dihydro-2-oxo-3H-indol-3-ylidene)amino]-N-(4,6-dimethyl-2-pyrimidinyl)-benzene sulphonamide and its derivatives, Eur J Pharm Sci, 14, 2001, 313-316.
[27] G.M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments, J Comput Aided Mol Des, 27, 2013, 221-234.
[28] R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, J Med Chem, 47, 2004, 1739-1749.
[29] R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T. Halgren, P.C. Sanschagrin, D.T. Mainz, Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein- ligand complexes, J Med Chem, 49, 2006, 6177-6196.
[30] T.A. Halgren, R.B. Murphy, R.A. Friesner, H.S. Beard, L.L. Frye, W.T. Pollard, J.L. Banks, Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening, J Med Chem, 47, 2004, 1750-1759.
[31] B. Debnath, S. Ganguly, Synthesis, biological evaluation, *in silico* docking, and virtual ADME studies of 2-[2-Oxo-3-(arylimino) indolin-1-yl]-N-arylacetamides as potent anti-breast cancer agents, Monatsh Chem, 147, 2016, 565-574.

Supporting information

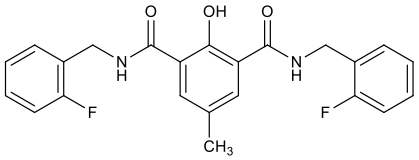
SwissTargetPrediction

Target	Common name	Uniprot ID	ChEMBL ID	Target Class	Probability*	Known actives (3D/2D)
Dopamine D2 receptor (by homology)	DRD2	P14416	CHEMBL217	Family A G protein-coupled receptor	0.120225750913	0 / 405
Serotonin 2a (5-HT2a) receptor (by homology)	HTR2A	P28223	CHEMBL224	Family A G protein-coupled receptor	0.120225750913	0 / 140
11-beta-hydroxysteroid dehydrogenase 1	HSD11B1	P28845	CHEMBL4235	Enzyme	0.120225750913	0 / 150
MAP kinase p38 alpha	MAPK14	Q16539	CHEMBL260	Kinase	0.120225750913	0 / 37
Serine/threonine-protein kinase Chk1	CHEK1	O14757	CHEMBL4630	Kinase	0.120225750913	1 / 0
Serine/threonine-protein kinase WEE1	WEE1	P30291	CHEMBL5491	Kinase	0.120225750913	4 / 0
Trypsin I	PRSS1	P07477	CHEMBL209	Protease	0.120225750913	0 / 28
Epoxide hydratase	EPHX2	P34913	CHEMBL2409	Protease	0.120225750913	0 / 57
Poly [ADP-ribose] polymerase-1	PARP1	P09874	CHEMBL3105	Enzyme	0.120225750913	0 / 84
Cytochrome P450 2C9	CYP2C9	P11712	CHEMBL3397	Cytochrome P450	0.120225750913	0 / 7
Cytochrome P450 2C19	CYP2C19	P33261	CHEMBL3622	Cytochrome P450	0.120225750913	0 / 7
Poly [ADP-ribose] polymerase 2	PARP2	Q9UGN5	CHEMBL5366	Enzyme	0.120225750913	0 / 22
Calcitonin gene-related peptide 1	CALCA	P06881	CHEMBL5293	Unclassified protein	0.120225750913	6 / 0
C-C chemokine receptor type 4	CCR4	P51679	CHEMBL2414	Family A G protein-coupled receptor	0.0	2 / 0
Interleukin-8 receptor A	CXCR1	P25024	CHEMBL4029	Family A G protein-coupled receptor	0.0	12 / 0
C-C chemokine receptor type 6	CCR6	P51684	CHEMBL4423	Family A G protein-coupled receptor	0.0	2 / 0
CCR4-NOT transcription complex subunit 7	CNOT7	Q9UIV1	CHEMBL3616361	Hydrolase	0.0	1 / 0
Glycine transporter 2	SLC6A5	Q9Y345	CHEMBL3060	Electrochemical transporter	0.0	0 / 41
Mu opioid receptor	OPRM1	P35372	CHEMBL233	Family A G protein-coupled receptor	0.0	0 / 411
Kappa Opioid receptor	OPRK1	P41145	CHEMBL237	Family A G protein-coupled receptor	0.0	0 / 219
Delta opioid receptor (by homology)	OPRD1	P41143	CHEMBL236	Family A G protein-coupled receptor	0.0	0 / 349
Plasminogen	PLG	P00747	CHEMBL1801	Protease	0.0	0 / 3
C-C chemokine receptor type 3	CCR3	P51677	CHEMBL3473	Family A G protein-coupled	0.0	0 / 21

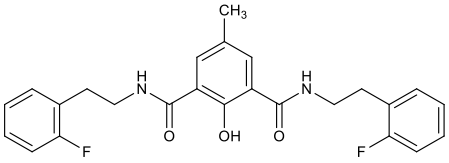
SI-1. Target Prediction results of the compound 12



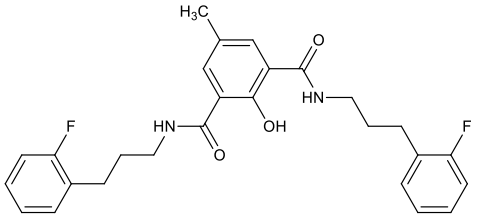
Compound 1



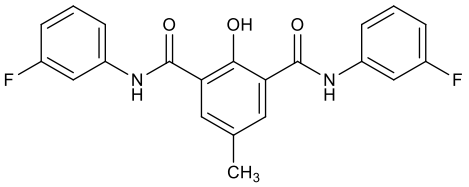
Compound 2



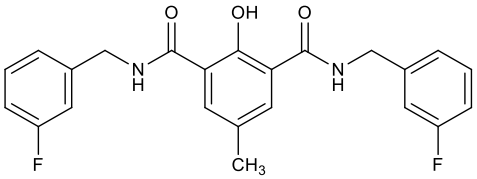
Compound 3



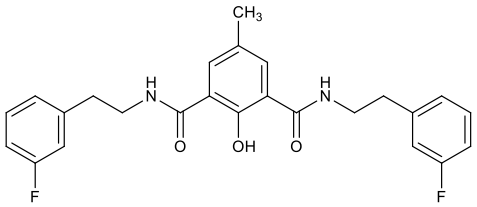
Compound 4



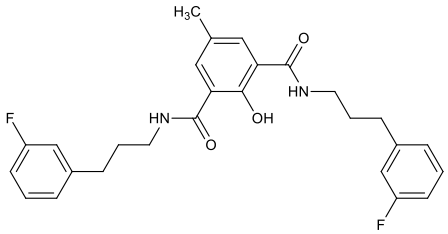
Compound 5



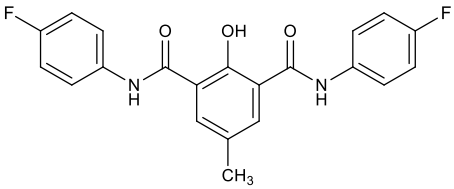
Compound 6



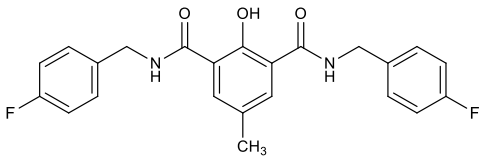
Compound 7



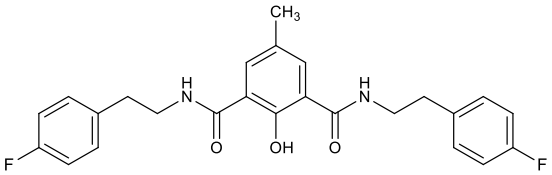
Compound 8



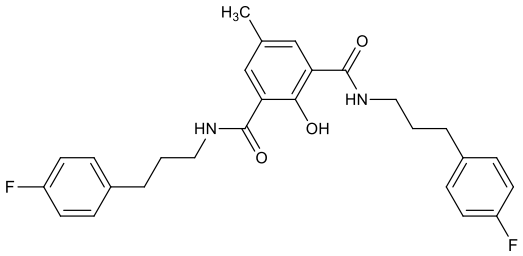
Compound 9



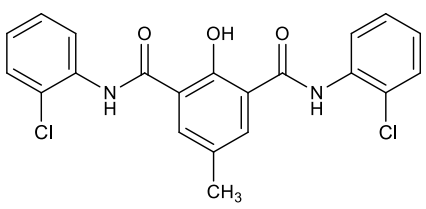
Compound 10



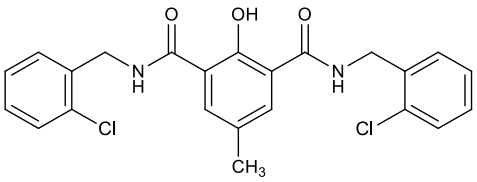
Compound 11



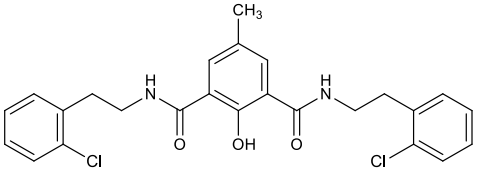
Compound 12



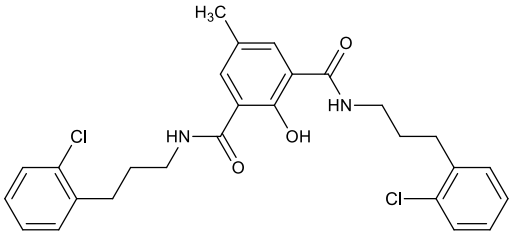
Compound 13



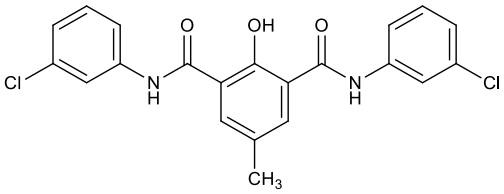
Compound 14



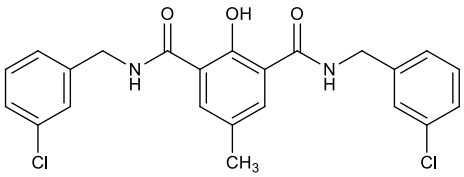
Compound 15



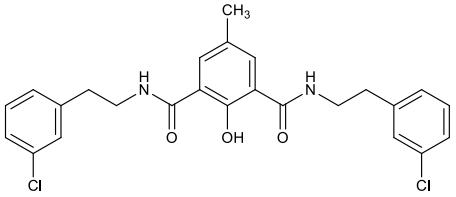
Compound 16



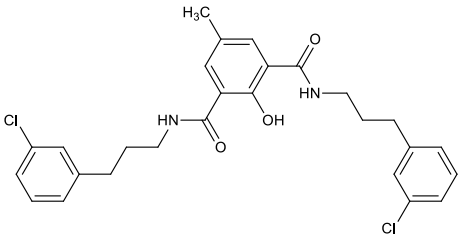
Compound 17



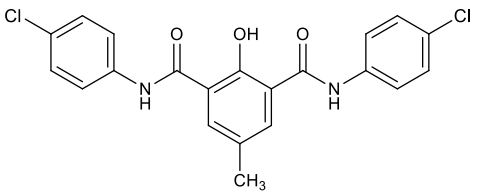
Compound 18



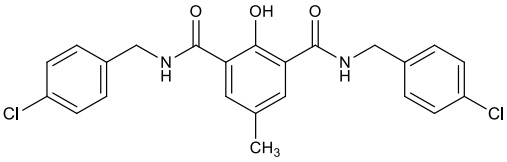
Compound 19



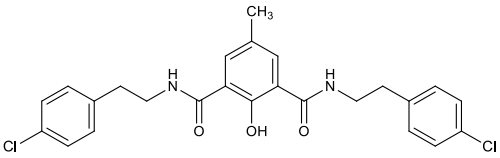
Compound 20



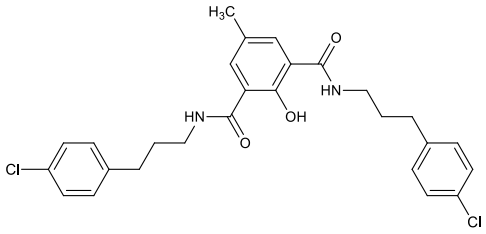
Compound 21



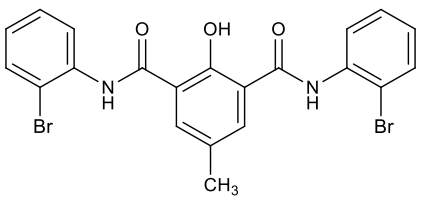
Compound 22



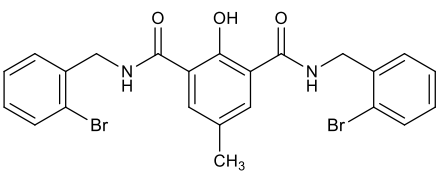
Compound 23



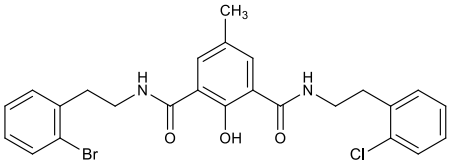
Compound 24



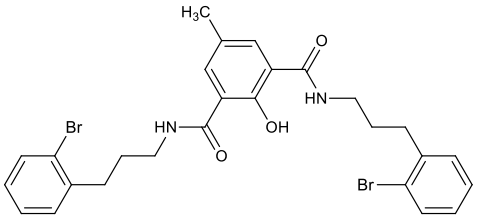
Compound 25



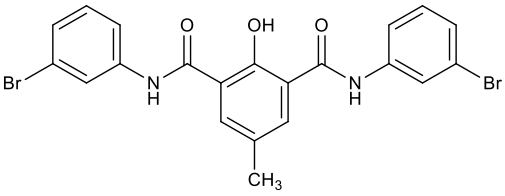
Compound 26



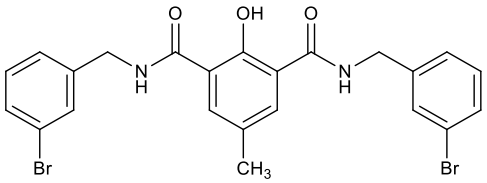
Compound 27



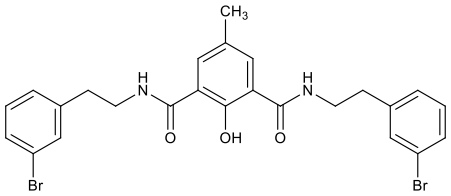
Compound 28



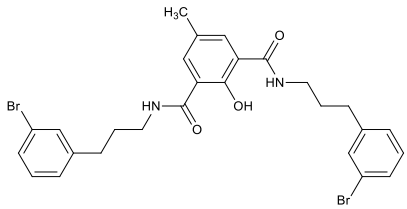
Compound 29



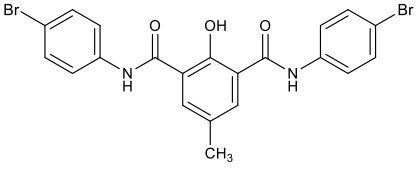
Compound 30



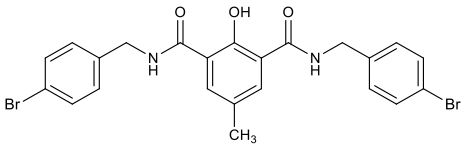
Compound 31



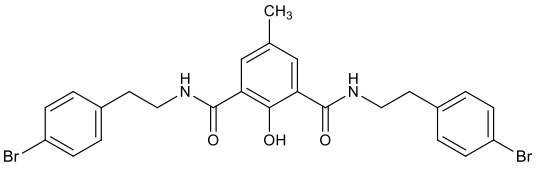
Compound 32



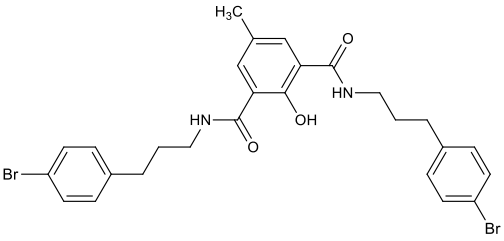
Compound 33



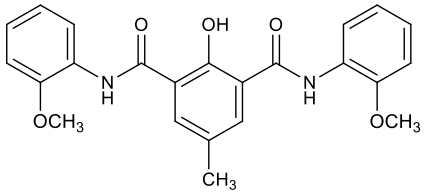
Compound 34



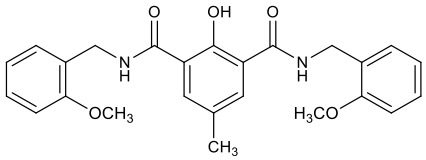
Compound 35



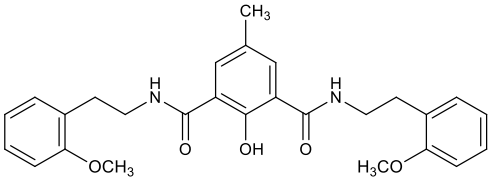
Compound 36



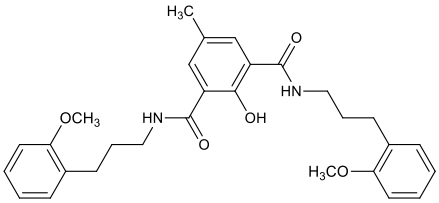
Compound 37



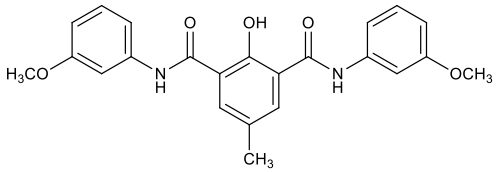
Compound 38



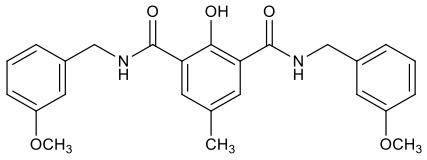
Compound 39



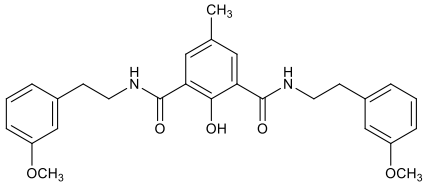
Compound 40



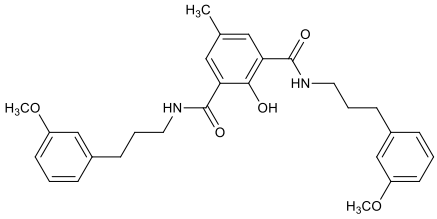
Compound 41



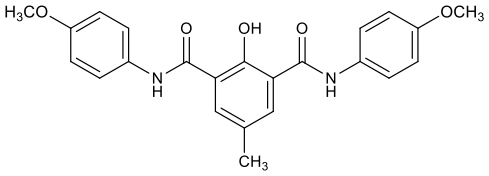
Compound 42



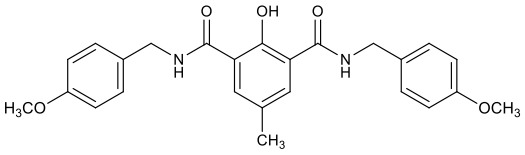
Compound 43



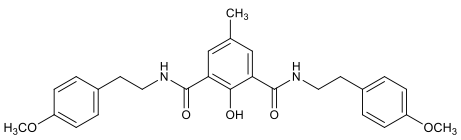
Compound 44



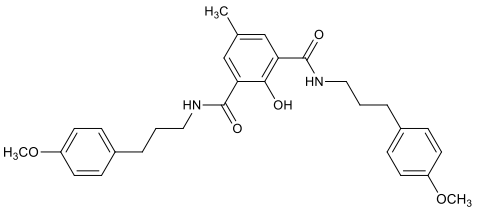
Compound 45



Compound 46



Compound 47



Compound 48