Future appeal of comparative studies on putative binding sites of HIV-1 virus-encoded proteolytic enzyme inhibitor of different Food and Drug Administration-approved compounds

Zainularifeen Abduljaleel1,2, Mohammad Athar1,2, Faisal A. Al-Allaf, Saied Al-Dehlawi1, Sami Melebari3, Waseem El-Huneidi4

1Department of Medical Genetics, Faculty of Medicine, Umm Al-Qura University, Makkah, Kingdom of Saudi Arabia
2Science and Technology Unit, Umm Al-Qura University, Makkah, Kingdom of Saudi Arabia
3Molecular Diagnostics Unit, Department of Molecular Biology, the Regional Laboratory, Ministry of Health (MOH), Makkah, Kingdom of Saudi Arabia
4Department of Basic Medical Sciences, College of Medicine, University of Sharjah, Sharjah, United Arab Emirates

Abstract

Introduction: Human immunodeficiency virus (HIV) protease enzyme is one of the most promising therapeutic targets for acquired immunodeficiency syndrome (AIDS) treatment. Due to mutation of the virus, there is always a room for new agents.

Material and methods: The aim of in silico molecular docking study was to analyze and compare the binding mode of seven Food and Drug Administration (FDA)-approved HIV protease enzyme inhibitors, and to understand their structural requirements to inhibit an enzyme by using Schrodinger model as well as to evaluate a free energy of binding of these inhibitors with an enzyme.

Results: The binding mode analysis showed that the active site was present at the interface of two chains A and B of the enzyme and the crucial amino acid remained responsible for the binding of inhibitors to the HIV-1 protease, which could help to classify the inhibitors as better drug targets. Results of this comparative binding mode analysis of seven FDA-approved drugs could be potential and useful for designing of a new effective inhibitor of HIV-1 protease. Out of seven inhibitors drugs, only two drugs present the best inhibition. HIV protease-nelfinavir complex with PDB: 2Q64 and HIV protease D30N, and R41A double mutant-tipranavir complex in PDB: 1D4S double mutant V82F and I84V, were used as templates for applying the mutations on HIV protease active site. Furthermore, the structure-based computer-assisted search for the comparison of the two inhibitors of HIV protease was completed. On the other hand, tipranavir seems to be a broad specificity inhibitor, as no changes in the bond lengths with the introduction of mutations were observed.

Conclusions: Tipranavir could be targeted more effectively for designing future drug analogues, as it is less vulnerable to mutations. HIV mutants reported in this study could also be used for preliminary identification of specific inhibitors, as drugs that may alter the HIV protease activity for medicinal use.

Key words: therapeutic targets, molecular docking, binding site comparison, enzyme inhibitors.

Address for correspondence: Dr. Zainularifeen Abduljaleel, PhD, Department of Medical Genetics, Faculty of Medicine, Umm Al-Qura University, P.O. Box 715, Makkah 21955, Kingdom of Saudi Arabia, phone: + 966 570552361, e-mail: zainulbio@gmail.com, zaabduljaleel@uqu.edu.sa

Article history:
Received: 08.04.2019
Received in revised form: 06.06.2019
Accepted: 07.07.2019
Available online: 19.06.2020
Introduction

The Joint United Nations Programme on HIV/AIDS (UNAIDS) is determined to end the public health threat of global human immunodeficiency virus (HIV) epidemic by 2030. To achieve this goal, an estimated budget of US$ 26.2 billion will be required for the HIV response in 2020 (http://www.unaids.org/en/goals/unaidsstrategy), which may gradually reduce to US$ 22.3 billion by 2030 (http://www.unaids.org/en/goals/unaidsstrategy).

HIV affects cells in the immune system and central nervous system. The main type of cells, which are infected by HIV are T helper lymphocytes. These cells play a crucial role in the immune system by coordinating the actions of other immune system cells. A large reduction in the number of T helper cells seriously weakens the immune system.

There are certain grades of HIV (Figure 1), with each grade having its own symptoms. Knowing the symptoms of HIV while being exposed to the virus could accelerate the treatment and extend one's life. The primary stage of HIV infection may be asymptomatic for up to 10 years [1]. Most people in the early stages of HIV experience these symptoms, but in more severe ways [2]. Sometimes, these symptoms can start in the first month of being exposed to the virus and can last from a few days to several weeks. Clinical latency is the second stage, where the virus is present in a body, causing no symptoms. The virus grows at very slow rate but is still active in the body. The amount of time HIV patients are in the latency stage varies depending on their individual immune system and treatments they receive. The third stage of HIV virus is acquired immunodeficiency syndrome (AIDS) [3]. There are specific symptoms indicating that HIV has turned into AIDS.

Figure 1. Stages of HIV. A) Acute infection: during this stage, large amount of virus is produced in the body. Many people develop flu or influenza-like symptoms in this stage (www.theleak.co.za); B) Clinical latency: during this stage, HIV virus multiplies at very slow rate, but is still active (www.webmd.com); C) This stage is considered as AIDS when CD4 count falls below 200/ mm$^3$ (www.redorbit.com)
as people with AIDS present greatly weakened immune systems. However, some of the symptoms described above can indicate other diseases.

World Health Organization (WHO) global report on HIV drug resistance suggests that in several low- and middle-income countries (LMICs), over one in ten HIV-infected patients initiating antiretroviral therapy (ART) have pretreatment HIV drug resistance to efavirenz and nevirapine [4]. This resistance is associated with poor virological outcomes, impaired immune recovery, reduced durability of non-nucleoside reverse-transcriptase inhibitor (NNRTI)-based regimens, and increased mortality in adults and children.

Intermolecular interactions produced by functional residues of the protein are investigated with the use of protein homology tools to identify the functional residues, creating single, double, and triple mutations using different bioinformatics approach, and observing the changes brought by docking [5]. Despite the structural similarities, HIV-1 and HIV-2 proteases demonstrate dramatic disparities in susceptibility to HIV-1 protease inhibitors. HIV-2 is at least partially resistant to the majority of HIV-1 drugs approved by the U.S. Food and Drug Administration (FDA); only lopinavir (LPV), darunavir (DRV), and saquinavir (SQV) are active at clinically useful concentrations [6, 7]. Six substrates of eight FDA-approved drugs are shown in Table 1.

The motion in the flaps of HIV-1 protease plays a critical role in substrate binding and catalysis. Because of the intrinsic flexibility of HIV-1 protease, the free enzyme contains conformations of various flaps, usually termed as semi-open, open, and closed [8]. The structures show a semi-open flap conformation; however, the flap must open for binding to a natural substrate. The protonation states of two catalytically important aspartates, Asp-25 and Asp-251, have been investigated to understand the catalytic mechanism and inhibitor recognition [9]. Because hydrogen atoms of Asp-25 and Asp-251 can contribute to an important hydrogen bond, protonation states of the residues need to be considered in structure-based drug design. In addition, molecular dynamics (MD) simulations to investigate the protease enzyme, MD and multi-scale quantum mechanics and molecular mechanics (QM and MM) methods to investigate protonation states of catalytic key targeted residues, and Brownian dynamics (BD) simulations to model large-scale conformational changes in HIV-1 protease for substrate recognition need to be considered.

Material and methods

Protein structure validation

The stereochemistry of protein structure quality, and the verification of mainly both side and main chain were examined by employing WHAT_CHECK [10], PROCHECK [11], ERRAT [12], VERYFY_3D [13], and PROVE [14]. These servers help to verify the stereo-chemical quality of a protein structure by analyzing a residue-by-residue geometry and overall structure geometry. Many stereochemical parame-
ers of the residues in the protein model can be found, with the statistics of non-bonded interactions between different atom types and plots analyses, calculated by a comparison with statistics from highly refined structures. Additionally, the compatibility of an atomic model (3D) with its own amino acid sequence (1D) can be determined by assigning a structural class based on its location and environment (alpha, beta, loop, polar, non-polar, etc.), volumes of atoms in macromolecules using an algorithm, which treats the atoms like hard spheres and calculates a statistical Z-score deviation for the model from highly resolved (2.0 Å) can be determined.

Ligand structures and pharmacophore preparation

Seven of FDA-approved compounds to analyze the best active inhibitors were selected, including tipranavir (TPV), atazanavir (ATV), darunavir (DRV), fosamprenavir (FPV), ritonavir (RTV), nelfinavir (NFV), and saquinavir (SQV), which were downloaded from PubChem (http://pubchem.ncbi.nlm.nih.gov/) in the SDF file format. All compounds were prepared by using LigX module in MOE (molecular operation environment) 2009. LigX has a strong property of conducting interactive ligands modification and energy minimization in the active site of a flexible receptor. The verified drug ligand step will isolate from all such molecules. The chosen drug ligands were identified to be treated as though they were a part of receptor.

Binding site prediction

The active site was predicted by using MOE. An alpha shape algorithm was used to determine potential active sites in 3D protein structures. We analyzed active site of a receptor, protein surface calculations, and molecular docking used to search the favorable binding configurations between one or more small flexible ligands and protein target. That are most likely to contribute to a tight protein–ligand binding. Typically, scoring functions emphasize favorable hydrophobic, ionic, and hydrogen bond contacts. In our study, we detected a candidate for protein-ligand binding sites using a fast-geometric algorithm, based on Edelsbrunner’s alpha shapes. Protease enzyme binding site on a macromolecular structure was ranked according to its accessible hydrophobic contact surface and active sites analysis to identify polar, hydrophobic, acidic, and basic residues. Visualize solvent of exposed ligand atoms and residues in close contact with ligand atoms as well as sidechain and backbone acceptor were evaluated.

Molecular docking and molecular dynamics simulations

These selected drug molecules were docked into the active site of protease enzyme (HIV-1) protein docking server (http://www.dockingserver.com/web) [15]. This server was based on AutoDock version 4, and compared with MOE and
HIV & AIDS Review 2020/Volume 19/Number 2

SYBYL (SYBYL 8.1, Tripos International, St. Louis, USA). This application was analyzed for accurate ligand geometry optimization, energy minimization, charge calculation, docking calculation, and protein-ligand complex representation. It showed more accurate partial charge calculation and consequently, more accurate docking was achieved by using quantum chemical methods. MD simulations were performed in the NPT ensemble, with a target temperature of 300 K and target pressure of 1 atm. A Berendsen coupling time of 0.2 ps was used to maintain the right temperature.

Table 1. The chemical structures of Food and Drug Administration (FDA)-approved drugs

<table>
<thead>
<tr>
<th>Compound Name</th>
<th>Molecular Structure &amp; Accession No</th>
<th>Molecular Formula</th>
<th>Chemical Formula</th>
<th>Total bond energy (ev)</th>
<th>Heat of formation (ev)</th>
<th>Dipole moment (D)</th>
<th>Molecular weight (MW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tipranavir (TPV)</td>
<td>DB00932</td>
<td>C_{6}H_{8}F_{2}N_{2}O_{5}S</td>
<td>2-Pyrindinesulfonylamide, N-[[3-[[1R]-1-[(6R)-3,6-dihydropyridine-2-oxo-6-(2-phenylethyl)-6-propyl-2H-pyran-3-yl]propyl][phenyl]-5-(trifluoromethyl)]</td>
<td>349.4878</td>
<td>-16.6186</td>
<td>1.7931</td>
<td>602.7</td>
</tr>
<tr>
<td>Atazanavir (ATV)</td>
<td>DB01072</td>
<td>C_{28}H_{31}N_{6}O_{17}</td>
<td>(3S,8S,9S,12S)-3,12-Bis[[1,1-7 dimethyllethy]l-8-hydroxy-4,11-dioxo-9-phenylmethyl-][4-[(2 pyridinyl)methyl]-6,2,5,6,10,13 pentazatetracenediac acid dimethyl ester, sulfate (1:1).</td>
<td>460.8513</td>
<td>-13.6406</td>
<td>2.9911</td>
<td>704.9</td>
</tr>
<tr>
<td>Darunavir (DRV)</td>
<td>DB01264</td>
<td>C_{15}H_{14}F_{2}N_{4}O_{4}S</td>
<td>[[1S,2R]-3-[[[(4-aminophenyl)sulfonyl][2- methylpropyl]amino]-2-hydroxy- [phenyl(methyl)propyl]- carbamic acid[3R,3a,5a,R]]- hexahydrorufan-3-yl ester monoethanolate.</td>
<td>331.4491</td>
<td>-11.6575</td>
<td>5.0302</td>
<td>593.75</td>
</tr>
<tr>
<td>Fosamprenavir (FPV)</td>
<td>DB01319</td>
<td>C_{15}H_{14}F_{2}N_{4}O_{4}S</td>
<td>fosamprenavir calcium is (3S)-tetrahydrofurran-3-yl (1S,2R)-3-[[[(4-aminophenyl)sulfonyl]l(isobutyl)amino]- [1-benzyl-2-(phosphonoxy) propyl]carbonate monocalcium salt</td>
<td>331.316</td>
<td>-20.1940</td>
<td>6.8842</td>
<td>623.7</td>
</tr>
<tr>
<td>Ritonavir (RTV)</td>
<td>DB00503</td>
<td>C_{15}H_{14}F_{2}N_{4}O_{4}S_{2}</td>
<td>10-Hydroxy-2-methyl-(1- methylethyl)-1-[2-[1(1-methyllethyl)-4 thiiazolyl]-3,6-dioxo-8,11 bis(phenylmethyl)-2,4,7,12-tetraazatricane-13-oxic acid, 5- thiactzymlethy]l ester, [5S (3R,3a,5a,R,10R*,11R*)].</td>
<td>440.074</td>
<td>-8.7487</td>
<td>5.4764</td>
<td>720.95</td>
</tr>
<tr>
<td>Nelfinavir (NFV)</td>
<td>DB00220</td>
<td>C_{16}H_{14}F_{4}N_{3}O_{3}</td>
<td>[3S-(22S*, 35*)]-N-(1,1 dimethyllethylidene)hydroxy-2-[2- hydroxy-3-[[3-hydroxy-2 methyl(2- benzyl)amino]-4 (phenylthio)R-(2)]-3-isochinoline carbosamide mono-methanesulphonate (salt)</td>
<td>375.1660</td>
<td>-7.9088</td>
<td>6.3703</td>
<td>663.90</td>
</tr>
<tr>
<td>Saquinavir (SQV)</td>
<td>DB01232</td>
<td>C_{16}H_{14}F_{4}N_{3}O_{3}</td>
<td>N- tert-butyl-decalohydro-2-[2(R)-hydroxy-4-phenyl-3(S)]-N-(2- quinolylcarbonyl)L- L-asparaginyl[jamino]benzyl(4-adamantane-3(S)-carboxamide</td>
<td>445.8475</td>
<td>-8.3205</td>
<td>9.8985</td>
<td>670.86</td>
</tr>
</tbody>
</table>
and pressure of the system [16]. The system was gradually heated from 10 to 300 K over 20 ps. Initial velocities were assigned from a Maxwellian distribution at the starting temperature. The system was equilibrated at 300 K for 100 ps, followed by a production phase between 100 and 500 ps.

Results

The replication cycle of HIV (Figure 2) that schematically represents all principle steps of viral life cycle, starting from cell entry to maturation of new infectious viral particles were reported. The Merck laboratories [17] were the first to obtain a crystal structure of HIV protease; however, more accurate structure was reported by Kent et al. The structure of protease (Figure 3) is a homodimer and consists of subunits with 99 amino acid residues [18, 19] that operate as a homodimer with only one active site, which is C2-symmetric in free form. Moreover, new HIV treatment drugs can reduce drug resistance and one important factor of adherence may affect the impact of HIV treatment outcome [20].

Binding site prediction and molecular docking

The crystallographic structure of protease enzyme (PDB ID: 2Q64) was used for docking calculations of the FDA-selected compounds inhibitors that were targeted with protease enzyme, which selected seven compounds (TPV, ATV,
Comparative studies on binding sites of proteolytic enzyme

DRV, FPV, RTV, NFV, SQV). Molecular operating environment, version 2010.10 (MOE 2010.10) software was selected for the calculations of possible ligand-binding site and was detected through application of a Site Finder of MOE, with a connection distance parameter set to 1.9 Å. The output results suggested that the presence of two homopiperazine rings, and the relationship between the homopiperazine rings and the side structures at X position were important for the inhibition of protease [21]. The binding site of structures was identified as the absence of crystal structure of the ligand, and a blind docking was performed with the 2Q64 protein structure (Figure 3).

Two main interacting binding sites were identified, including one near the entrance of central beta-barrel of the enzyme and the other, near the N-terminal of the protein. The MOE docking protocol and parameters setting could be used to search for binding modes of all seven FDA compounds accordingly. Using the same docking protocol, all initial hits were docked into the binding pocket of 2Q64 protease. A maximum of 10-20 conformations allowed to be saved for each ligand using the default parameters of MOE. The top ranked conformations of all docked compounds were saved in a separate database. Based on docking score, top ranked compounds were selected for further evaluation. The resulted binding interactions between these 100-300 hits and the protein were visually observed using LigPlot implemented in MOE, and those molecules, which revealed significant interactions with the most important binding pocket residues (Asp30, Asp29, Asp25, Asp30, Ile50) of a protease site were selected as promising hits. Among these top compounds, four showed a significant interaction with important residues of target protein.

These seven compounds were further subjected to the binding energy and binding affinity calculation. In the free enzyme state, the flaps assumed a semi-open conformation [18], and with a ligand in the active site, they assumed a closed conformation [22]. It has been reported that a network of weakly polar interactions between the flaps...
Figure 3. Structure of human immunodeficiency virus-1 (HIV-1) protease enzyme. A) The 2Q64 structure is presented using the helical option in Tripos Sybyl [16], with chain $A$ and $B$ shown in violet color. This highlights the secondary structure of protein, which is largely $\beta$-sheet. The binding sites range are between Asp25 (dimer interface in terminal $\beta$-sheet) and 2$\beta$-hairpin flaps Glu48 and Ile 50, and are shown as molecular surface in cloudy white color. B) The same structure is shown, this time of targeted inhibited region as shown in molecular surface atom color. C) The terminal targeted inhibited drug molecule TPV overview is shown in atom color. D) The active site and motion orientation of 1.4 Å with complex of TPV in chain A. E) Other chain $B$ orientation of 1.4 Å complex TPV. F) The theoretical B values of protease enzyme were analyzed and determined by CCP4 program as shown in the Wilson plot. The graph shows R-factor (blue) and free R-factor (green) for cycles of refinement. G) The cross correlation of poorly defined loop regions, weak RS correlation, and excessive B-factor analysis based on the residues.
keeps them in a semi-open conformation [18]. The docked structures with inhibitors of six compounds presented binding energy in ranges as follows: atazanavir MMFF94 binding energy minimization from 201.818 kcal/mol to 79.302 kcal/mol, darunavir from 22.674 kcal/mol to –27.649 kcal/mol, fosamprenavir from 116.657 kcal/mol to –225.363 kcal/mol, nelfinavir from 256.785 kcal/mol to 121.261 kcal/mol, ritonavir from 64.125 kcal/mol to –52.031 kcal/mol, and saquinavir from 243.789 kcal/mol to –138.872 kcal/mol. It was observed that the 3D structure exhibited high binding energy of atazanavir, ranging from 201.818 kcal/mol to 79.302 kcal/mol, and the 3D structure exhibited binding energy of ritonavir, ranging from 64.125 kcal/mol to –52.031 kcal/mol. The low binding and intermolecular free energies of darunavir ranged from 22.674 kcal/mol to –27.649 kcal/mol. The protease active site cavity included Arg8, Leu23, Asp25, Gly27, Ala28, Asp29, Asp30, Val32, Lys45, Ile47, Met46, Gly48, Gly49, Ile50, Phe53, Leu76, Thr80, Pro81, Val82, and Ile84 residues. The majority of residues forming the substrate binding site were hydrophobic; the exceptions were the catalytic Asp25 and Asp29, which formed hydrogen bonds with peptide main chain groups, and Arg8, Asp30, and Lys45, which could interact with polar side chains or distal main chain groups in longer peptides [23]. The result of structure showed an interaction with the Asp30, Asp29, Asp25, Asp30, and Ile50 residues that were present in the binding site (Figure 3).

Furthermore, the Lamarckian genetic algorithm method was applied for docking calculations using default parameters. AutoDock uses a semi empirical free energy force field to evaluate conformations during docking simulations. The optimized orientations represent possible binding modes of the ligand within the site. The first three terms are Van der Waals (vdW), hydrogen bonding, and electrostatics. The term is used for rotation, translation, desolvation upon binding, and for hydrophobic effect. The docking result, 10 solutions were assembled into groups with RMS deviations lower than 1.0Å. The clusters were ranked by the lowest energy representative of each cluster (Figure 4).

**Molecular dynamics and simulation**

MD simulation has been used to study the movement of flap region of the HIV-1 protease with a ligand [24]. The flaps initially opened to all-atom RMSD of 25 Å within

---

**Figure 4.** The human immunodeficiency virus-1 (HIV-1) protease structure in complex with different Food and Drug Administration (FDA)-approved inhibitors. Lowest energy docked structure for six compounds with HIV-1 protease. A closer view showing the binding mode of compound 20 inside the PDB: 2Q64 protease active site. Some atoms are deleted for clarity. The pharmacophore prediction is colored in yellow. The inhibitors residues Asp30, Asp29, Asp25, Asp30, and Ile50 are presented as colored sticks. The 3D presentation for computational results as showed in the structure form A) atazanavir, B) darunavir, C) fosamprenavir, D) nelfinavir, E) ritonavir, and F) saquinavir.
200 ps, became completely open at the end of a 10 ns simulation. In this study, the flaps opened and moved away from X-ray structure PDB: 2Q64 from 0.54 Å at 0.1 ps to 3.30 Å RMSD at 10 ps. The flap RMSD was calculated from residue 40 to 60 of each protein chain. These movements, after 0.1 ps of simulation, were inversely correlated with the quality of binding energy prediction. The correlation coefficient significantly decreased from 0.87 at 0.1 ps to 0.74 at 10 ps, as the all-atom flap RMSD increased from 0.54 to 3.30 at 0.1 and 10 ps. Complementarity between the ligand and the binding site is the basic concept behind ligand binding, which is shown as a steric complementarity. The shape of the ligand is mirrored in the shape of the binding site, allowing molecular interactions between two molecules (http://www.fst.rdg.ac.uk/courses/fs916/lect5/lect5.htm). MD simulations permit a re-arrangement of protease side chain, especially on the active site surface, which improves the interacting surface complementarities of the complex; after 0.1 ps, the time scale that produced the best correlation coefficient. The average all-atom RMSD of the complex was only 0.35 Å, while the average all-atom RMSD of the flap region was 0.54 Å.

Conclusions

Evaluating the interaction energy for the specific ligand-receptor complex under study usually identifies the best ligands. From the results of this study it can be observed that among the studied ligands, only TPV and NFV can be considered as promising drugs due to their good affinity with protease enzyme and fine interaction energy (> 99.9%), when compared to other ligands. It can be concluded that molecular docking helps in drug designing and provides a good understanding of the mechanism of interaction of the drug and target protein.

Acknowledgment

The authors would like to thank the Makkah Regional Laboratories and the Umm Al-Qura University, Kingdom of Saudi Arabia for the continuous support of the research.

Conflict of interest

The authors declare no conflict of interest with respect to the research, authorship, and/or publication of this article.

References

20. Jacob SA, Jacob DG, Jugulete G. Improving the adherence to antiretroviral therapy, a difficult but necessary task for a successful HIV treatment-clinical points of view and practical considerations. Front Pharmacol 2017; 8: 831.