Insilico Analysis of Zinc Database to Discover New Potent HIV1 Protease Inhibitors

Mohammad Hossein Heidari1, Neda Adibpour2, Sajjad Gharaghani3 and Mohaddesse Rahimi1

1. Department of Pharmacy, Faculty of Pharmacy, Zanjan University of Medical Sciences, Zanjan, Iran
2. Department of Medicinal Chemistry, Faculty of Pharmacy, Zanjan University of Medical Sciences, Zanjan, Iran
3. Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran

Abstract: AIDS (Acquired immune deficiency syndrome) is the leading cause of death in modern societies. HIV (human immunodeficiency virus) by binding to CD4 receptors on the surface of lymphocytes, T helper and by preventing the throughput of lymphocytes is due to the severe weakening of the immune system and causes AIDS. One of the most critical enzymes of HIV is named PR (Protease) that converts polyprotein precursor to HIV proteins and enzymes. In the meantime, researchers are also struggling and after extensive researches they succeeded in finding the structure and function of enzymes that are critical for HIV, specially Protease. Scientists reported PR enzyme as a desirable target to achieve therapeutic goals, because by inhibition of PR enzyme virus lifecycle will be disrupted. To now ten effective drugs are approved by the FDA (Food and Drug Administration) for HIV protease inhibition. Today research for designing the inhibitors of HIV PR is increasing. In this respect we are particularly trying in this article to find new compounds with high potential protease inhibition. We finally managed to identify several effective compounds and report them. The methods are based on in silico skills. HIV protease structure was prepared from RCSB protein data bank (4rvj ID protein with amprenavir-an enzyme inhibitor-complex) and the structure of all ten drugs that are approved by the FDA were prepared from PubChem and similar compounds were found by using the zinc online server. After similar compounds passed lipinsky filter, 86 compounds were extracted. After these processes compounds were docked on protease-amiprenavir complex by PyRx software and then analysis of medicinal chemistry was done by using ligplot software. We analyzed the interactions between compounds and residues of the active site and we compared these interactions with approved drugs, and then we reported effective compounds. From all similar compounds, those showing best binding affinities are mentioned below: 87402021, 68688069, 90305202, 89150243, 59959863, 53682869, 54691125. After medicinal chemistry analysis we guess that these compounds are potent to being a drug. Medicinal chemistry analysis was done by using ligplot software. We used in silico skills and validated processes to find these compounds and we guess that these compounds inhibit HIV1 protease and are useful in HIV1 drug design process.

Key words: HIV, PR, medicinal chemistry.

1. Introduction

AIDS (acquired immune deficiency syndrome) is one of the major causes of death in many communities that occurs by the HIV. HIV entered to lymphocytes by binding to CD4 receptors (on T helpers) and entry into the cells results in a dysfunction of the immune system. Fifteen proteins and an mRNA are responsible for filling in and completing all phases of the cell cycle of HIV. IN (integrase) PR (protease) and RT (reverse transcriptase) are the three important enzymes in the virus life cycle. HIV after binding to the CD4 receptors by using RT enzyme converts its single-stranded RNA into double-stranded DNA then, integrates it in the host genome [1]. After that with any replication of host genome, HIV mRNA will be produced by cell, inadvertently. After completing transcription and translation, it is time to produce viral proteins that depends on PR enzyme function. PR can make the active and efficient proteins of the virus from the primary polyprotein. Thirty years ago, several therapies to prevent the progression of HIV and combating it has emerged that among these ways,
PR enzyme inhibition was one of the most efficient ways. In the meantime many inhibitors are in three phases of clinical trial, but there are only ten drugs that have been approved by the FDA (Food and Drug Administration) [2]. Today, researchers are trying to propose new compounds for the prevention of HIV. Because one of the most important methods in HIV treatment is “highly active antiretroviral therapy” (ART) that is based on the principles of multi-drug combination to prevent the occurrence of viral resistance, there is not any way to achieve ART just only to synthesize and reveal new and different categorized pharmaceutical compounds [3]. In this respect researchers in the computational drug design field are struggling to find new compounds by using CADD (computer aided drug design) method that it has an important role in the ART procedures and plays its role by elimination of additional costs and saving time and money [4]. The goal of the present article is to use CADD methods for achieving new HIV protease inhibitors.

1.1 HIV1 Protease Structure and Function

Fortunately, the protease is a well-known enzyme and among the processes of developing the knowledge of this enzyme, Insilico methods had a great impact. Protease has two chains and protease active site contains amino acids Asp, Thr, Gly which are 25 to 27, respectively.

Reaction between substrate and active side residues depends on Asp25 and Asp25′. Unfortunately, accurate performance of the four other amino acids in the active site is not known very well, but it is believed that the strong bond between Thr26 and Thr26′ preserves the formation of the active site, as well as two amino acids Gly27 and Gly27′ responsible for keeping the substrate in the active site for attack of Asp to the amid group of substrate.

It is also observed that the effects of various inhibitors are mediated by contacting the electrostatic interactions with Asp30 [5].

2. Materials and Methods

HIV1 protease preparation.

HIV1 X-ray crystallography complex with Amprenavir was extracted from RCSB protein data bank (4rvj) [6]. We set 4rvj as macromolecule. Autodock force field was applied by PyRx program.

2.1 Molecular Dynamic Simulation

For the purpose of achieving best estimations it was decided to relax the enzyme (HIV1 protease) by an MD test. MD test was applied by gromax software [7]. The selected forcefield for all atoms of the protein was “all atom OPLS”. The box has been filled by water molecules. The temperature of protein was fluctuating between 298-300 K in 100 ps and rapidly reached the target value (300 K). RMSD plot of protein explains the equilibration of enzyme as shown in Fig. 1.

2.2 Validation of Docking Program

For achieving the best pose and binding energy estimation from docking program it was decided to set a redock method before starting the docking process. Asp25, Asp25′, Thr26, Thr26′ were selected as flexible residues of active site. Indinavir (the FDA applied drug HIV1 protease inhibitor) was docked in active site of the enzyme. The docking pose was compared to X-ray pose of indinavir in complex with HIV1 PR. The excellent RMSD (0.57, Fig. 2) was demonstrated the best estimation of PyRx for our enzyme.

2.3 Data Set

To now ten drugs have been approved by the FDA, as mentioned above. The structure of all ten drugs was extracted from PubChem server [8] and then similarity search was done by zinc similarity search engine [9]. We defined a query for all ten drugs one by one. Needed chemistry features such as rotatable bounds and dihedral bounds were applied on ligands by autodock tools 1.5.6 [10]. All ligands were converted
Insilico Analysis of Zinc Database to Discover New Potent HIV1 Protease Inhibitors

Fig. 1  RMSD plot of protein in 100 ps.

Fig. 2  X-ray crystallography pose of indinavir colored orange and docking pose colored yellow. Program is validated by RMSD < 2 (in this case RMSD is 0.57).

After all these processes 86 compounds were extracted but three drugs did not show any similarity in zinc compounds. After optimization and energy minimization of these compounds we docked all of them on HIV1 protease (4rvj).

2.4 Docking Process

Docking process of all 86 compounds was carried out by PyRx (autodockvina 4.2). PyRx software is a validated useful docking application that is based on autodockvina with an improved speed of docking, multithreading and with a new scoring function [11]. The resolution was set on 0.375 angstrom and the grid box was set on active site.

The validation of application was done by redock method. Redock showed a good accommodation between X-ray crystallography pose and docking pose.

Docking of these 86 compounds on 4rvj HIV1 protease showed very good results that are mentioned in Table 1.

As seen in Table 1, six compounds show good docking score and we mentioned them below:
Table 1  Docking score (D.S) of 86 similar compounds with seven approved drugs.

| Compound name/number | D.S  | Compound name/number | D.S  |
|----------------------|------|----------------------|------|
| Amprenavir           | -5.9 | 57611570             | -6.5 |
| 18667185             | -5.7 | 58800086             | -6.8 |
| 21917287             | -5.9 | 59116036             | -6.1 |
| 22105710             | -5.9 | 59116038             | -6.1 |
| 57288314             | -5.5 | 59116040             | -5.9 |
| 57294950             | -6.8 | 59116052             | -6.2 |
| 66905729             | -5.4 | 59897678             | -7.0 |
| 68461188             | -2.4 | 59959863             | -11.1|
| 68729841             | -5.9 | 59964282             | -6.1 |
| 87402021             | -5.5 | 71027128             | -6.4 |
| 87508645             | -5.4 | 91030358             | -6.7 |
| 18667184             | -5.8 | 91159384             | -5.4 |
| Atazanavir           | -6.9 | saquinavir            | -7.8 |
| 67994614             | -7.1 | 22860385             | -7.2 |
| 68687674             | -6.0 | 53682869             | -8.2 |
| 68687678             | -6.1 | 53780913             | -6.9 |
| 68688067             | -7.5 | 53954227             | -8.1 |
| 70311191             | -6.7 | 53983607             | -5.7 |
| Darunavir            | -6.8 | 54168821             | -7.0 |
| 90305205             | -6.5 | 67707417             | -6.1 |
| Fosamprenavir        | -6.2 | 69412349             | -7.0 |
| 89150243             | -6.3 | 69778445             | -7.2 |
| Nelfinavir           | -6.2 | 69829058             | -6.9 |
| 20624766             | -6.3 | 69829652             | -7.4 |
| 20624778             | -6.6 | 69829700             | -6.3 |
| 53686546             | -6.9 | 70067258             | -6.3 |
| 54239162             | -6.9 | 70067260             | -6.6 |
| 71014160             | -8.0 | 54691092             | -7.3 |
| 71042502             | -8.0 | 54691125             | -7.6 |
| 71467772             | -7.7 | 54691234             | -6.7 |
| 73962838             | -7.3 | 54696348             | -6.9 |
| 73974830             | -7.2 | 54711615             | -6.5 |
| 89054597             | -5.6 | 54714127             | -6.5 |
| 110165102            | -7.2 | 59131322             | -7.5 |
| Tipranavir           | -6.2 | 59131752             | -7.2 |
| 24826330             | -6.9 | 59928596             | -6.8 |
| 54685540             | -7.0 | 67674706             | -6.3 |
| 54685677             | -7.0 | 67697170             | -6.9 |
| 54689663             | -7.3 | 70079161             | -6.5 |
| 54691022             | -7.1 | 70080839             | -6.9 |
| 54691023             | -7.1 | 70081353             | -6.4 |
| 54691024             | -6.7 | 70081574             | -6.6 |
| 54691086             | -7.4 | 70930613             | -7.4 |
| 54691087             | -7.4 | 71211841             | -7.4 |
| 54691088             | -7.3 | 71226286             | -6.9 |
| 54691091             | -7.5 | 90726916             | -7.0 |
| 91563706             | -7.4 |                     |      |

Medicinal chemistry analysis was done by using ligplot software [12].

3. Results and Discussion

In this study, we received ten drugs from PubChem server applied for protease enzyme inhibition by the FDA and we relaxed enzyme for achieving more reliable data. Then we set a query with lipinsky filter and found similar compounds from zinc.docking.org server. Finally we found 86 similar compounds. All of them were docked on 4rvj. Six of these compounds demonstrated good results. As seen in Table 1, six compounds show good docking score and we mentioned them below: 53682869, 54691125, 59959863, 68688069, 89150243, 90305205.

Medicinal chemistry analysis was done by using ligplot software (Fig. 3).

As seen in the case of inhibitory drugs, at least one of the amino acids Asp25, Asp29, Asp30 and/or Gly27 has been mediated in drug interaction plot. This criterion is also seen, exactly in about six compounds but it should be noted that among many drugs and six screened compounds two amino acids Arg87 and Arg45 can link compounds to PR inhibition. We are guessing that a new mechanism of PR inhibition is to contact an electrostatic interaction with at least one of these two amino acids.

4. Conclusions

In this study, we used ten drugs and zinc.docking.org server for finding new similar compounds. All of 86 compounds were docked on HIV1 protease enzyme (4rvj pdb ID) and medicinal chemistry analysis was done and six compounds were extracted from those 86 datasets. Because of their excellent results these six compounds must be synthesized for further experimental tests. We hope that these compounds can inhibit HIV1 Protease as their results have shown.
Fig. 3 Hydrogen bond and hydrophobic interactions between compounds/drugs and active site residues.
References

[1] Barré-Sinoussi, F., Ross, A. L., and Delfraissy, J. F. 2013. “Past, Present and Future: 30 Years of HIV Research.” *Nature Reviews Microbiology* 11 (12): 877-83.

[2] Brik, A., and Wong, C.-H. 2003. “HIV-1 Protease: Mechanism and Drug Discovery.” *Organic & Biomolecular Chemistry* 1 (1): 5-14.

[3] Murri, R. 2005. “Highly Active Antiretroviral Therapy.” *BMJ* 330 (7493): 681-2.

[4] Kapetanovic, I. M. 2008. “Computer-Aided Drug Discovery and Development (CADD): Insilico-Chemico-Biological Approach.” *Chem Biol Interact* 171 (2): 165-76.

[5] Mager, P. 2001. “The Active Site of HIV-1 Protease.” *Med. Res. Rev.* 21: 348-53.

[6] Rose, P., Prlić, A., BiCi, Bluhm, W. F., Christie, C. H., Dutta, S., Green, R. K., Goodsell, D. S., Westbrook, J. D., Woo, J., Young, J., Zardecki, C., Berman, H. M., Bourne, P. E., and Burley, S. K. 2014. “The RCSB Protein Data Bank: Views of Structural Biology for Basic and Applied Research and Education.” *Nucleic Acids Res.* 43: D345-56.

[7] Abraham, M., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. 2015. “GROMACS: High Performance Molecular Simulations through Multi-level Parallelism from Laptops to Supercomputers.” *SoftwareX* 1-2: 19-25.

[8] Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., and Shoemaker, B. A. 2015. “PubChem Substance and Compound Database.” *Nucleic Acids Research* 44: D1202-13.

[9] Sterling, T., and Irwin, J. 2015. “Zinc 15—Ligand Discovery for Everyone.” *J. Chem. Inf. Model* 55 (11): 2324-37.

[10] Morris, G., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. 2009. “AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility.” *J. Comput. Chem.* 30 (16): 2785-91.

[11] Trott, O., and Olson, A. J. 2009. “Software News and Update AutoDockVina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading.” *Journal of Computational Chemistry* 31 (2): 455-61.

[12] Wallace, A. C., Laskowski, R. A., and Thornton, J. M. 1995. “LIGPLOT: A Program to Generate Schematic Diagrams of Protein-Ligand Interactions.” *Protein Engineering* 8 (2): 127-34.