In silico drug designing of protease inhibitors to find the potential drug candidate for HIV1

Alka Dwivedi[1, *], Vijay Laxmi Saxena[2]

[1]National Bioinformatics Infrastructure Facility Center of (DBT), Ministry of Science and Technology (Govt. of India), D.G (PG) College civil lines Kanpur, INDIA

[2]Department of Zoology (HOD), (Coordinator) National Bioinformatics Infrastructure Facility Center of (DBT), Ministry of Science and Technology (Govt. of India), D.G (PG) College civil lines Kanpur, INDIA

Email address:

Alkabioinfo964@gmail.com(A. Dwivedi), vijaykanpur@rediffmail.com(V. L. Saxena)

Abstract: Acquired immunodeficiency syndrome (AIDS)was first reported by the us centre of disease (CDC),a few years later it was found that's a retrovirus called human immune deficiency virus (HIV) and this causative agent in AIDS, the study of HIV protease is one of the most important approaches for the therapeutic intervention in HIV infection and their development is regarded as   major success of design, HIV attacks on the CD4+ (T helper cells in human) lymphocyte and these are key component of the body's immune system. The present anti retroviral HIV drugs targets based on three protein reverse transcriptase, protease, integrase in this project work on protease enzyme for block the protein malfunction who is responsible for this activity with the help of computer aided drug designing and the best dynamical and statical parameters like homology modeling ,model verification , binding site identification, docking , according there procedure active   and effective site of the protease is determined and dock with suitable ligand with receptor and calculate the statical values , NVP is the most suitable ligand   and can be use as a inhibit the activity of protease enzyme NVP molecule with a drug likeness property can be considered for in vitro and finally it can acts as a potential lead inhibitor for HIV1.

Keywords: Protease, T-Helper Cells, CO-Receptor, Reverse Transcriptase, Integrase, AIDS, HIV1

1. Introduction

Acquired immunodeficiency syndrome (AIDS) was the first major epidemic caused by a previously unknown pathogen to appear during the 20th century, the period corresponding to the modern development of pharmaceutical sciences. This disease is caused by human immunodeficiency virus type 1 (HIV-1), a member of the family of

Retroviruses. At the onset of the epidemic in the early1980s, no existing drug was known to be useful against AIDS and completely new pharmaceutical agents had to be created.

Although azidothymidine (AZT), the first drug shown to counteract the effects of HIV-1 infection, was previously known as a potential anticancer agent, the rapid progress in the understanding of the structure and life cycle of the virus led to unprecedented development of other drugs targeted to a variety of viral proteins. The retroviral enzymes—reverse transcriptase (RT), integrase (IN), and protease (PR)—were the obvious targets for drug discovery. The first drugs to be identified were inhibitors of RT , which were discovered and Developed long before the Structure of RT itself was solved .However, newer RT-targeted drugs, no nucleoside inhibitors, are being developed bearing the enzyme structure in mind. Even now, only fragmentary structural data have been described for inhibitor and no drugs are available. Retroviral protease, however, was identified early as a potential target, and the discovery and development of its inhibitors are an unqualified success of modern pharmacology and structural biology. Analysis of the nucleotide sequence of the HIV-1 genome (88) led to the discovery that the virus encodes an aspartic protease (HIV-1 PR). Inactivation of HIV-1 PR by either mutation or chemical inhibition leads to the production of immature, noninfectious viral particles thus the function of this enzyme was shown to be essential for proper virion assembly and maturation. It is not surprising, then, that HIV-1 PR was identified over a decade ago as the prime target for structure-assisted (sometimes called "rational") drug design. Drugs against HIV include

antiretroviral. Therapy *.Human Immuno deficiency Virus*.(HIV) is the cause of Acquired Immunodeficiency Syndrome (AIDS). HIV directs the synthesis of several polyproteins, which each consist of several tandemly linked proteins. The maturation of the virus to its infectious form requires that these polyproteins be cleaved to their component proteins. HIV-1 protease. A homodimeric enzyme, is responsible for doing so and is therefore crucial to the virus'sinfectious.capacity [1-7,]

2. Methodology

Database NCBI (National Center for Biotechnology Information) (www.ncbi.nlm.nihgov) [8]. PDB (Protein Data Bank) (www.pdb.org) [9, 10, 11]. Drug Bank (www.drugbank ca.) [12, 13, 14,]. Tools: BLAST (Basic Local Alignment Search Tool),(blast.ncbi.nlm.nih.gov)[15]. Model validation: SAVES (Structural Analysis and Verification Server) (nihserver.mbi.ucla.edu/SAVES/) Model visualization: Chimera, Rasmol, Pymol, discovery studio,Binding site analysis: QsiteFinder, Pocket Finder, (www.modelling.leeds.ac.uk/qsitefinder/)[16-22]
Dockingtool: AutoDock, hex, PATCHDOCK, (hexserver.loria.fr/)Automated Docking Server: Online different type of docking server, The first step in methodology is collection of sequences data from NCBI. Sequence alignment: The protein sequences of protease protein were obtained from NCBI/PDB (http://www.ncbi.nlm.nih.gov/). The homologous structure of protease was identified, which was used as template for the homology modeling. The sequence alignment was done using the online version of Clustal W (http://www.ebi.ac.uk/Tools/msa/clustalw2/) after that the homology modeling of sequence is done then selection of the best model is done with the help of core region and model validation a binding site is also predicted via online tools then go for docking for identification of potential ligand with minimum energy.

3. Results and Discussion

The results analysis base on Sequence of protease Protease [Human immunodeficiency virus 1] GenBank: AAV56605.1 > gi|55705328|gb|AAV56605.1| protease [Human immunodeficiency virus 1] PQITLWQRPLVTIKIGGQLKEALLDTGADDTI LEEMNLPGRWKPKLIGGIGGFIKVRQYDQILVEICGH K XIGTVLVGPTPANIIGRNLLTQIGCTLNF
Expectation Value = 0.001, Search Tool = blast, Mask Low Complexity=yes) via BLASTP .this blast mainly use for protein the results of Computer Aided Drug Designing of Protease Inhibitors to find the potential drug candidate for HIV1 based on different type potential parameters Homology modeling, also known as comparative modeling of protein, refers to constructing an atomic-resolution model of the "*target*" protein from its amino acid sequence and an experimental three-dimensional structure of a related homologous protein. In this project homology modeling completed with Geno3Dit is an automatic web server for protein molecular starting with a query protein sequence. And it's based on Lipinski rules. Models energy (kcal/mol): model 5: -3963.05

Table 1. Model residues in the following four regions

Model number	core	Allowed	generously	disallowed
Model.1	74.7%	24.1%	0.0%	1.3%
Model.2	79.7%	20.3%	0.0%	0.0%
Model.3	78.5%	21.5%	0.0%	0.0%
Model.4	82.3%	17.7%	0.0%	0.0%
Model.5	89.9%	10.1%	0.0%	0.0%
Model.6	81.0%	19.0%	0.0%	0.0%
Model.7	77.2%	22.8%	0.0%	0.0%
Model.8	81.0%	19.0%	0.0%	0.0%
Model.9	86.1%	13.9%	0.0%	0.0%
Model.10	78.5%	21.5%	0.0%	0.0%

Table: 1 explains the core regions with percentage and allowed generously disallowed Parameters potential model highlighted by yellow colour

3.1. Model Validation

Model validation completed with the help of SAVES server. It is type of online web server for model validation and for analyzing protein structure for validity and assessing how correct they are it's based on six programs Results**:** the Model 5 is pass by the SAVES server, and the main value is, verify_3D score: 87.00% of the residues had an average 3D-1D score >0.2 and next value is Errat score: overall quality factor is 91.667

3.2. Mode. Visualization

Selected model analysis based on 2D and 3D structure visualization and the model visualization completed with the help of different type of software's for example: Chimera, Rasmol, and discovery studio

3.3. Binding Site Prediction: According

Steps model 5 submitted in a pocket finder for best receptor ligangbindingsite.

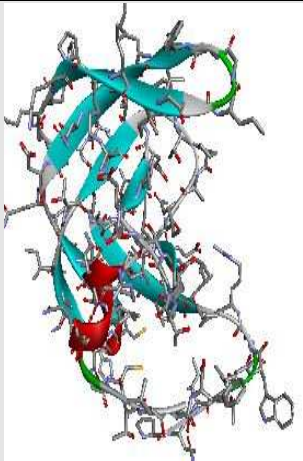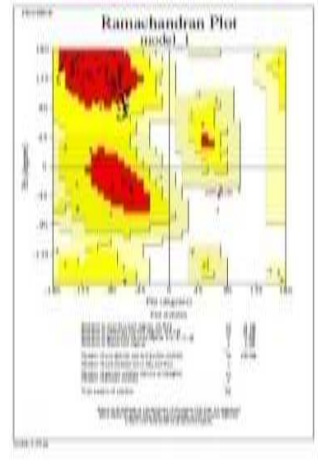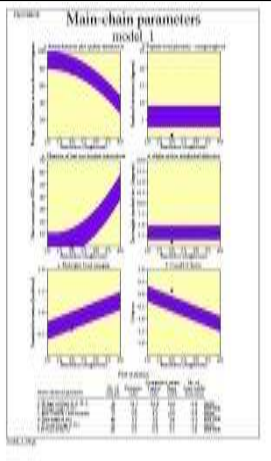***Table 2.*** *Potential structural analysis of Geno3D*

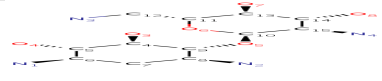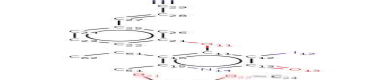| s.no	Model number	3Dstructure	Ramchandran plot	Main chain parameters
1	**Model.5**			

Table:2 Explains 3D model structure with different type of coloring bands with ramchandran plot and main chain Parameters.andRamachandran plot of the model.5 showing maximum residue in the core region (red colour), allowed region (dark yellow coloured), generously allowed (pale yellow colour).Geno3D web server is an online homology modelling server, it gave the ten model for query sequence in all the three model .5 have a maximum residues (89.9%)in the core region than compare to the model 1 to 10based on energy values and core region, so model number 5 is selected for further work

Table 3. *Site name and site volume in angstroms*

s.no	Site name	Site Volume (Angstroms)	Protein Volume Angstroms	Binding Box Around Selected Sites (max)	(min)
1.	Site 1	115 Cubic	8932 Cubic	(-19, 30, 39)	(-32, 17, 25)
2.	Site 2	108 Cubic	8932 Cubic	(-2, 11, 18)	(-16, -2, 6)
3.	Site 9	80 Cubic	8932 Cubic	(-8, 31, 36)	(-22, 16, 24)
4.	Site 10	72 Cubic	8932 Cubic	(-19, 30, 39)	(-32, 17, 25)

Table: 3 Explains selected site volume and protein volume in angstroms with binding boxes based on max and min parameters.

Table 4. *showing the legend's chemical query*

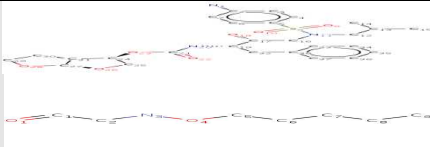| S.no	Ligand name ID	Chemical formula	molecular weight	Chemical structure
1	NVP	C15 H14 N4 O	266.30 g/mol	
2	XXX	C12 H26 N4 O6	322.36 g/mol	
3	THK	C23 H27 Cl F N O S	452.05 g/mol	
4	RIT	C37 H48 N6 O5 S2	720.94 g/mol	
5	P16	C21 H16 Cl2 N4 O2	427.28 g/mol	
6	P14	C16 H31 N7 O2	353.46 g/mol	
7	3AC	C17 H15 I N2 O2	406.22 g/mol	
8	3T1	C26 H23 Cl N2 O3	446.93 g/mol	

| 9 | O17 | C27 H37 N3 O7 S | 547.66 g/mol | |
| 10 | AOP | C7 H15 N O2 | 145.20 g/mol | |

Table: 4 Explains the query about ligand with ligand ID in PDB,with chemical structure and chemical formula and molecular weight.

3.4. Suitable Ligand Selection for Receptor Ligand Binding

In this project the ligand   selection for potential receptor ligand binding based on protein data   bank .then select the five potential ligand for example:table:4.
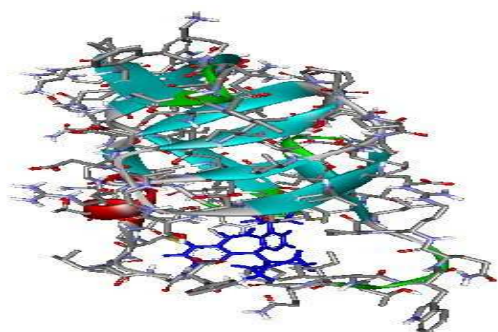

Fig 1. Showing the interaction between chemical and protein molecule blue color is a type of chemical molecule as a ligand

3.5. Docking

Prediction of the optimal physical configuration and energy between two molecules and the Docking is the phenomenon which enables the interaction between a receptor molecules and the ligand molecule, and mainly docking in this project based on receptor and ligand molecule the receptor molecule is large and ligand is small molecule.thedocking completed in this project with the help of Hex, PATCHDOCK.

And the model5 is a potential receptor molecule and model5 complete docking with potential ten ligands out of which ligand number 1(NVP) is fully bounded with model number five and also with minimum energy and all important parameters showing.
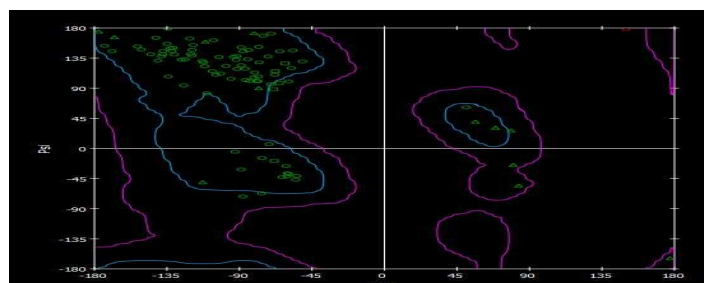

Fig 2. ramchandran plot of fig 1, Ramachandran plot explain the conditions of residues .and its shows in theory which values, or conformations, of the ψ and φ angles are possible for an amino-acid residue in a protein

Table 5. *Important query about ligand*

s. no	Name	Ligand name	Zinc _id
1.	Nevirapine(NVP)	11-CYCLOPROPYL-5,11-DIHYDRO-4-METHYL-6H-DIPYRIDO[3,2-B:2',3'-E][1,4] DIAZEPIN-6-ONE	Zinc_00004778

Table 6. *Different type of structure of protein receptor binding (docking structure)*
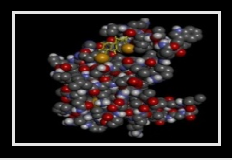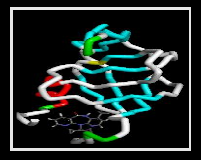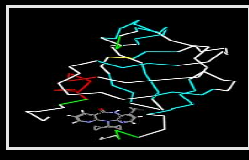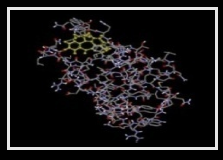
s.no	Stick and CPK structure	Stick and tube structure	Line and Ca wire structure	Ball and sticks structure (wire frame)

Table:6 Explains column 1 ligang showing by sticks and protein showing by cpk, and column 2 showing ligand with sticks and protein by tubes and the column 3 showing ligand with line and protein by ca wire and column fourth showing the wire frame structure.

4. Conclusion

Finally in this study of HIV protease structure that can be concluded10 number of model and numbers of binding sites and 10 number of potential ligand with 3Dstructure

which is predicted through the Insilico approaches and homology modeling and the docking of HIV1 protease with various ligand determine and the interaction between protein and ligand that's bind on active site of the protease enzyme ,although docking process is very complicated because its depends on various parameters the main resultant obtained by different type of docking tools and docking completed with the help of HEX,PATHADOCK for identify the suitable HIV1protease inhibitor in all the ligands which are docked with the protease only 10 numbers of ligand given the minimum energy and out of these 10 ligands NVP is playing an important role in HIV1and NVP ligand have minimum binding energy(-124.27) and its work as a potential protease inhibitor. Perhaps the ultimate solution is to develop a potential drug candidate against this devastating epidemic.

Acknowledgements

References

[1] Kohl N.E. Emini E.A. Schleif W.A. Davis L.J. Heimbach J.C. Dixon R.A. Scolnick E.M.Sigal I.S, "Active human immunodeficiency virus protease is required for viral infectivity" Proc.Natl Acad. Sci. U S A , 8: 4686-4690,1988.

[2] Winters M.A and Merigan T, "Insertions in the human immunodeficiency virus type 1 proteaseand reverse transcriptase genes:clinical impact and molecular mechanisms" Antimicrob Agents Chemother, 49:2575-2582,2005.

[3] Kozisek M. Saskova K.G. Rezacova P. Brynda J. van Maarseveen. N.M. de Jong. D Boucher. C.A. Kagan. R.M Nijhuis. M Konvalinka, "Ninety-nine is not enough: molecular characterization of inhibitor-resistant human immunodeficiency virus type 1 protease mutants with insertions in the flap region" J. Virol, 82:5869-5878,2008.

[4] Doyon L. Croteau G. Thibeault D. Poulin F. Pilote Lamarre, "Second locus involved inhuman immunodeficiency virus type 1 resistance to protease inhibitors" J. Virol, 70:3763-3769, 1996.

[5] Mammano F. Petit C. Clavel F, "Resistance-associated loss of viral fitness in human immunodeficiency virus type 1: phenotypic analysis of protease and gag coevolution in protease inhibitor-treated patients" J. Virol, 72:7632-7637, 1998.

[6] Navia MA. Fitzgerald PM. McKeever BM. Leu CT. Heimbach JC, "Three-dimensional structure of a spartyl protease from human immune deficiency virus HIV-1" Nature, 337:615–20, 1989.

[7] Pearl LH and Taylor WR, "A structural model for the retroviral proteases"Nature.329:351–54, 1987.

[8] Olson M. Hood L. Cantor C. Botstein D, "A common language for physical mapping of the human genome" Science, 245(4925):1434–1435, 1989.

[9] Berman H, "The Protein Data Bank: a historical perspective". ActaCrystallographica Section A: Foundations of Crystallography" A64, (1): 88–95, 2008.

[10] Meyer EF, "The first years of the Protein Data Bank. Protein Science" Cambridge University Press) 6 (7): 1591–1597, 1997.

[11] Berman HM. Westbrook J. Feng Z. Gilliland G. Bhat TN. Weissig H. Shindyalov IN. Bourne PE, "The Protein Data Bank" Nucleic Acids Res, 28 (1): 235–242,2000.

[12] Knox C. Law V. Jewison T. Liu P. Ly S. Frolkis A. Pon A. Banco K. Mak C. Neveu V. Djoumbou Y. Eisner R. Guo AC. Wishart DS, "DrugBank 3.0: a comprehensive resource for 'omics' research on drugs" Nucleic Acids Res, 39:D1035-41,2011.

[13] Wishart DS. Knox C. Guo AC. Cheng D. Shrivastava S. Tzur D. Gautam B. Hassanali M, "DrugBank: a knowledgebase for drugs, drug actions and drug targets" Nucleic Acids Res, 36:D901-6,2008

[14] Wishart DS. Knox C. Guo AC. Shrivastava S. Hassanali M. Stothard P, Chang Z. Woolsey J, "DrugBank: a comprehensive resource for in silico drug discovery and exploration" Nucleic Acids Res, Jan 1; 34:D668-72, 2006.

[15] Altschul, S Gish, W Miller, W Myers, E Lipman, D (1990). "Basic local alignment search tool". Journal of Molecular Biology 215 (3): 403–410, 1990.

[16] Brady G, Stouten P,(2000) Fast prediction and visualization of protein binding pockets with PASS. J Comput Aided MolDes 14:383-401, 2000.

[17] Laurie A, Jackson R: Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. Bioinformatics 2005, 21:1908-1916, 2005.

[18] Vincent Le Guilloux, Peter Schmidtke and Pierre Tuffery. (2009).Fpocket: An open source platform for ligand pocket detection", BMC Bioinformatics. 10:168, 2009.

[19] Kawabata T, "Detection of multi-scale pockets on protein surfaces using mathematical morphology" Proteins, 78: 1195-1121, 2010.

[20] Capra JA. Laskowski RA. Thornton JM. Singh M. Funkhouser TA, "Predicting Protein Ligand Binding Sites by Combining Evolutionary Sequence Conservation and 3D Structure"PLoSComputBiol, 5:(12),2009.

[21] D.W. Ritchie and V. Venkatraman, "Ultra-Fast FFT Protein Docking On Graphics Processors" Bioinformatics, 26: 2398-2405, 2010.

[22] Moore M. S.,"The GTP-binding protein Ran/TC4 is required for protein import into the nucleus" Nature, 365:661–663,1993.