DOI:10.21276/sjpm

Saudi Journal of Pathology and Microbiology
Scholars Middle East Publishers
Dubai, United Arab Emirates
Website:http://scholarsmepub.com/

ISSN 2518-3362 (Print)
ISSN 2518-3370(Online)

Original Research Article

Insights into Integrase resistance to Dolutegravir, Elvitegravir and Raltegravir-strand transfer inhibitors of HIV-1: A computational approach

Ameeruddin Nusrath Unissa*, Sukumar Swathi, Appisetty Ramya Lakshmi, Luke Elizabeth Hanna
Post Doctoral Fellow, Department of Biomedical Informatics, National Institute for Research in Tuberculosis (NIRT), Indian Council of Medical Research (ICMR), No. 1, Mayor Sathyamoorthy Road, Chetput, Chennai 600 031, Tamil Nadu, India

***Corresponding Author:**
Ameeruddin Nusrath Unissa
Email: nusrathunissa@gmail.com

Abstract: Integrase (IN) is an important enzyme for the replication of the type-1 human immunodeficiency virus (HIV-1), and an essential target for the development of anti-HIV drugs. The enzyme is currently targeted by first and second-generation inhibitors [raltegravir (RAL), elvitegravir (EVG) and dolutegravir (DTG)]. Of these, resistance to RAL and EVG are associated with three main pathways involving key mutations at positions N155H, Q148K/R/H, and Y143R/C within the IN gene. Although new resistance mutations appear to confer only low levels of cross-resistance to second-generation drugs (DTG), the Q148 pathway with numerous secondary mutations has the potential to significantly decrease susceptibility to all inhibitors of IN. In order to get insights into the development of IN resistance to first and second generation inhibitors using *in silico*approaches, in this study, one of the clinically essential mutant (MT)-Q148R of IN was modeled and docked with inhibitors. The MT model was built using the template and wild type (WT) 4E1M of HIV-1 integrase. Docking results indicate that in MT-Q148R the score was lower with respect to all inhibitors in comparison to the WT. In this study, the MT-Q148R of IN from HIV-1 displayed low affinity, this could be attributed to less number of interactions principally hydrogen and halogen bonds with the inhibitors (DTG, EVG and RAL), compared to the WT. Therefore, it can be suggested that the MT could not bind efficiently with inhibitors owing to its structural changes due to substitution to mediate its inhibitory activity and eventually leads to DTG, EVG and RAL resistance.
Keywords: HIV, IN resistance, DTG, EVG, RAL.

INTRODUCTION

Acquired immunodeficiency syndrome (AIDS) is the leading cause of mortality amongst the list of infectious diseases. AIDS is caused by human immunodeficiency virus (HIV), a disease that is currently treated with a combination therapy of drugs targeting essential enzymes (reverse transcriptase, protease, integrase) of HIV. Compared to drugs targeting reverse transcriptase and protease, drugs that targets type-1 HIV (HIV-1) integrase (IN), can prevent latent virus infection. As IN is a key enzyme for the replication of the HIV-1 and complete inhibition of IN could block chronic HIV-1 infection [1, 2]. IN inhibitors, also known as integrase strand transfer inhibitors (INSTIs), are a class of antiretroviral drug designed to block the action of IN, that inserts the viral genome into the DNA of the host cell. As integration is a primary step in retroviral replication, blocking it can inhibit the viral transmission. Since IN inhibitors target a distinct step in the retroviral life cycle, they may be taken in combination with other types of HIV drugs to minimize adaptation by the virus. IN inhibitors are potent antiretroviral drugs that efficiently decrease viral load in patients [3]. However, emergence of resistance mutations against this new class of drugs represents a threat to their long-term efficacy.

Clinically resistance to IN inhibitors such as dolutegravir (DTG), elvitegravir (EVG) and raltegravir (RAL)] has been documented [4, 5]. Of these, resistance to RAL and EVG (first generation INSTIs) based regimens failure are associated with three main pathways involving key mutations at positions N155H, Q148K/R/H, and Y143R/C within the IN gene, with polymorphisms among subtypes that may affect resistance and viral-replication capacity [6-11].

Genetic barrier to resistance for DTG (second-generation INSTIs) is very high compared to EVG and RAL. DTG has not selected for resistance in HIV-positive individuals when used in first-line therapy [12]. HIV containing the primary RAL/EVG resistance substitution N155H could select for R263K substitution under DTG pressure which displayed low-level resistance to DTG and the virus was found fit [13]. Although DTG has demonstrated greater resilience than

other antiretroviral drugs at withstanding the emergence of HIV-1 resistance mutations, such substitutions can develop, albeit rarely, in treatment-experienced IN inhibitor-naïve individuals. The most common substitution in integrase under those circumstances is R263K whereas other substitution that was selected against DTG in tissue culture was G118R and S153Y [5, 13]. Despite the fact that these new resistance mutations (R263K, G118R and S153Y) confer only low-level cross-resistance to second-generation drugs, the Q148 pathway with numerous secondary mutations has the potential to significantly decrease susceptibility to all members of the INSTI family of drugs [5, 14]. Selection of mutations *in vitro* with second-generation INSTIs suggests that only low level cross-resistance may exist between these new drugs and first-generation members of this class. The emergence of mutations at position Q148 should be monitored whenever possible and more data are needed to assess the long-term efficacy of second-generation INSTIs (DTG) in patients who may have failed first-generation INSTIs such as EVG and RAL.Deep sequencing, biochemical analyses and structural modeling are methods that currently help in the understanding of the mechanisms of resistance conferred by these various substitutions [5]. Hence, in the light of above, to understand the differences in the binding affinity between wild-type (WT) and one of the clinically essential mutant (MT- Q148R) of HIV-1 integrase protein that could lead to the cause of numerous secondary mutations that significantly decrease susceptibility to all members of the INSTI, in our study, this protein was allowed to interact with three drugs such as DTG, EVG and RAL through structural modeling approaches in comparison to WT.

MATERIALS AND METHODS

Protein: Homology modeling of mutant

The target IN protein sequence region (P12497) from Gag-pol polyprotein of HIV-1 was obtained from the Uniprot database [http://www.uniprot.org/], in the present study. The obtained protein sequence was submitted to protein alignment program (BLASTp) [15] and searched against protein database (PDB) [http://www.rcsb.org/pdb/home/home.do]. On the basis of the search, the crystal structure of IN protein 4E1M of HIV-1 was considered as template and wild type (WT) [16].

Model building

Residue at position 148 of IN of HIV-1 in template protein sequence was substituted from Gln to Arg for the creation of MT (Q148R) protein. Further, in the template protein 4E1M, A chain was retained and the heteroatoms such as water and others were removed. Sequence alignment between MT and template protein was performed with the command line options, and a series of commands were provided for model building using the software MODELLER9v14 [17].

Model evaluation

Validation of the model was done by Ramachandran plot [18]. Further the deviation between the WT and the MT-model upon structural superimposition was determined using PDBeFOLD [19].

Ligands

The ligands DTG, EVG and RAL used in this study were obtained from Chemspider database [http://www.chemspider.com]. Chemsketch software [20] was used to obtain the structure of the ligands in Mol format and the ligands were saved as Mol2 file using the software Discovery Studio [21].

Discovery Studio (DS)

The software v.2.0 was used for visualization purpose of modeled proteins and docking complexes [21].

Docking protocol

Docking was carried out with the help of software-GOLD [22]. The GOLD protocol is based on the principle of genetic algorithm wherein the receptor is held rigid while the ligands are allowed to flex during the refinement process.The input files for both the protein and the ligands were generated. Hydrogen atoms were added tothe models and ligands using auto edit option in GOLD before docking followed by energy minimization. The cavity atom file containing the atom number of binding residues (Asp64, Asp116, Glu152, Gln168, Ala169, Glu170, His171 and Thr174) was prepared for ligands such as DTG, EVG and RAL. The binding residues were selected based on comparison between the binding regions of 4E1M and MT protein [16].

Dockings were performed under 'Standard default settings' mode- number of islands was 5, population size was 100, number of operations was 100,000, niche size was 2, and selection pressure was 1.1. Ten docking poses were obtained for each ligand. Poses with highest GOLD score were used for further analysis. The docked poses of the ligands were visualized using DS. The scoring function of GOLD provides a way to rank the ligands relative to one another. Ideally, the score should correspond directly to the binding affinity of the ligand for the protein, so that the best scoring ligand poses are the best binders.

Biovia

This software was used to determine the interactions between the ligands and proteins [23].

RESULTS

Template selection and homology modeling of MT of IN

Using BLASTp search against PDB, 4E1M was identified as the template and WT as it displayed maximum identity with the IN protein sequence (Fig. 1). Mutant model of IN was built based on substitution at codon 148 in the WT- protein sequence (P12497) of IN of HIV-1, using MODELLER9v14 (Fig. 2). The generated model was validated by structural superimposition (Fig. 3). The root mean square deviation (RMSD) between the WT and MT was 0.5 Å, indicating least deviation and the reliability of the MT-model. This was also supported by Ramachandran plot analysis [18] which showed 96% of the residues in the favoured regions (Fig. 4).

NCBI Blast:Protein Sequence (288 letters)

Chain A, Crystal Structure Of Hiv-1 Integrase With A Non-Catayltic Site Inhibitor

Sequence ID: 4E1M_A **Length:** 166 **Number of Matches:** 1

e 1 more title(s)

Range 1: 1 to 166 GenPept Graphics Next Match Previous Match

Related Information

Structure - 3D structure displays

Identical Proteins - Identical proteins to 4E1M_A

Score	Expect	Method	Identities	Positives	Gaps
325 bits(834)	6e-114	Compositional matrix adjust.	157/166(95%)	160/166(96%)	0/166(0%)

```
Query  47   GEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAG  106
            G  MHGQVDCSPGIWQLD THLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAG
Sbjct  1    GSHMHGQVDCSPGIWQLDXTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAG  60

Query  107  RWPVKTIHTDNGSNFTGATVRAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVR  166
            RWPVKT+HTDNGSNFT ATV+AA WWAGIKQEFGIPYNPQSQGV+ESMNKELKKIIGQVR
Sbjct  61   RWPVKTVHTDNGSNFTSATVKAAXWWAGIKQEFGIPYNPQSQGVIESMNKELKKIIGQVR  120

Query  167  DQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKE  212
            DQAEHLKTAVQMAVFIHN KRKGGIGGYSAGERIVDIIATDIQTKE
Sbjct  121  DQAEHLKTAVQMAVFIHNKKRKGGIGGYSAGERIVDIIATDIQTKE  166
```

Fig-1: BLASTp result showing 95% identity between template (4E1M) and the target WT-IN sequence (P12497) of HIV-1

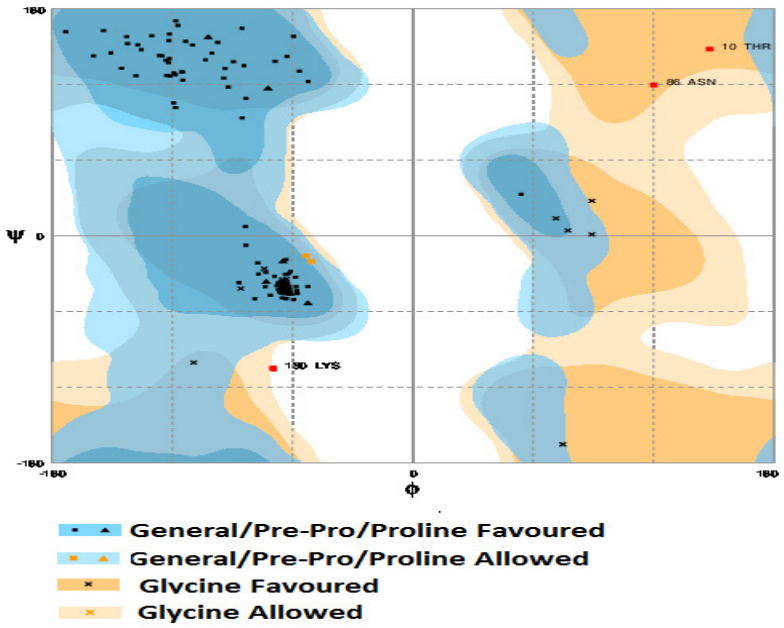

Fig-2: Ramachandran Plot

Evaluation of residues

Number of residues in favoured region (~98.0% expected) : 139 (96.5%)

Number of residues in allowed region (~2.0% expected) : 2 (1.4%)

Number of residues in outlier region : 3 (2.1%)

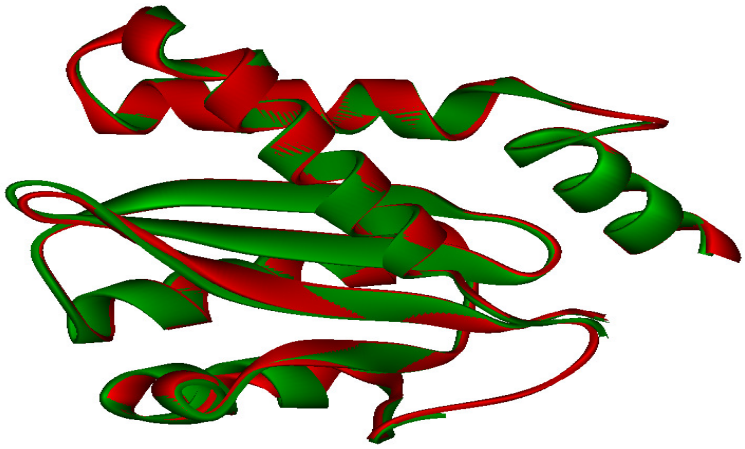

Fig-3: Superimposition of template-4E1M (green) and MT (red) showing an RMSD of 0.5 Å

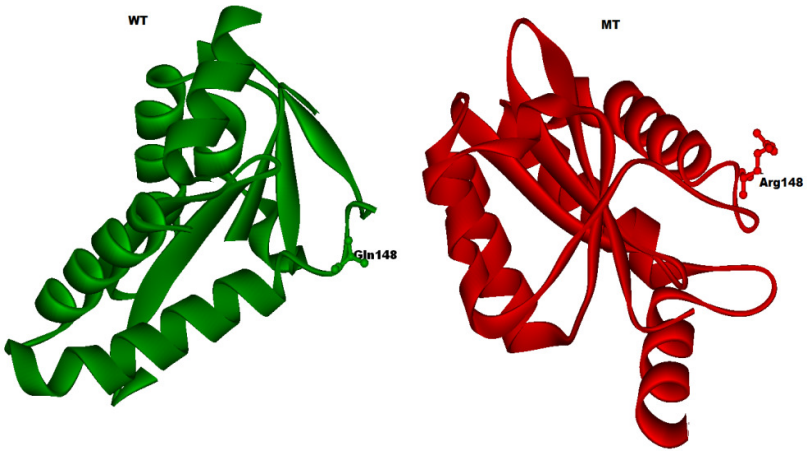

Fig-4: Structure of IN -WT (green) and 3-D model of MT (red)

Docking between INs and ligands

The generated model and WT proteins were used for docking with three ligands such as DTG, EVG and RAL, and the docked complexes were visualized using DS (Fig. 5, 6, and 7).Docking of ligands with INs resulted in ten poses. Of the ten poses, the best ligand pose was selected based on top GOLD score. In case of WT-IN, the score was consistent for all the three drugs showing values in the range of fifties (58, 53, 55 kcal/mol) for DTG, EVG and RAL, respectively. While in case of MT, the score (52 and 51 kcal/mol) was high for DTG and RAL, and low for EVG (45 kcal/mol) (Fig. 8).

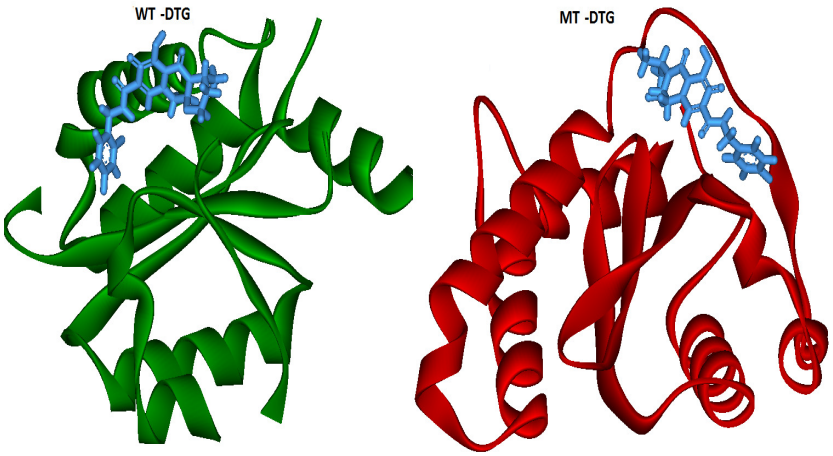

Fig-5: Docking complex WT and MT with DTG (blue)

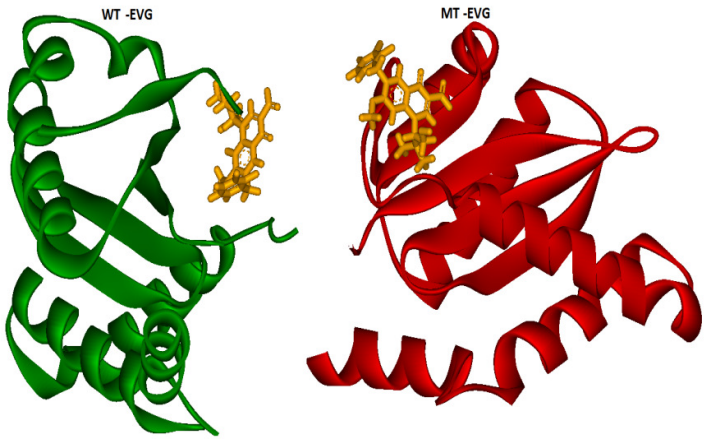

Fig-6: Docking complex of WT and MT with EVG (mustard)

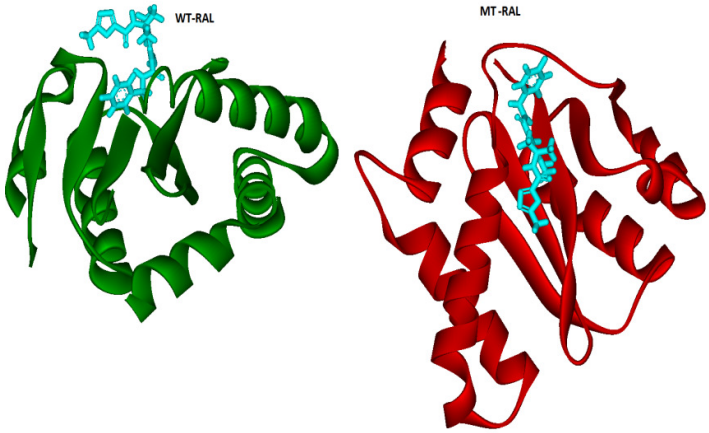

Fig-7: Docking complex of WT and MT with RAL (cyan)

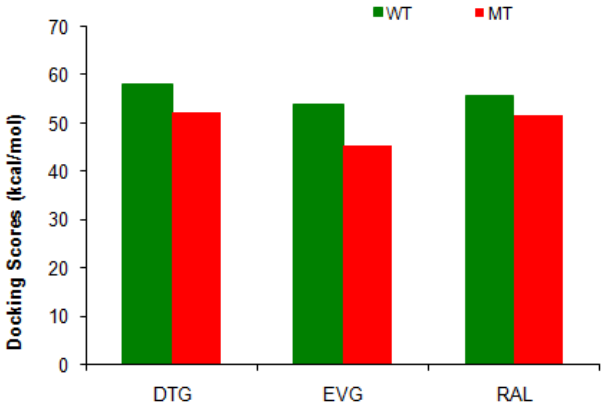

Fig-8: Docking score ofWT and MT with DTG, EVG and RAL

Interactions of IN proteins with the drugs at their binding site

In case of WT-IN of HIV-1, van der Waals interactions were found more in number compared to other types. Two Hydrogen (H) bonds (carbon H) between the drug (DTG) and one with Gln62; another with Asp116 residue of WT-IN were found and 3 conventional H bonds were formed between DTG and Asp116, Gly149, and Glu152 residues. Interestingly, a halogen bond between fluorine (F) atom of the drug and

residue Asp116 was formed. Further, Pi-anion bonds were formed between the drug and residues Asp116 and Glu152 respectively. Of note, it was found that the residue Asp116 was very actively participated in many interactions. Also, alkyl bonds were formed between the drug and residues Ile141 and Ile151, respectively (Fig. 9). Interactions in MT in complex with DTG followed a different way; the MT showed a conventional H bond formed between the drug and the residue Ile141 (modeling number is 84). Pi-Pi T shaped was formed between the drug and residue Phe139 (82). Other interactions are evident as shown in the Fig. 9.

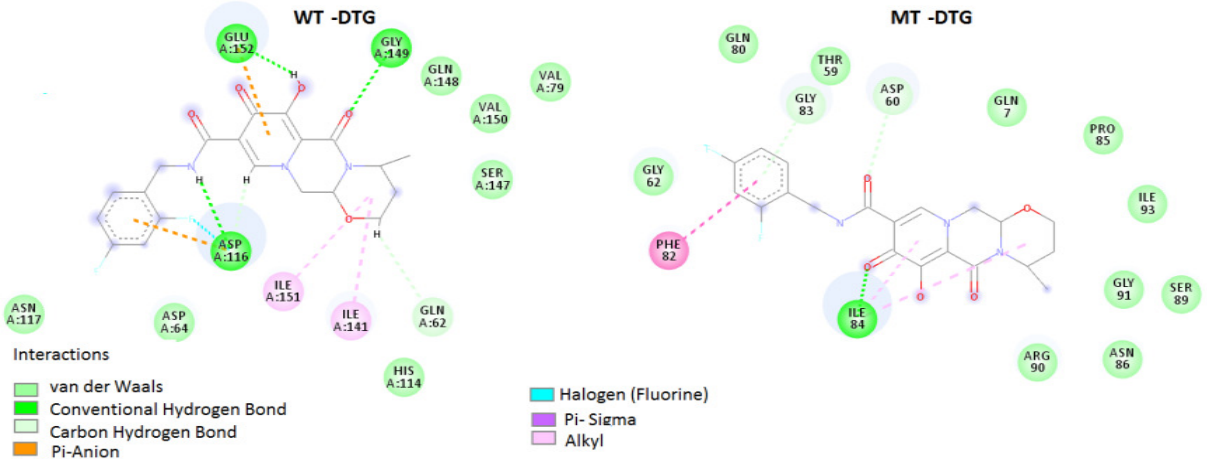

Fig-9: Interactions ofWT and MT with DTG at its binding site

In case of the ligand-EVG complexed with WT-IN protein, a conventional H bond between the drug (EVG) and the residue Asp116 was found; also two carbon H bonds were formed with the same residue (Asp116). Interestingly, two halogen (F) bonds were formed between EVG and with residues Pro142 and Asn144 of WT, respectively. Further, a bond with Chlorine atom was also found. Of note, several types of alkyl and Pi-alkyl interaction were found as evident in Fig. 10. In contrast in MT model, no H bonds were observed. More importantly, an unfavorable negative-negative interaction was formed between the drug and the residue Asp116 (modeling number is 60). Besides, no Halogen bonds were observed (Fig. 10).

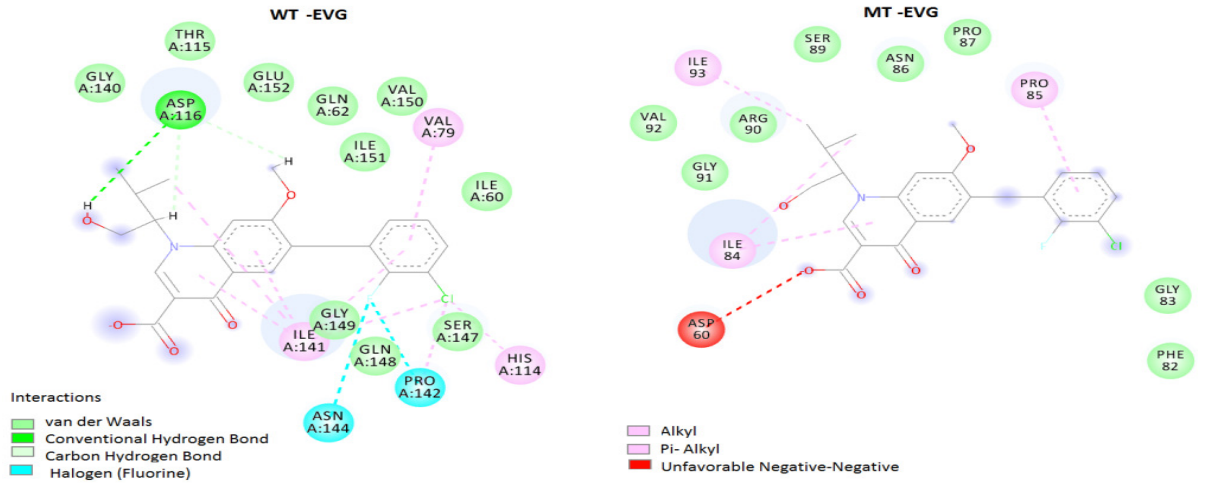

Fig-10: Interactions ofWT and MT with EVG at its binding site

In case of the RAL with WT, more of van der Waals interactions were found, followed by four conventional H bonds that were significantly contributed from Asp116 (2 bonds), Ile151 and Glu152 residues. Similar to EVG, F atom of RAL was also involved in interaction with Pro142 residue. Other types of interactions are shown in Fig. 11. On the other hand in case of MT-Q148R, followed by van der Waals interactions 3 conventional H bonds were imparted from Gln62 (modeling number 7), Asp116 (60) and Glu152 (94). Two carbon H bonds were formed between RAL and Asp116 (60), Ser147 (89) residues, respectively. More importantly, the mutated residue Arg148 (90) was directly involved in interaction with RAL (Fig. 11).

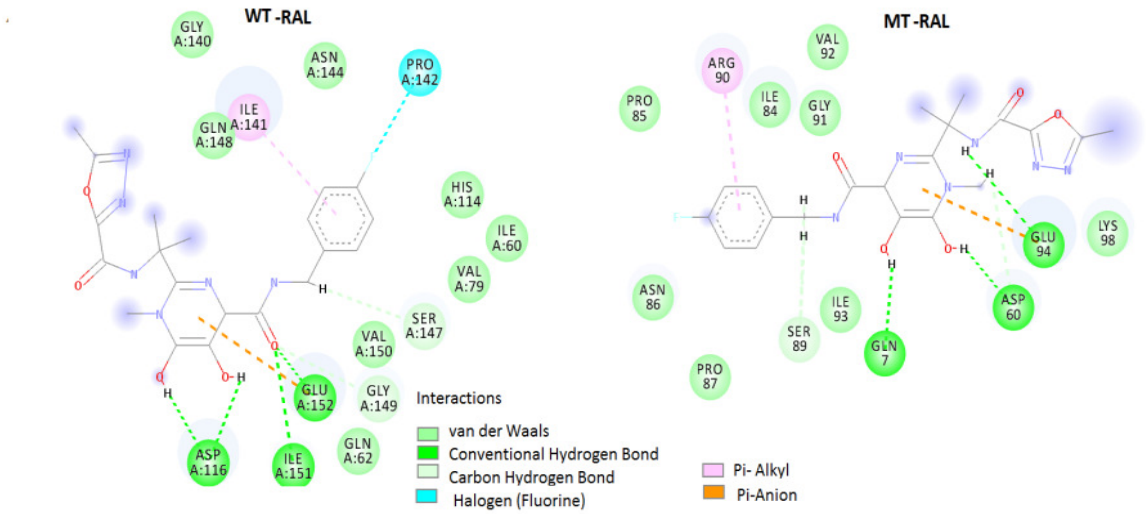

Fig-11: Interactions ofWT and MT with RAL at its binding site

DISCUSSION

Integrase (IN) is an important enzyme for the replication of the HIV-1 and responsible for chronic infection of the virus. IN is a protein of 32,000 Daltons, with three functional domains. The amino-terminal domain (amino acids 1-50) contains a conserved and essential zinc binding motif HHCC (histidines 12 and 16, cysteines 40 and 43) that binds one zinc atom. One known function of this region is protein multimerization [24]. The catalytic core domain (amino acids 50-212) contains the conserved motif consisting of the active site residues Asp64, Asp116, and Glu152. Mutation of any one of these three residues inactivates the enzyme. The third domain is the carboxyl-terminal domain (amino acids 213-288), which is important for nonspecific DNA binding of sub-terminal viral DNA [25, 26] or possibly host DNA.

IN inhibitors provides a new treatment option to patients who have developed resistance to protease and reverse transcriptase inhibitors. However, resistance was also found in patients under treatment based on IN inhibitors. First and second-generation INSTIs have been identified that were found to be associated with several resistance pathways. Although new resistance mutations appear to confer only low levels of cross-resistance to second-generation drugs, the Q148 pathway with numerous secondary mutations has the potential to significantly decrease susceptibility to all drugs of the INSTI family [5]. Structural modeling, biochemical analyses and deep sequencing are methods that currently help in the understanding of the mechanisms of resistance conferred by these mutations [5].In light of this, in the present study, the wild type (WT) and mutant (MT) Q148R of IN was modeled, and docked with three drugs, as interaction between these drugs with this clinically relevant MT (Q148R) was not explored previously, to our knowledge.

The results of the study indicate that in comparison to MT, WT showed more affinity with all the three drugs (DTG, RAL and EVG). The reason for the high score in the WT with DTG, RAL and EVG could be attributed to the presence of halogen bonds (fluorine and chlorine) and many H bonds (both conventional and carbon) compared to MT. As H bonds contributes to the stability and integrity of the protein ligand complex, in comparison to other types of interactions (alkyl, Pi anion) and halogen bonding exhibits a strong type of interactions. In other words, the low score in the MT, could be correlated with less number of important H bond interactions, lack of halogen bonding and the presence of unfavorable negative-negative interaction, as in the case of EVG. The reason for the less number of interactions could in turn be due to the substitution Arg in the MT protein in place of Gln in WT, as Arg contains a functional group (guanido) instead of Gln which is a basic amino acid, might have induced structural changes in the protein side chain, which was obvious in the changes in the pattern of interactions and consequently reflected in less score. Although, the effect of docking could be better explained after performing molecular dynamics for understanding their function precisely, yet, the information provided over here can be useful to understand the impact of such substitution and consequent changes in binding ability. Therefore, in this study, an effort was taken to understand the effect of binding affinity of WT and MT proteins of IN with DTG, RAL and EVG, which showed more affinity towards the WT compared to the resistant MTs. Therefore, it can be suggested that the mutant displayed less affinity with the drugs (DTG, EVG and RAL), because of the substitution that induces structural changes and drugs could not bind efficiently to mediate its inhibitory activity, thereby leads to DTG, EVG and RAL resistance. However, further studies are needed to

get a deeper understanding of the mechanism of IN resistance to inhibitors that will aid in development of inhibitors that are selective against IN which can circumvent the problem of inhibitors resistance.

ACKNOWLEDGMENT

Dr. A. NusrathUnissa received Post Doctoral Fellowship from Indian Council of Medical Research.

REFERENCES

1. Fesen, M. R., Kohn, K.W., & Leteurtre, F. (1993). Inhibitors of human immunodeficiency virus integrase. *ProcNatlAcadSciUSA, 90*, 2399-2403.
2. Nair, V. (2002). HIV integrase as a target for antiviral chemotherapy. *Rev Med Vir,12*, 179-193.
3. Johnson, A. A., Marchand, C., & Pommier Y. (2004). HIV-1 integrase inhibitors: a decade of research and two drugs in clinical trial. *Curr top med chem, 4*,1059-1077.
4. Wainberg, M. A., Zaharatos, G. J., & Brenner, B. G. (2011). Development of antiretroviral drug resistance. *N Eng J Med, 365*, 637-46.
5. Quashie, P. K., Mesplede, T., &Wainberg, M. A. (2013). HIV Drug Resistance and the Advent of Integrase Inhibitors. *Curr Infect Dis Rep*,15, 85-100.
6. Cooper, D. A., Steigbigel. R. T., Gatell, J. M., Rockstroh, J. K., Katlama, C., &Yeni, P. (2008). Subgroup and resistance analyses of raltegravir for resistant HIV-1 infection. *N Engl J Med*, 359*(4),*355–365.
7. Fransen, S., Karmochkine, M., & Huang. W. (2009). Longitudinal analysis of raltegravir susceptibility and integrase replication capacity of human immunodeficiency virus type 1 during virologic failure. *Antimicrob Agents Chemother,* 53, 4522-4524.
8. Charpentier, C., Karmochkine, M., &Laureillard, D. (2008). Drug resistance profiles for the HIV integrase gene in patients failing raltegravir salvage therapy. *HIV Med*, 9, 765-770.
9. Buzon, M. J., Dalmau, J., &Puertas, M. C. (2010).The HIV-1 integrase genotype strongly predicts raltegravir susceptibility but not viral fitness of primary virus isolates. *AIDS, 24,*17-25.
10. Fransen, S., Gupta, S., &Danovich, R. (2009). Loss of raltegravir susceptibility by human immunodeficiency virus type 1 is conferred via multiple nonoverlapping genetic pathways. *J. Virol, 83,*11440-11446.
11. Brenner, B. G., Lowe, M., Moisi, D., Hardy, I., Gagnon, S., Charest, H., Baril, J. G., Wainberg, M. A., & Roger, M. (2011).Subtype diversity associated with the development of HIV-1 resistance to integrase inhibitors. *J Med Virol, 83,* 751-759.
12. Oliveira, M., Mesplede, T., Moisi, D., Ibanescu, R. I., Brenner, B., &Wainberg, M. (2015). The dolutegravir R263K resistance mutation in HIV-1 integrase is incompatible with the emergence of resistance against raltegravir. *AIDS, 29,* 2255-2260.
13. Anstett, K., Fusco, R., Cutillas, V., Mesplède, T., &Wainberg, M. A. (2015). Dolutegravir-Selected HIV-1 Containing the N155H and R263K Resistance Substitutions Does Not Acquire Additional Compensatory. Mutations under Drug Pressure That Lead to Higher-Level Resistance and Increased Replicative Capacity. *J Virol, 89,* 10482-10488.
14. Goethals, O., Van Ginderen, M., Vos, A., Cummings, M. D., Van Der Borght, K., Van Wesenbeeck, L., Feyaerts, M., Verheyen, A., Smits, V., Van Loock, M., &Hertogs, K. (2011). Resistance to raltegravir highlights integrase mutations at codon 148 in conferring crossresistance to a second-generation HIV-1 integrase inhibitor. *Antiviral Res*, *91,*167–176.
15. Altschul, S. F., Madden, T. L., Schäffer, A. A. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research, 25,* 3389-3402.
16. Tsiang, M., Jones, G. S., Niedziela-Majka, A., Kan, E., Lansdon, E. B., Huang, W., Hung, M., Samuel, D., Novikov, N., Xu, Y.., & Mitchell, M. (2012). New class of HIV-1 integrase (IN) inhibitors with a dual mode of action. *J BiolChem,287*, 21189-21203.
17. Sali, MODELLER: Implementing 3D Protein Modeling, mc2, vol. 2:5, Molecular Simulations Inc., 1995.
18. Lovell, S. C., Davis, I. W., &Arendall, W. B. (2003). Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. *Proteins: Structure, Function, Bioinformatics*, *50,*437-50.
19. Krissinel, E., Henrick, K. (2003). Protein structure comparison in 3D based on secondary structure matching (PDBeFold) followed by Cα alignment, scored by a new structural similarity function. In Proceedings of the 5th International Conference on Molecular Structural Biology, Vienna. Vol. 88.
20. ACD/Chemsketch, Version-10.0. (2006).Advanced Chemistry Development Labs.
21. Discovery studio, Version 2, Accelrys Inc., San Diego, CA, 2007 http://www.accelrys.com.
22. Jones, G., Willett, P., & Glen, R. C. (1995). Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J MolBiol, 31*, 43-53.
23. DassaultSystemes BIOVIA. (2015). Discovery Studio Modeling Environment, Release 4.5, San Diego: DassaultSystèmes,
24. Zheng, R., Jenkins, T. M., &Craigie, R. (1996). Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. *Proc. Natl. Acad. Sci.USA,* 93,13659-13664.
25. Esposito, D., &Craigie, R. (1998). Sequence selectivity of viral end DNA binding by HIV-1

integrase reveals critical regions for protein- DNA interactions. *EMBO J, 17,*5832-5843.

26. Jenkins, T. M., Esposito, & D., Engelman, A. (1997). Critical contacts between HIV-1 integrase and viral DNA identified by structure-based analysis and photo-crosslinking. *EMBO J, 16,* 6849-6859.