In silico Mutational analysis and Drug interaction studies of HIV Reverse Transcriptase

Rajasekhar Pinnamaneni
Department of Biotechnology, Sreenidhi Institute of Science and Technology, Yamnampet, Ghatkesar, R.R.District, Andhra Pradesh, India

Kondeti Subramanyam
Plant Molecular Biology Laboratory, Department of Biotechnology, Bharathidasan University, Tiruchirappalli – 620024, Tamil Nadu, India.

M. Thirumavalavan
Department of Biotechnology, Sreenidhi Institute of Science and Technology, Yamnampet, Ghatkesar, R.R.District, Andhra Pradesh, India

Nateshan Anil
Department of Biotechnology, Sreenidhi Institute of Science and Technology,Yamnampet, Ghatkesar, R.R.District, Andhra Pradesh, India

Koona Subramanyam
Department of Biotechnology, Sreenidhi Institute of Science and Technology, Yamnampet, Ghatkesar, R.R.District, Andhra Pradesh, India

ABSTRACT

Human immunodeficiency virus is a retrovirus that causes acquired immunodeficiency syndrome, a condition in humans in which the immune system begins to fail, leading to life threatening opportunistic infections. Reverse transcriptase inhibitors are class of antiretroviral drugs used to treat HIV infection. RTIs inhibit activity of reverse transcriptase, a viral DNA polymerase enzyme that HIV needs to replace. Protein P04585|POL_HV1H2 Gag-Pol polyprotein - Human immunodeficiency virus type 1 (isolate HXB2 group M subtype B) (HIV-1) was docked with RTIs and thebfree energies of the docking complexes were analyzed. It was concluded that M184V mutations has least impact than that of other mutations and Y115F mutation has great impact on the drug interaction with Abcavir. This has to be correlated with the real effect on the patients. This serves a best model for evaluating the impact of mutations in changing the folding of protein and further its effect on the inhibition. This is shown by superimposing the wild type structures and mutant structures.

Key Words: Reverse transcriptase, Inhibitors.

1. INTRODUCTION

HIV (Human immunodeficiency virus) is a retrovirus that causes acquired immunodeficiency syndrome, a condition in humans in which the immune system begins to fail, leading to life threatening opportunistic infections. Infection with HIV occurs by transfer of blood, semen, vaginal fluid, pre-ejaculate, breast milk, unprotected sexual intercourse, contaminated needles and from mother. The retrovirus HIV, and its subsequent progression to AIDS, is a rapidly growing worldwide epidemic. HIV-1 reverse transcriptase is one of the key players in the mechanism of infection by this retrovirus. The HIV-1 reverse transcriptase enzyme is responsible for copying a single-stranded viral RNA genome into double-stranded DNA [1].

The newly created DNA can then be incorporated into the host genome; the host is mainly the human in the case of HIV. The HIV-1 reverse transcriptase enzyme contains two main domains: a DNA polymerase domain and a ribonuclease H (RNase H) domain. The DNA polymerase is able to copy either an RNA or DNA template. The function of the RNase H domain is to cleave and degrade the template RNA after DNA synthesis so that the newly made DNA can generate a second DNA strand. The RNase H domain is also responsible for the integration of the duplex DNA into the host cell chromosome. Here we describe the crystal structure of HIV-1 reverse transcriptase complexed with two FAB-28 monoclonal antibody fragments and a DNA: RNA hybrid. The FAB-28 heavy chain and FAB-28 light chain are not actual components of the reverse transcriptase enzyme. The antibody fragments are complexed with the enzyme during the crystallization procedure in order to stabilize the enzyme structure [1].

Reverse transcriptase inhibitors (RTIs) are a class of antiretroviral drug used to treat HIV infection, tumors and cancer [2]. RTIs inhibit activity of reverse transcriptase, a viral DNA polymerase enzyme that retroviruses need to reproduce. When HIV infects a cell, reverse transcriptase copies the viral single stranded RNA genome into a double-stranded viral DNA. The viral DNA is then integrated into the host chromosomal DNA, which then allows host cellular processes, such as transcription and translation to reproduce the virus. RTIs block reverse transcriptase's enzymatic function and prevent completion of synthesis of the double-stranded viral DNA, thus preventing HIV from multiplying.

The mode of action of Nucleoside analog reverse transcriptase inhibitors (NRTIs) and Nucleotide analog reverse transcriptase inhibitors (NtRTIs) is essentially the same; they are analogues of the naturally occurring deoxynucleotides needed to synthesize the viral DNA and they compete with the natural deoxynucleotides for incorporation into the growing viral DNA chain. However, unlike the natural deoxynucleotides substrates, NRTIs and NtRTIs lack a 3'-hydroxyl group on the deoxyribose moiety. As a result, following incorporation of an NRTI or an NtRTI, the next incoming deoxynucleotide cannot form the next 5'-3' phosphodiester bond needed to extend the DNA chain. Thus,

when an NRTI or NtRTI is incorporated, viral DNA synthesis is halted, a process known as chain termination. All NRTIs and NtRTIs are classified as competitive substrate inhibitors.

In contrast, NNRTIs have a completely different mode of action. NNRTIs block reverse transcriptase by binding at a different site on the enzyme, compared to Nucleoside analog reverse transcriptase inhibitors (NRTIs) and Nucleotide analog reverse transcriptase inhibitors (NtRTIs). Non-nucleoside reverse transcriptase inhibitors (NNRTIs) are not incorporated into the viral DNA but instead inhibit the movement of protein domains of reverse transcriptase that are needed to carry out the process of DNA synthesis. NNRTIs are therefore classified as non-competitive inhibitors of reverse transcriptase.

Inhibitors of HIV reverse transcriptase (RT) are important drugs for the treatment of acquired immuno-deficiency syndrome (AIDS). One approach to identify novel inhibitors of HIV-1-RT is the screening of natural compounds. Many natural products have been shown to be active as RT inhibitors. These compounds belong to a wide range of different structural classes, e.g., coumarins, flavonoids, tannins, alkaloids, lignans, terpenes, naphtho- and anthraquinones, and polysaccharides. The life forms from which the bioactive compounds were isolated are as equally diverse and comprise terrestrial and marine plants, micro-organisms, and marine animals. From the most extensive screening effort, carried out by the NCl, calanolide A, isolated from the terrestrial plant *Calophyllum lanigerum* (Guttiferae), has been discovered as the most interesting natural RT inhibitor. The promise of this natural product probably relates to a novel mechanism of action [3].

Here it is focused on why these nucleoside inhibitors are not effective in some of the patients. So by creating the mutations in enzyme and by modeling it, the docking was done between the mutant enzyme and wild type enzyme with those drugs Later, the obtained energy values were analyzed and the result was concluded

2. MATERIALS AND METHODS

The protein sequence of Protein P04585|POL_HV1H2 Gag-Pol polyprotein - Human immunodeficiency virus type 1 (isolate HXB2 group M subtype B) (HIV-1) was known and its protein sequence was subjected to PSI-BLAST at NCBI. Protein parameters were analysed by using the tool Prosite [4]. All the protein parameters with respect to amino acid composition, secondary structure prediction, hydrophobicity, isoelectric point etc were analyzed. The generated model was subjected to several repeated cycles of energy minimization using SPDBV software [5] and the final model was subjected to docking.

2.1 Inhibitors used in the present study

- Zidovudine, also called AZT, ZDV, and azidothymidine, has the trade name Retrovir. Zidovudine was the first antiretroviral drug approved by the FDA for the treatment of HIV.
- Didanosine, also called ddI, with the trade names Videx and Videx EC, was the second FDA-approved antiretroviral drug.
- Stavudine, also called d4T, has trade names Zerit and Zerit XR.
- Lamivudine, also called 3TC, has the trade name Epivir.
- Abacavir, also called ABC, has the trade name Ziagen, is an analog of guanosine.
- Emtricitabine, also called FTC, has the trade name Emtriva (formerly Coviracil).
- Tenofovir, also known as tenofovir disoproxil fumarate, has the trade name Viread.

2.2 SOPMA

SOPMA is a secondary structure prediction program that uses multiple alignments. SOPMA correctly predicts 69.5% of amino acids for a three-state description of the secondary structure (alpha-helix, beta-sheet and coil) in a whole database containing 126 chains of non-homologous (less than 25% identity) proteins. Joint prediction with SOPMA and PHD correctly predicts 82.2% of residues for 74% of co-predicted aminoacids.SOPMA (Self-Optimized Prediction Method with Alignment) is based on the homologue method of [6] The improvement takes place in the fact that SOPMA takes into account information from an alignment of sequences belonging to the same family [7].

2.3 PubChem

PubChem is a database of chemical molecules. The system is maintained by the National Center for Biotechnology Information (NCBI), a component of the National Library of Medicine, which is part of the United States National Institutes of Health (NIH). PubChem can be accessed for free through a web user interface. Millions of compound structures and descriptive datasets can be freely downloaded via FTP. PubChem contains substance descriptions and small molecules with fewer than 1000 atoms and 1000 bonds. The American Chemical Society tried to get the U.S. Congress to restrict the operation of PubChem, because they claim it competes with their Chemical Abstracts Service. More than 80 database vendors contribute to the growing PubChem database.

2.4 Molegro Virtual Docker

Molegro Virtual Docker is an integrated platform for predicting protein - ligand interactions. Molegro Virtual Docker handles all aspects of the docking process from preparation of the molecules to determination of the potential binding sites of the target protein, and prediction of the binding modes of the ligands. Molegro Virtual Docker offers high-quality docking based on a novel optimization technique combined with a user interface experience focusing on usability and productivity. The Molegro Virtual Docker (MVD) has been shown to yield higher docking accuracy than other state-of-the-art docking products (MVD: 87%, Glide: 82%, Surflex: 75%, FlexX: 58%) [8].

3. RESULTS AND DISCUSSION

Protein: P04585|POL_HV1H2 Gag-Pol polyprotein - Human immunodeficiency virus type 1 (isolate HXB2 group M subtype B) (HIV-1).

3.1 Aminoacid composition

Length = 1435 amino acids

Molecular Weight = 162033.71 Daltons

Amino Acid		Number	Mol%
Ala	A	96	6.69
Cys	C	20	1.39
Asp	D	63	4.39
Glu	E	107	7.46
Phe	F	35	2.44
Gly	G	108	7.53
His	H	27	1.88
Ile	I	100	6.97
Lys	K	122	8.50
Leu	L	110	7.67
Met	M	29	2.02
Asn	N	58	4.04
Pro	P	77	5.37
Gln	Q	95	6.62
Arg	R	69	4.81
Ser	S	59	4.11
Thr	T	88	6.13
Val	V	96	6.69
Trp	W	37	2.58
Tyr	Y	39	2.72

3.2 Pepstats

Protein P04585|POL_HV1H2 Gag-Pol polyprotein - Human immunodeficiency virus type 1 (isolate HXB2 group M subtype B) (HIV-1) Pepstats of 1-578 constitutes 578 residues whose molecular weight is 66681.24 and the average residue weight is115.365. The overall protein charge is 8.0 and its Isoelectric point is 8.4613. The molar extinction Coefficient is 136270 and the improbability of expression in inclusion bodies is 0.639. The total number of negatively charged residues (Asp + Glu) is 78 and the total number of positively charged residues (Arg + Lys) is 80. The atomic composition of the peptide is 3034 carbons, 4739 Hydrogens, 799 Nitrogens, 872 Oxygens and 10 Sulphur atoms.

PEPSTATS of from 1 to 578
Molecular weight = 66681.24
Residues = 578
Average Residue Weight = 115.365 Charge = 8.0
Isoelectric Point = 8.4613

A280 Molar Extinction Coefficient = 136270
A280 Extinction Coefficient 1mg/ml = 2.04
Improbability of expression in inclusion bodies = 0.639

Residue	Number	Mole%	DayhoffStat
A = Ala	28	4.844	0.563
B = Asx	0	0.000	0.000
C = Cys	2	0.346	0.119
D = Asp	27	4.671	0.849
E = Glu	51	8.824	1.471
F = Phe	15	2.595	0.721
G = Gly	33	5.709	0.680
H = His	12	2.076	1.038
I = Ile	38	6.574	1.461
K = Lys	57	9.862	1.494
L = Leu	47	8.131	1.099
M = Met	8	1.384	0.814
N = Asn	19	3.287	0.764
P = Pro	36	6.228	1.198
Q = Gln	33	5.709	1.464
R = Arg	23	3.979	0.812
S = Ser	27	4.671	0.667
T = Thr	36	6.228	1.021
V = Val	45	7.785	1.180
W = Trp	19	3.287	2.529
X = Xaa	0	0.000	0.000
Y = Tyr	22	3.806	1.119
Z = Glx	0	0.000	0.000

Property	Residues	Number	Mole%
Tiny	(A+C+G+S+T)	126	21.799
Small	(A+B+C+D+G+N+P+S+T+V)	253	43.772
Aliphatic	(I+L+V)	130	22.491
Aromatic	(F+H+W+Y)	68	11.765
Non-polar	(A+C+F+G+I+L+M+P+V+W+Y)	293	50.692

```
Polar          (D+E+H+K+N+Q+R+S+T+Z)        285          49.308
Charged        B+D+E+H+K+R+Z)               170          29.412
Basic          (H+K+R)                       92          15.917
Acidic         (B+D+E+Z)                     78          13.495
```

Atomic composition

```
Carbon         C            3034
Hydrogen       H            4739
Nitrogen       N             799
Oxygen         O             872
Sulfur         S              10
```

Formula: $C_{3034}H_{4739}N_{799}O_{872}S_{10}$
Total number of atoms: 9454

3.3 Pepwheel

Protein P04585|POL_HV1H2 Gag-Pol polyprotein - Human immunodeficiency virus type 1 (isolate HXB2 group M subtype B) (HIV-1) Pepwheel of 1-578 residues is composed of 208 helices whose percentage is 37, 131 β- sheets accounting 23.3 percent, 107 turns that account 19 percent and 132 coils making 23.5 percent (Fig 1).

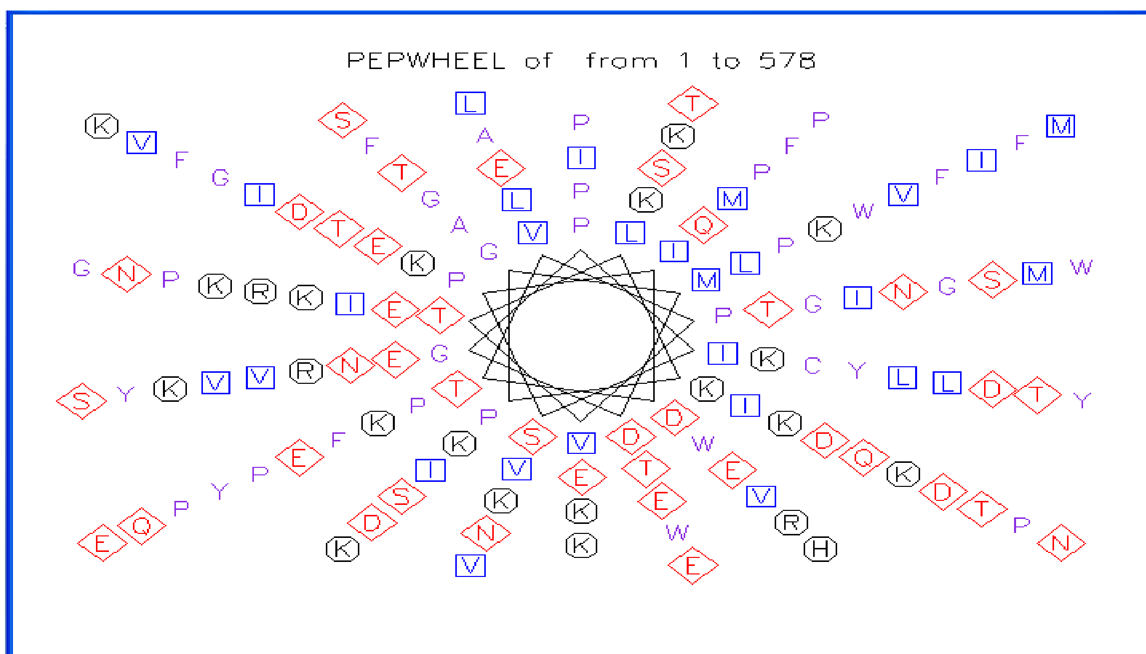

Fig 1: Pepwheel of 1-578 residues

Sequence: from: 1 to: 578
HitCount: 169
DCH = 0, DCS = 0
Please cite:
Garnier, Osguthorpe and Robson (1978) J. Mol. Biol. 120:97-120
#=======================================

```
          .   10    .   20    .   30    .   40    .   50
      pispidtvpvtlkpgmdgpkvkqwplteekikalteickemekegkiski
helix                             HHHHHHHHHHHHHHHHHHHHHHH
sheet    E EEEEEEEE         E                           E
```

```
turns T   T        TTT  TT                             T
coil  CC             CC  CC C                       C
          .   60    .   70    .   80    .   90    .  100
      gpenpyntpvfaikkkdstkwrklvdfrelnkrtqdfwevqlgiphpagl
helix         HHHHHHHH HHHHH   HHHHH                HHHH
sheet        E              EEE            EE
turns  T  TT          T             TTT   T  TTT
coil C CC  C                          CCC      CCCC
          .  110    .  120    .  130    .  140    .  150
      kkkksvtvldvgdayfsvpldesfrkytaftipsmnnetpgiryqynvlp
helix HH                  HHHHH
sheet    EEEEEEEEE EEEEE          EEE          EEEEEEEE
turns   T         T            TT            TT        T
coil                   CC       C   CCCCCCCC
          .  160    .  170    .  180    .  190    .  200
      qgwkgspaifqssmtkilepfriknpemviyqymddlyvgsdleigqhrt
helix          HH     H  HH                    HHHHHHH
sheet         E        EE        EEEEEEEEEEEE
turns TTT                  TT   T
coil    CCCCC   CCCCC       CCC             CC       CC
          .  210    .  220    .  230    .  240    .  250
      kieelrahllrwgfttpdkkhqkeppflwmgyelhpdrwtvqpielpekd
helix HHHHHHHHHH                 HHH
sheet                                          E
turns           TTTT   TTT T   TT   TT   TTTT        TTT
coil               CCC   C CCC       CCC    CC CCCCC
          .  260    .  270    .  280    .  290    .  300
      swtvndiqklvgklnwasqiyagikvkqlcklllrgakaltevvplteeae
helix  HHHHHH         HH     HHHHHH HHHHHHHHH   HHHHHHHH
sheet        EEEEE      EEE        E         EEE
turns T             T      TT
coil             CC C
          .  310    .  320    .  330    .  340    .  350
      lelaenreilktpvhgvyydpskdlvvevqkqgqdqwtyqiyqepfknlk
helix HHHHHHHHHH             H  HHHH                  H
sheet           E E EEEEE     EE           EEEEE
turns              T      TTT       TTTTTTT       TTTT T
coil            C       C                      CC
          .  360    .  370    .  380    .  390    .  400
      tgkyarkrsahtndvrqlaevvqkvatesiviwgkipkfrlpiqretwet
helix H HHHH        HHHHHHHHHHHHHH   HH
sheet       E                     EEE        E
turns  T     T     T                   TT  TT       T  T
coil         CCCCC                      CC   CCCCCC CC
          .  410    .  420    .  430    .  440    .  450
      wwmeywqatwipewefvntpplvklwyqlekdpivgaetfyvdgaasret
helix                       H   HHH      HH    HHHHHHHHH
sheet                        EEE        E  EEEE
turns T  TTT      TT TT   TT       TT
coil  CC   CCCCCC  C  CCC           CCC
          .  460    .  470    .  480    .  490    .  500
      klgkagyvtnrgrqkvvsltettnqktelhaihlalqdsgsevnivtdsq
helix HH                      HHHHHHHHHHHH
sheet      EEEE     EEEEEE                      EEEEE
turns   TT     TTTT                                  T T
coil     C        C      CCCC            CCCCCC      C
          .  510    .  520    .  530    .  540    .  550
      yalgiiqaqpdrsesevvnqiieelikkekvylswvpahkgiggneqvdk
helix                HHHHHHHHHHHHHHHHHHHHHH
sheet   EEEEEE                                      EEEE
turns T                                    T
coil  C      CCCCCCC                       CCCCCCCC
          .  560    .  570
      lvssgirkvlfldgidkaqeeheryhsn
helix   H          HHHHHHHHHHH
sheet EE   EEEEEEE
turns     T       T           TT T
coil    C                      C
```

```
#---------------------------------------
#
#  Residue totals: H:208   E:131   T:107   C:132
#         percent: H: 37.0 E: 23.3 T: 19.0 C: 23.5
#
#---------------------------------------
```

3.4 SOPMA

The parameters of secondary structure by SOPMA are

Window width : 17
Similarity threshold : 8
Number of states : 4

The secondary structure analysis was carried out by different tools like Garnier and sopma. Both of them have almost given the same results. It contains various types of secondary structures like 208 Alpha helices, 113 extended strands, 40 sheets and 217 Random coils (Fig 2).

```
PISPIDTVPVTLKPGMDGPKVKQWPLTEEKIKALTEICKEMEKEGKISKIGPENPYNTPVFAIKKKDSTK
TCCCCCCEEEEECCCCCCCCCCCCCCCCHHHHHHHHHHHHHHHHTTCCCCCCCCCCCCCCEEEEECCCCCC

WRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDESFRKYTAFTIPSMNNETP
HEEEEHHHHHHHHHHHHHHEEECCCCCCCCCCCCEEEEEECCCCEEECCCCCCCCCEEEEEEECCCCCCCT

GIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRIKNPEMVIYQYMDDLYVGSDLEIGQHRTKIEELRAHLL
TCEEEEEECCTTTTCCCHHHHHHHHHHHHHHHHTTTTCEEEEEHHHHEEECCCCCHHHHHHHHHHHHHHHH

RWGFTTPDKKHQKEPPFLWMGYELHPDRWTVQPIELPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLC
HTTCCCCCCCCCCCCCCEEEEEECCTTCCEEEECCCCCCCCCCCHHHHHHHHHHHHHHHHCCTTCCHHHHH


KLLRGAKALTEVVPLTEEAELELAENREILKTPVHGVYYDPSKDLVVEVQKQGQDQWTYQIYQEPFKNLK
HHHTTCCHHHHHHHHHHHHHHHHHHHHHHHHHCTTCEEECTTCHHEEEEHHCCCCCCEEEEEECCCCCCEE

TGKYARKRSAHTNDVRQLAEVVQKVATESIVIWGKIPKFRLPIQRETWETWWMEYWQATWIPEWEFVNTP
TTCCCCCCCCCCCHHHHHHHHHHHHHHHHHEEEETCCCCEECCCCCHHHHHHHHHHHHHTCCCCCEEECCC

PLVKLWYQLEKDPIVGAETFYVDGAASRETKLGKAGYVTNRGRQKVVSLTETTNQKTELHAIHLALQDSG
CHHHHHHHCCCCCCCCCEEEEETCCCCCCCTTCEEEEETTTCCEEEEEHCCCCCHHHHHHHHHHHHHTCC

SEVNIVTDSQYALGIIQAQPDRSESEVVNQIIEELIKKEKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVL
CCEEEEEECCHHHHEEEECCCCCCCCHHHHHHHHHHHHHTTHEEEEECCCCCCCCCCHHHHHHHHTTCCHEE

FLDGIDKAQEEHERYHSN
EEHCCCHHHHHHHHHHHE
```

Sequence length: 578

```
   Alpha helix     (Hh) :        208 is  35.99%
   3_10  helix     (Gg) :          0 is   0.00%
   Pi helix        (Ii) :          0 is   0.00%
   Beta bridge     (Bb) :          0 is   0.00%
   Extended strand (Ee) :        113 is  19.55%
   Beta turn       (Tt) :         40 is   6.92%
   Bend region     (Ss) :          0 is   0.00%
   Random coil     (Cc) :        217 is  37.54%
   Ambigous states (?)  :          0 is   0.00%
   Other states         :          0 is   0.00%
```

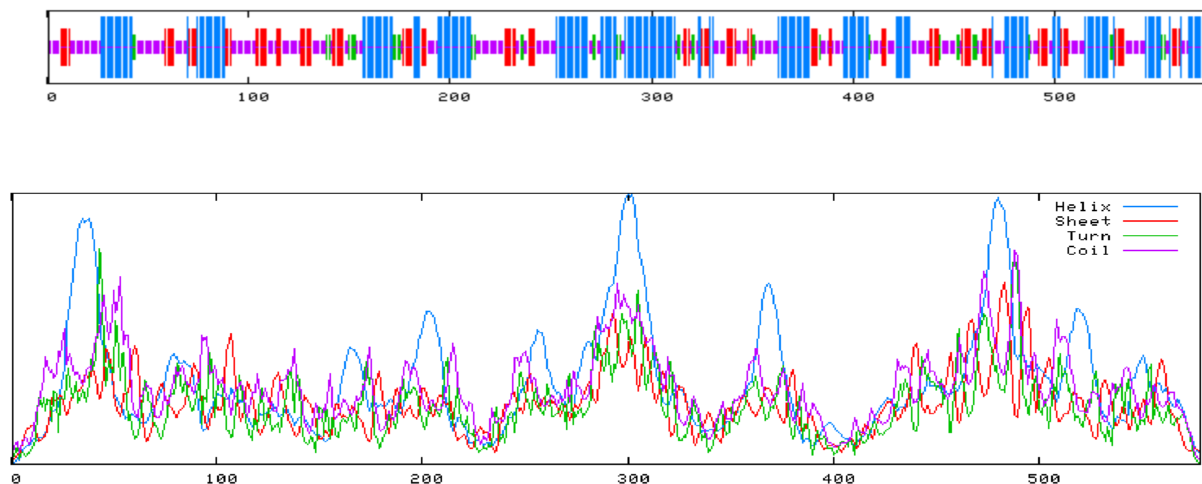

Fig. 2: SOPMA Analysis

3.5 Molegro Virtual Docking Results

Table 1: Molegro Virtual Docking Results

Drug	Enzyme	Energy Value
ABCAVIR	WILD	-143.471
	K65R	-152.369
	L74V	-147.575
	Y115F	-160.561
	M184V	-152.784
DIDANOSINE	WILD	-160.085
	K65R	-145.129
	L74V	-144.942
EMTRICITABINE	WILD	-147.491
	K65R	-153.996
	M184V	-141.148
LAMIVUDINE	WILD	-148.334
	K65R	-153.174
	M184V	-142.283
	M184I	-150.474
	WILD	-146.143

TENOFOVIR	K65R	-153.447
	K70E	-148.196
STAVUDINE	WILD	-149.533
	M41L	-149.139
	D67N	-150.558
	K70R	-148.456
	L210W	-156.946
	T215V	-156.7
	K219Q	-151.074
ZIDOVUDINE	WILD	-150.795
	M41L	-162.146
	D67N	-148.077
	K70R	-145.474
	L210W	-151.471
	T215V	-154.441
	K219Q	-152.203

Docking was carried out to evaluate the effect of mutations in protein with drug interaction. It was reported that mutant protein would lead to loss or no interaction with drugs which has been shown by docking analysis (Table 1).

The free energies of the docking complexes are analysed. The drug Abcavir on mutant Y115F showed positive effect (-160.561) and showed least impact on wild type. The drug Didanosine on wild type showed positive effect (-160.085) and showed least impact on L74V mutant. The drug Emtricitabine on K65R mutant showed positive effect (-153.447) and showed least impact on M184V mutant. The drug Lamivudine on K65R mutant showed positive effect (-153.174) and showed least impact on M184V mutant. The drug Tenofovir on K65R mutant showed positive effect (-153.447) and showed least impact on wild type. The drug Stavudine on L210W mutant showed positive effect (-156.946) and showed least impact on K70R mutant. The drug Zidovudine on T215V mutant showed positive effect (-154.441) and showed least impact on K70R mutant.

This has to be correlated with the real effect on the patients. This serves a best model for evaluating the impact of mutations in changing the folding of protein and further its effect on the inhibition. This is shown by superimposing the wild type structures and mutant structures.

4. CONCLUSION

It was concluded that M184V mutations has least impact than that of other mutations and Y115F mutation has great impact on the drug interaction with Abcavir. This study helps in exploiting the said drug as an effective drug target. The intake of this retroviral drug by an HIV infected person over a long time may combat the multiplication of the virus in an individual. It has to be further correlated by clinical trials.

5. REFERENCES

[1] Sarafianos SG, Das K, Tantillo C, Clark AD Jr,Ding J, Whitcomb JM, Boyer PL, Hughes SH and Arnold E. (2001). Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBOJ.* 15;20(6):1449-61.

[2] Matteo Landriscina, Annarita Fabiano, Settimia Altamura, Cinzia Bagalà, Annamaria Piscazzi, Alessandra Cassano, Corrado Spadafora, Francesco Giorgino, Carlo Barone and Mauro Cignarelli. (2005). Reverse Transcriptase Inhibitors Down-Regulate Cell Proliferation in Vitro and in Vivo and Restore Thyrotropin Signaling and Iodine Uptake in Human Thyroid Anaplastic Carcinoma. *Endocrine Care,* 90 (10): 5663.

[3] Matthee G, Wright A D, Konig G M.. (1999). HIV reverse transcriptase inhibitors of natural origin. *Planta Med*, 65(6):493-506.

[4] www.expasy.org/prosite

[5] Guex N and Peitsch MC. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis,* 18(15):2714-23.

[6] Levin, J.M., Robson, B. & Garnier, J., An algorithm for secondary structure determination in proteins based on sequence similarity, FEBS Lett. 1986, *205*, 303-308.

[7] Geourjon C and Deleage G. (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput Appl Biosci,*11(6):681-4.

[8] Robert D. Busam, Ann-Gerd Thorsell, Alex Flores, Martin Hammarstrom, Camilla Persson, and B. Martin Hallberg, (2006). First Structure of a Eukaryotic Phosphohistidine Phosphatase. *The Journal of Biological Chemistry,* Vol. 281, NO. 45, pp. 33830 –33834.