

# Modeling of human CCR5 as target for HIV-I and virtual screening with marine therapeutic compounds

Selvaraj Manikandan<sup>1,\*</sup> and B. K. Malik<sup>1</sup>

<sup>1</sup>Institute of Genomics and Integrative Biology, Delhi-110 007, India;  
Selvaraj Manikandan\* - Email: smanii@hotmail.com, Phone: 91 9994475169; \* Corresponding Author

received July 29, 2008, accepted September 21, 2008; published October 25, 2008

### Abstract:

Marine derivatives are of great pharmaceutical interest as inhibitory compound and search of bioactive compounds from Marine organism which is relatively new to medicinal chemistry. Our main aim in the study is to screen possible inhibitors against CCR5 which acts as co-receptor M-tropic HIV-1, through virtual screening of 122 Marine derived compounds from various organisms known to have biological activity. Homology Model of CCR5 was constructed using MODELLER and the Model was energy minimized and validated using PROCHECK to obtain a stable structure, which was further used for virtual screening of Marine derived compounds through molecular Docking studies using GOLD. The Docked complexes were validated and Enumerated based on the GOLD Scoring function to pick out the best Marine inhibitor based on GOLD score. Thus from the entire 122 Marine compounds which were Docked, we got best 4 of them with optimal GOLD Score. (LAMIVUDINE: 45.0218, BATZELLINE-D: 44.3852, ACYCLOVIR: 43.1362 and THIOOACETAMIDE: 42.7412) Further the Complexes were analyzed through LIGPLOT for their interaction for the 4 best docked Marine compounds. Thus from the Complex scoring and binding ability its deciphered that these Marine compounds could be promising inhibitors for M-tropic HIV-1 using CCR5 as Drug target yet pharmacological studies have to confirm it.

**Keywords:** CCR5; marine derivatives; HIV-I; GOLD

### Background:

In AIDS therapy, there are currently a number of compounds available for multiple targets approved by the FDA and in clinical trials, e.g. protease inhibitors, reverse transcriptase inhibitors (NRTI, NNRTI) and CCR5, CCR4, fusion inhibitors, among others [1-4]. CCR5 is a member of the serpentine-receptor superfamily [5] and is expressed by monocytes, memory T lymphocytes, and preferably by Th1 cells and NK cells [6-8]. CCR5 act as a co-receptor for macrophage tropic(R5) strains of HIV-1 [9, 10] via interacting with viral envelope glycoprotein gp120, causing envelope fusion and viral entry [11-13]. Certain individuals who are homozygous for an internal 32-bp deletion in the gene encoding CCR5 produce a truncated form of the receptor that is not expressed on the cell surface, providing a high degree of resistance to M-tropic strains of HIV-1 [14, 15]. The physiological Ligand of CCR5 has been shown to block the co-receptor activity of M-tropic HIV-1 [11, 16-18].

In contrast, the viral chemokine vMIP-II [19], an antagonist of CCR5, has been shown to have limited virtue as an inhibitor of co-receptor activity [20]. This variation may be ascribed to the binding of these Ligands to different regions of CCR5, indicating the importance of the

N-terminal extracellular region (N-ter) and the second extra cellular loop (ECL2) in co-receptor activity [21, 22]. Both the N-ter and the body of CCR5 have been shown to affect the co-receptor activity of CCR5 [23].

The N-ter domain significance has been shown in several studies, which entangle specific acidic and aromatic residues in co-receptor activity [24-26]. Substitution of the N-ter regions of chemokine receptors, such as CCR1 or CCR2b, with that of CCR5 has been shown to allow virus fusion and entry [23, 27-30]. However, replacement of the N-ter region of CCR5 with that of chemokine receptors CCR1 or CCR2b has not been found to abolish co-receptor activity considerably [23, 27-30], indicating that both regions play vital roles in co-receptor activity. Ligand binding and HIV-1 entry are also affected by the posttranslational sulfation of Y10 and Y14 of the N-ter, which have been shown to bind gp120 at micromolar affinities when included in synthetic peptides [31, 32], contrasted by a complete lack of binding of the non-sulfated forms. Thus, two regions of CCR5 have been postulated to exhibit important functions in cell fusion, the N-ter domain, and the extracellular loops [29]. Transmembrane-spanning domains 4 [33] and 5 [34] have been shown to play a role in co-receptor function, but it is unclear if these domains play a role in direct interaction

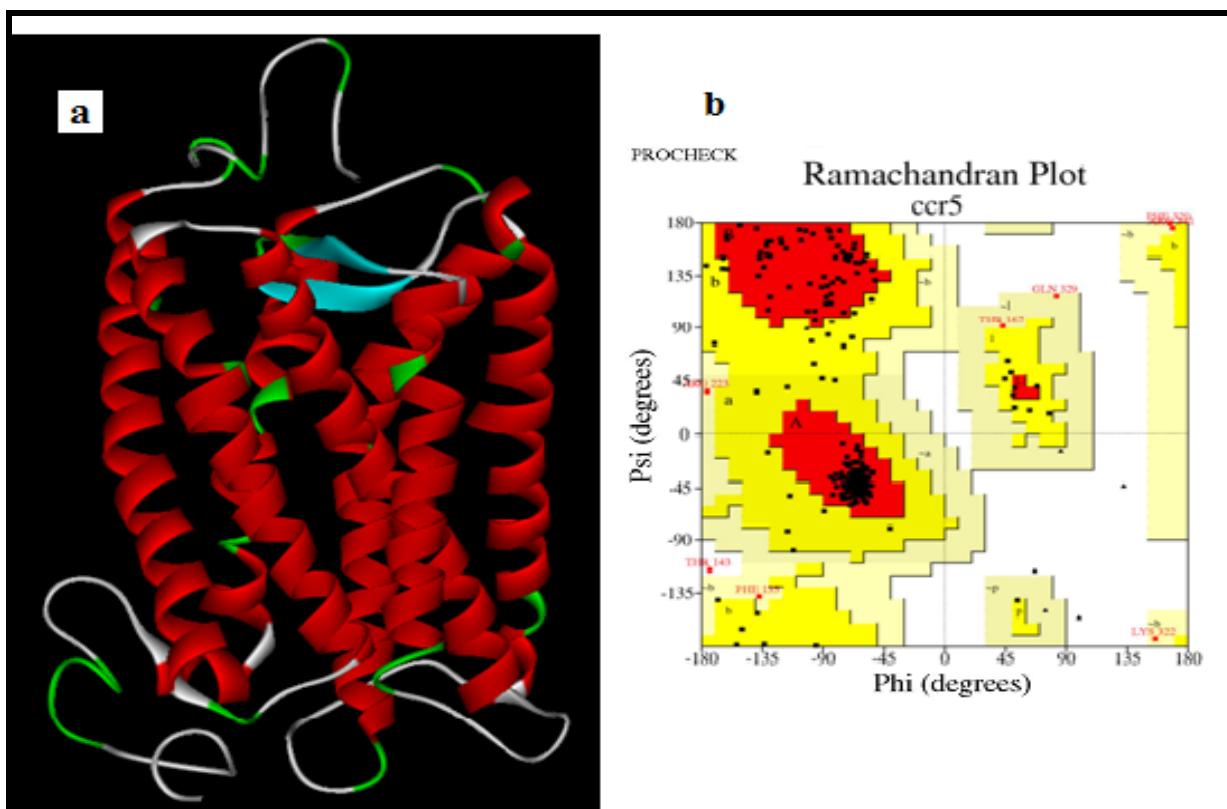
with gp120, or are important for regulating receptor conformation.

A small molecule inhibitor of CCR5 co-receptor function, TAK-779 [35], has been shown to require residues mapped to TM1, 2, 3, and 7 [36]. Other compounds such as SCH-C, SCH-351125, and SCH-350581, as well as members of another set of chemically unrelated small molecule inhibitors, 2-aryl-4-(piperidin-1-yl) butanamines and 1,3,4-trisubstituted pyrrolidines, appear to utilize a non-identical overlapping binding site involving TMs 2, 3, 6, and 7 [37,38]. Thus, it is likely that these CCR5 antagonists have an allosteric effect on receptor conformation, in contrast to direct blockade of the binding site for physiologic and pathologic Ligands.

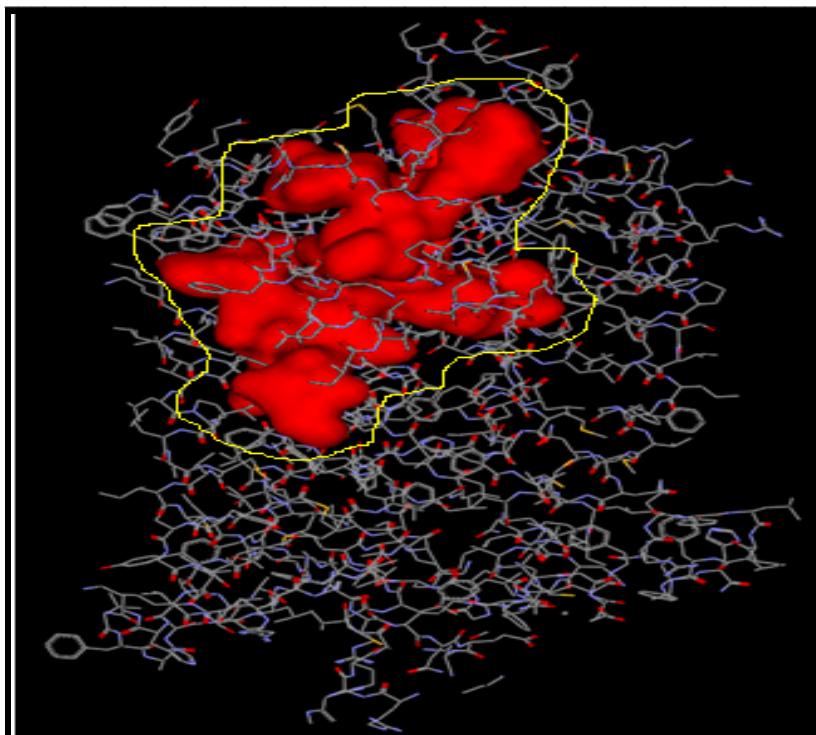
Similar to plant derived compounds for the therapeutic uses, Marine derived compounds do also have similar kind of effect from the pharmacological point of perspective and hence it's been used in cure for many deadly diseases. Marine sources were not utilized to any extent until the mid 20<sup>th</sup> century the last 30 years has seen an explosive growth of natural products from Marine organisms and these products are invariably described with some sort of

biological activity. The search of bioactive compounds from Marine organisms relatively a new field. However, the biodiversity of Marine environment far exceeds that of its terrestrial counter partner, So the oceans represents an enormous resource for new biologically Active compounds. There are 15 Marine natural products currently in clinical trials (Modern Drug discovery 2002) up to date. Our research work focus on virtual screening of such Marine compounds targeting HIV-1.

The parameters used for genetic algorithm were population size (100), selection pressure (1.1), number of operations (100,000), number of islands (5), niche size (2), migrate (10), mutate (95) and cross-over (95). The default speed selection was used to avoid a potential reduction in docking accuracy. Fifty genetic algorithm runs with default parameter settings were performed without early termination. To estimate the protein-Ligand complexes, the scoring function, the entire sets of 122 Marine therapeutic compounds were docked against Modeled CCR5 using GOLD [Genetic Optimization of Ligand Docking] version 2.1.2 was used to Screen for the possible inhibitory Marine compounds against CCR5 using GOLD Score. [44]. LIGPLOT [45] was constructed for analyzing the interaction with in the Best Docked complex.



**Figure 1:** (a) The final model of stable structure of CCR5; (b) Validation of the model using Ramachandran plot computed with the PROCHECK program with 86.4% of the residues in favored and allowed regions.



**Figure 2:** Identified active site of CCR5 where the inhibitors are likely to bind

### Methodology:

#### Homology modeling of CCR5

The Homology Model of Human CCR5 with Swiss-Prot ID: P51681 was generated using MODELLER [39] with ClustalW alignment of Bovine Rhodopsin sequence [accession number P02699] and its crystal structure [PDB: 1F88] as Template [40] chosen from PDB BLAST hit. The obtained Model was validated using PROCHECK [41] and final Energy minimized using CHARMM to obtain stable structure for further studies.

#### Active site identification

After the final Model was built, the possible Active sites of CCR5 were identified using Discovery Studio 2.0 [42] using the option Find site from Receptor Cavities.

#### Ligand generation and optimization

Marine Therapeutic compounds were gathered from Marine bioactive database [43] and 3D structures were generated using Discovery Studio 2.0[42] further the compounds were submitted for Ligand Minimization using CHARMM Forcefield with Smart Minimizer Algorithm Steps 1000 and Optimized.

#### Virtual screening of marine compounds against CCR5 using molecular docking

The Ligand s, including all hydrogen atoms, were built and minimized using Discovery studio 2.0 and GOLD [Genetic

optimization for Ligand Docking] version 2.1.2 [44] was used for Docking the entire 122 set of Marine compounds [43]. The Active site was defined as the collection of protein residues enclosed within a 15 Å radius sphere. The annealing parameters for Van der Waals and hydrogen bonding were set to 4.0 Å and 2.5 Å respectively.

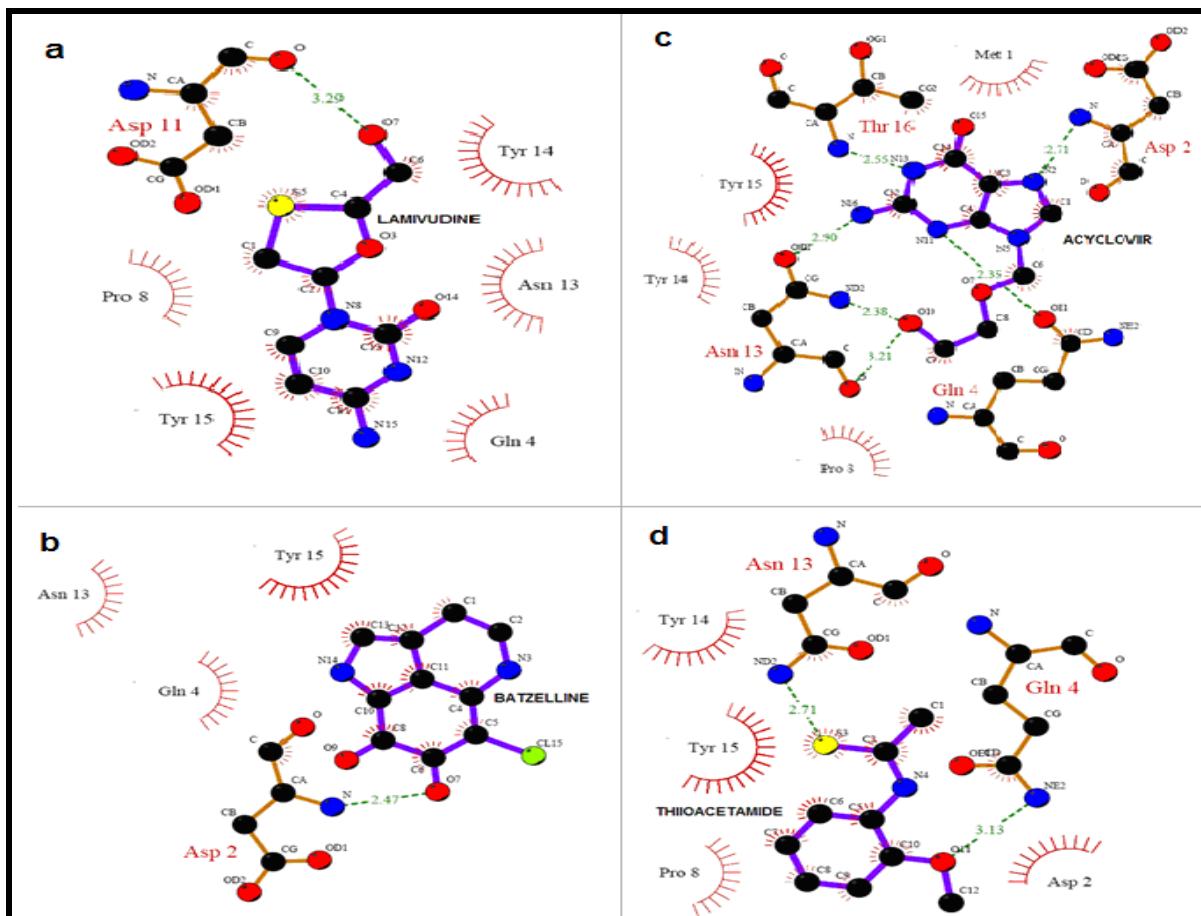
#### Results and discussion:

A confidence level sequence identity gives a reliable alignment between the target sequence and template structure. Our PDB BLAST Hit for CCR5 target sequence gave a best hit of Rhodopsin crystal structure (PDB: 1F88) .The Initial Model thus generated using MODELLER. Structurally conserved regions (SCRs) for the Model and the template were determined by multiple sequence alignment. Coordinates from the reference protein (1F88) to the SCRs, structurally variable regions (SVRs), N-terminal and C-terminal were assigned to the target sequence based on the satisfaction of spatial restraints. All side chains of the Modeled protein were set by rotamers. Thus, the obtained Model was refined by Energy minimization using CHARMM force field and the final stable structure of CCR5 is obtained (Figure 1).

After the refinement process, validation of the Model was carried out using Ramachandran plot calculations computed with the PROCHECK program. Altogether 86.4% of the residues were in favored and allowed regions (Figure 1). From

previous experimental evidence and from Figure 1 we come know that CCR5 receptor has 7 transmembrane helices with extracellular N-terminal end and intracellular C-terminal loop where in the extracellular loop plays an important role in the binding Gp120 to the co-receptor for M-tropic HIV-1 entry and which in turn also inhibitory site for the Antagonist. For the built CCR5 final Model the possible Active sites were identified using Discovery Studio 2.0. Fourteen possible Active sites were obtained using the option Find site from Receptor Cavities. These pockets were compared with the Active site of the template (Bovine Rhodopsin PDB: 1f88) and by careful analysis of

the entire predicted 14 Active sites, we found that 12<sup>th</sup> predicted Active site (shown in Figure 2) with residues Asp2, pro8, Asn13, Tyr14, Tyr15, Thr16, Gln21 of CCR5 were similar to that of the catalytic site in 1f88 and Therefore its chosen as the most biologically favorable site for Docking study and other Identified Active sites were neglected. (Figure 2) A set of 122 Marine Bioactive compounds were gathered from Marine bioactive database and 3D structures were generated using Discovery Studio 2.0 further Energy Minimized using CHARMM Forcefield and using Ligand's preparation option the best conformation were prepared for all 122 Marine compounds.



**Figure 3:** Receptor – ligand interaction is shown. **(a)** LAMIVUDINE O7 hydrogen bonded with extra cellular loop residue Asp11 OH at a distance of 3.29 Å and Gln4, Pro8, Asn13, Tyr14 and Tyr15 are non-ligand residues involved in hydrophobic contact(s). **(b)** BATZELLINE O7 hydrogen bonded with Extracellular loop residue Asp2 NH at a distance of 2.88 Å and Gln4, Asn13 and Tyr15 are non-ligand residues involved in hydrophobic contact(s). **(c)** ACYCLOVIR O10 hydrogen bonded with NH and with OH of Asn13 at a distance of 2.38 Å and 3.21 Å. Like wise NH hydrogen bonded with Gln4 at distance of 2.35 Å where as Met1, Pro8, Thy14 and Thy15 are non-ligand residues involved in hydrophobic contact(s). **(d)** THIOOACETAMIDE S3 hydrogen bonded with N of Asn13 at distance of 2.71 Å and OH hydrogen bonded with N of Gln4 at distance of 3.13 Å and NH hydrogen bonded with N at a distance of 2.71 Å where Asp2, Pro8, Thy14 and Thy15 are non-ligand residues involved in hydrophobic contact(s).

## Docking of marine compounds to CCR5

Docking of all 122 Marine compounds with Human CCR5 Model was performed using GOLD, which was set to 50 cycles of run without constraints between the Ligand s and the specific amino acids of the pocket. The algorithm exhaustively searches the entire rotational and translational space of the Ligand with respect to the receptors. The flexibility of the Ligand is given by dihedral angle variations. The various solutions evaluated by a score, which is equivalent to the absolute value of the total energy of the Ligand in the protein environment. The best Four Docking solution's GOLD score for each Compound were considered. It's noted that all the GOLD scores are very similar in all the top three solutions for each Marine compound. The GOLD score and RMSD differences between the consecutive top three runs of each best Docked Marine inhibitor is zero and this indicates the stabilization of particular inhibitor at the Active site. Out of 122 Marine compounds submitted for GOLD Docking run against Modeled Human CCR5 and we obtained 4 Marine compounds with optimal solution with Gold score Shown in Table [Rest of data not shown] (Table 1 in Supplementary material).

To understand the interaction between CCR5 and the four Marine compounds, Complexes were generated using the GoldMine suite for the outputs of Gold solutions. LIGPLOT were run for all the four complexes. It is evident from the analysis of the Docked complex that Marine inhibitors are located in the center of the Active site, and is stabilized by hydrogen bonding interactions and are bound in between 1<sup>st</sup> and 4<sup>th</sup>  $\alpha$ -helices Extracellular loop arresting the co-receptor activity of CCR5 for the entry of M-tropic HIV-1. (Figure 3) Through the interaction analysis, we know that Gln4, Asn13 and Thy14 are important anchoring residues for CCR5 and are the main contributors towards compound interaction. Though the interaction energy does not include the contribution from the water or the extended compound structure, this preliminary data along with the list of hydrogen bond interactions between the compounds and the Active site residues clearly supports that Gln4, Asn13 and Thy14 are more preferred residues in binding.

## Conclusion:

CCR5 is one of the most recent potent Drug target for HIV1. In this work, we have constructed a 3D Model of CCR5, using the MODELLER software and obtained a refined Model after energy minimization. The final refined Model was further assessed by PROCHECK program, and the results show that the Model was stable and reliable. The stable Model was further used for Virtual Docking of Marine Compounds. Docking results indicate that out of 122 Marine compounds, there were Four inhibitory compounds for CCR5 as Target for HIV-I. As it's well known, hydrogen bonding plays an important role for the structure and function of biological molecules, especially for Inhibition in a complex. Thus our study confirms LAMIVUDINE, BATEZELLINE-D, ACYCLOVIR and

THIOACETAMIDE are potential inhibitor for CCR5 as target for HIV-I forming a hydrogen bonding and with non-bonded interaction to act as a drug candidate yet Pharmacological study will yet confirm it to be promising.

## Acknowledgment:

This work was supported by CSIR (Council of Scientific and Industrial Research).

## References:

- [01] M. A. Noor *et al.*, *AIDS*, 18: 2137 (2004)
- [02] T. M. Dando and L. J. Scott., *Drugs*, 65: 285 (2005) [PMID: 15631548]
- [03] A. S. Veiga *et al.*, *J. Am. Chem. Soc.*, 127: 14758 (2004)
- [04] K. Princen *et al.*, *J. Virol.*, 78: 12996 (2004) [PMID: 1524989]
- [05] M. Samson *et al.*, *Biochemistry*, 35: 3362 (1996) [PMID: 8639485]
- [06] B. Lee *et al.*, *Proc Natl Acad Sci.*, 96: 5215 (1999) [PMID: 10409908]
- [07] R. Bonecchi *et al.*, *J Exp Med.*, 187: 129 (1998) [PMID: 950053]
- [08] F. Sallusto *et al.*, *J Exp Med.*, 187: 875 (1998) [PMID: 9500790]
- [09] R. I. Connor *et al.*, *J Virol.*, 67: 1772 (1993) [PMID: 8095306]
- [10] H. Schuitemaker *et al.*, *J Virol.*, 66: 1354 (1992) [PMID: 1738194]
- [11] G. Alkhatib *et al.*, *Science*, 272: 1955 (1996) [PMID: 8658171]
- [12] A. Trkola *et al.*, *Nature*, 384: 184 (1996) [PMID: 8906796]
- [13] L. Wu *et al.*, *Nature*, 384: 179 (1996) [PMID: 8906795]
- [14] M. Samson *et al.*, *Nature*, 382: 722 (1996) [PMID: 8751444]
- [15] R. Liu *et al.*, *Cell*, 86: 367 (1996) [PMID: 8756719]
- [16] H. Deng *et al.*, *Nature*, 381: 661 (1996) [PMID: 8649511]
- [17] F. Cocchi *et al.*, *Science*, 270: 1811 (1995) [PMID: 8525373]
- [18] M. Jansson *et al.*, *Proc Natl Acad Sci.*, 93: 15382 (1996) [PMID: 8986820]
- [19] P. S. Moore *et al.*, *Science*, 274: 1739 (1996) [PMID: 8939871]
- [20] C. Bosshoff *et al.*, *Science*, 278: 290 (1997) [PMID: 9323208]
- [21] J. M. Navenot *et al.*, *J Mol Biol.*, 313: 1181 (2001) [PMID: 11700073]
- [22] H. Wu *et al.*, *Nature*, 387: 527 (1997) [PMID: 9168119]
- [23] J. Rucker *et al.*, *Cell*, 87: 437 (1996) [PMID: 8898197]
- [24] C. Blanpain *et al.*, *J Biol Chem.*, 274: 34719 (1999) [PMID: 10574939]
- [25] T. Dragic *et al.*, *J Virol.*, 72: 279 (1998) [PMID: 9420225]
- [26] M. Farzan *et al.*, *J Virol.*, 72: 1160 (1998)

- [27] R. E. Atchison *et al.*, *Science*, 274: 1924 (1996) [PMID: 8943208]
- [28] P. D. Biéniasz *et al.*, *EMBO J.*, 16: 2599 (1997) [PMID: 9184207]
- [29] B. J. Doranz *et al.*, *J Virol.*, 71: 6305 (1997) [PMID: 9261347]
- [30] L. Picard *et al.*, *J Virol.*, 71: 5003 (1997) [PMID: 9188565]
- [31] E. G. Cormier *et al.*, *Proc Natl Acad Sci.*, 97: 5762 (2000) [PMID: 10823934]
- [32] M. Farzan *et al.*, *J Biol Chem.*, 275: 33516 (2000) [PMID: 10938094]
- [33] S. J. Siciliano *et al.*, *J Biol Chem.*, 274: 1905 (1999) [PMID: 9890944]
- [34] Z. Wang *et al.*, *J Biol Chem.*, 274: 28413 (1999) [PMID: 10497202]
- [35] M. Baba *et al.*, *Proc Natl Acad Sci.*, 96: 5698 (1999) [PMID: 10318947]
- [36] T. Dragic *et al.*, *Proc Natl Acad Sci.*, 97: 5639 (2000) [PMID: 10779565]
- [37] F. Tsamis *et al.*, *J Virol.*, 77: 5201 (2003) [PMID: 12692222]
- [38] L. A. Castonguay *et al.*, *Biochemistry*, 42: 1544 (2003) [PMID: 12578367]
- [39] J. Sali *et al.*, *J. Mol. Biol.*, 234: 779 (1993) [PMID: 8254673]
- [40] K. Palczewski *et al.*, *Science*, 289: 739 (2000) [10926528]
- [41] R. A Laskowski *et al.*, *J. Appl. Cryst.*, 26: 283 (1993)
- [42] <http://accelrys.com/products/discovery-studio/>
- [43] <http://www.niot.res.in/m5/mbic/pg/mcdb/antiviralrlist.php>
- [44] G. Jones *et al.*, *J. Mol. Biol.*, 267: 727 (1997) [PMID: 9126849]
- [45] A. C. Wallace *et al.*, *Prot. Eng.*, 8: 127 (1995)

**Edited by P. Kangueane**

**Citation:** Manikandan *et al.*, Bioinformation 3(2): 89-94 (2008)

**License statement:** This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

## Supplementary material

S. No	Marine compound	IUPAC name	Molecular formula and Mol. Wt.	Gold score
1	LAMIVUDINE	4-amino-1-[(2R,5S)-2-(hydroxymethyl)-1,3-oxathiolan-5-yl]pyrimidin-2-one	C8H11N3O3S 229.257g/mol	45.0218
2	BATZELLINE -D	Pyrrolo[4,3,2-de]quinoline-7,8-dione, 6-chloro-1,3,4,5-tetrahydro-	C10H7ClN2O2 222.628g/mol	44.3852
3	ACYCLOVIR	2-amino-9-(2-hydroxyethoxymethyl)-3H-purin-6-one	C8H11N5O3 225.205g/mol	43.1362
4	THIOACETAMIDE	Ethanethioamide	C <sub>2</sub> H <sub>5</sub> NS 75.132800 g/mol	42.7412

**Table 1:** Table shows the best docked Marine compounds and its GOLD score.