J Biochem Tech (2019) 10(3): 34-39
ISSN: 0974-2328

In silico Drug interaction studies on HIV Integrase

Rajasekhar Pinnamaneni*, Pravallika Tanneru, Navya Vaddavalli, Jahnavi Tatineni and Reddy Tharani

Received: 04 April 2019 / Received in revised form: 03 September 2019, Accepted: 08 September 2019, Published online: 26 September 2019

Abstract

The infection of HIV causes AIDS, which is a prevalent problem today. Retroviral infection happens upon the integration of a copy of the DNA of the viral genome into the host DNA at any location in the genome but is preferred in certain regions of the chromatin as a necessary step in the replication. The challenge in the treatment of retroviral infections is that the integrated viral DNA is sustainably maintained and replicated along with cellular DNA through the cell division. Although wide prospects have been made in antiviral therapy of HIV, the integrated virus persists in long-lived cells and eradication is an elusive goal. Integrase protein is the key enzyme in the virus that integrates retroviral DNA into the host genome. As the rapid development of drug resistance in the existing drug classes, discovery of new targets is the need of the hour. The three major enzymes of HIV-1 integrase IN (HIV-1 IN), protease, and reverse transcriptase have been of special interest. Integration of viral genetic material with the host genome is an important step in the process of viral replication catalyzed by HIV-1 IN. 6 chloro 2 oxo 4 phenyl 1,2 dihydroquinolin-3-yl acetic acid, Quinoline class of inhibitors of HIV-1 integrase, was used in the present study to analyze the drug-protein interaction.

Keywords: HIV-1 IN, 6 chloro 2 oxo 4 phenyl 1,2 dihydroquinolin-3-yl acetic acid, energy minimization.

Introduction

Retroviruses synthesize a DNA copy of their RNA genome after infection of the host cell. Integration of the viral DNA into host DNA is a fundamental step in the replication cycle of retroviruses especially HIV-1 (Brown, 1997; Asante-Appiah and Skalka, 1997; Hindmarsh and Leis, 1999). Integration of viral DNA and its transcription results in the progeny virions with the integration of RNA genome (Embaby et al., 2018). The replication cycle is completed upon integration into the host genome; virions bud from the cell surface and multiply by infecting the uninfected cells. HIV genome gets integrated into the DNA of the host CD4 cell by HIV-1 integrase enzyme (Mohy et al., 2018). Integrase strand transfer inhibitors (INSTIs) are a class of antiretroviral HIV drugs that block the integration as a crucial step in the HIV life cycle (Alsamarrai et al., 2018). Three domains are present in HIV-1 integrase, whose structures were determined based on the susceptibility of the linker regions to proteolysis (Engelman and Craigie, 1992), functional studies (Engelman and Craigie, 1992; Engelman *et al.,* 1993; van Gent *et al.,* 1993), and the x-ray crystallography or NMR (Amasha, 2018). The development of effective inhibitors of reverse transcriptase and protease in HIV replication is an effective antiviral therapy for the treatment of AIDS. The drugs targeted to integrase would be complementary to reverse transcriptase and protease. No drugs have been developed to target integrase (Craigie, 2001). The lack of good lead compounds is a bottleneck of drug development. Many compounds inhibit integrase but majority of them lack selectivity and inhibit other enzymes, too. Upon targeting integrase, HIV replication can be inhibited in cell culture (Hazuda et al., 2000). The structure of an inhibitor complexed with the active site of integrase has been determined (Goldgur et al., 1999). The binding modes in the absence and presence of DNA substrate are not identical and this issue led to suggest the selectivity of the DNA strand transfer reaction and the much higher affinity of these inhibitors for integrase in the presence of viral DNA (Espeseth et al., 2000). Hence, it is necessary to determine the structure of these inhibitors in complex with an integrase active site engaged with DNA to understand the interaction in detail in providing a platform for novel drug design. The basic knowledge in overcoming the mechanism of retroviral integration contributes to fight against AIDS. The present study aimed to study about 6 chloro 2 oxo 4 phenyl 1,2 dihydroquinolin-3-yl acetic acid, Quinoline class of HIV-1 integrase inhibitors in order to analyze the drug-protein interaction as a new series of HIV-1 IN inhibitors.

Materials and Methods

Rajasekhar Pinnamaneni*, Pravallika Tanneru, Navya Vaddavalli, Jahnavi Tatineni and Reddy Tharani

Department of Biotechnology, Koneru Lakshmaiah Education Foundation, Greenfields, Vaddeswaram, Guntur-522502, Andhra Pradesh, India.

***Email:** pinnamaneniraj @ yahoo.com,

Data Set

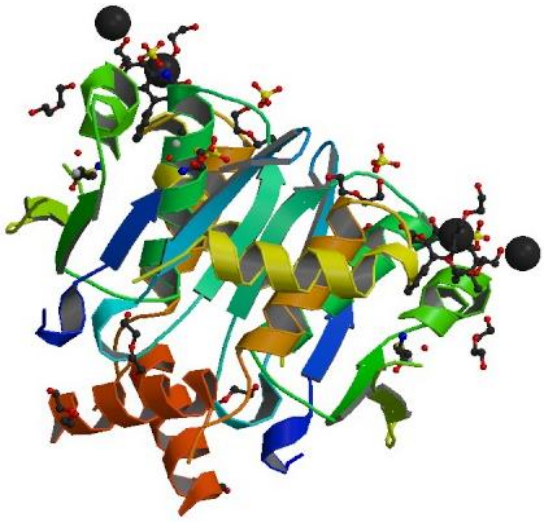

Fig.1: 3LPT|HIV integrase-Human immunodeficiency virus 1

The three-dimensional, known protein-ligand complexes identified experimentally were obtained from Protein Data Bank (PDB). Biological databases like PDB (protein data bank), pubchem, can be used for the same. The protein sequence of Protein 3LPT|HIV integrase-Human immunodeficiency virus 1, expressed in *E.coli* was retrieved (Fig.1). The Prosite tool was used to analyze the parameters of the protein at www.expasy.org/prosite. All the protein parameters concerning amino acid composition, secondary structure prediction, hydrophobicity, isoelectric point, etc. were analyzed. The final model was used for docking by subjecting the PDB model to several repeated cycles of energy minimization using SPDBV software (Guex and Peitsch, 1997).

ProtParam

From the protein sequence, the computation of various physicochemical properties can be deduced with no additional information on the protein. The protein in the form of a raw sequence or as specified with a Swiss-Prot/TrEMBL accession number or ID. White space and numbers were ignored. By default, the complete sequence was analyzed. The choice included the possibility to enter the start and end positions in two boxes and also the selection of mature domains, peptides, or chains from the Swiss-Prot feature table. ProtParam was used to compute the parameters including GRAVY (grand average of hydropathicity), theoretical pI, aliphatic index, molecular weight, instability index, estimated half-life, extinction coefficient, atomic composition, and amino acid composition. The theoretical pI and Molecular weight were calculated to compute pI/Mw. The atomic and amino acid compositions were self-explanatory (Gasteiger et al., 2005).

The extinction coefficient is the light absorbed by a protein at a specific wavelength. It is useful to have an estimation of this extinction coefficient of a protein using a spectrophotometer during protein purification. By knowing the amino acid sequence of a protein we can estimate the molar extinction coefficient of it. The following equation was used to compute the extinction coefficient of a native protein in water:

E (Prot) = Numb (Tyr)×Ext (Tyr) + Numb (Trp) × Ext (Trp) + Numb (Cystine) × Ext (Cystine)

The absorbance (OD) is calculated by using the following formula:

Absorb (Prot) = E (Prot) / Molecular weight

Based on the above equations, two values were obtained by ProtParam, which both were for proteins measured at 280 nm in water. The 1st one showed the computed value based on the assumption that all pairs of Cys residues were form cystines (i.e. all cysteine residues appear as half cystines), and the 2nd one assuming all Cys residues are reduced (i.e. assuming that no cysteine appears as half cystine). Experiences show that the computations for proteins containing Trp residues are quite reliable; however, there may be more than 10% error for proteins without Trp residues.

SOPMA

SOPMA (Self-Optimized Prediction Method with Alignment) (Levin et al., 1986) is a secondary structure prediction based on multiple alignments, by which 69.5% of amino acids of the secondary structure for a 3-state description (coil, beta-sheet, and alpha-helix) in a whole database containing 126 chains (<25% identity) of non-homologous proteins were predicted. 82.2% of residues for 74% of co-predicted amino acids can be predicted by SOPMA. Sequences alignment information belonging to the same family is taken into account in SOPMA (Geourjon and Deleage, 1995).

Inhibitors used in the present study
6 chloro 2 oxo 4 phenyl 1,2 dihydroquinolin-3-yl acetic acid

Quinoline class of HIV-1 integrase inhibitors has an ancillary aromatic ring and a quinoline subunit, which are linked by functionalized spacers including urea, hydrazide, amide, and 1-hydroxyprop-1-en-3-one moieties be potent HIV-1 IN inhibitors *in vitro*. Many polynucleotidyl transferases that are related to HIV-1 IN, have 2 divalent metal cations, with 4 Å separation in their active sites. Potential IN inhibitors based on the quinoline substructure link to an aryl nucleus with different hydroxy substitution patterns and easily interact at this site. Since styrylquinoline derivatives inhibit the disintegration assay performed by the active site with deletion mutation, they target the IN core domain.

These styrylquinoline derivatives target the core domain of IN since they inhibited the disintegration assay by the active site containing deletion mutation (Bénard et al., 2004).

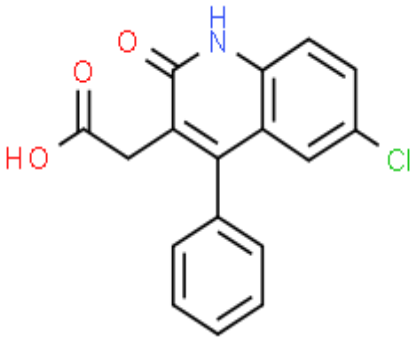

Fig 2: 6 chloro 2 oxo 4 phenyl 1,2 dihydroquinolin-3-yl acetic acid

Generation of 6 chloro 2 oxo 4 phenyl 1,2 dihydroquinolin-3-yl acetic acid 3D structure

ChemSpider server was used to generate 3D structure of 6 chloro 2 oxo 4 phenyl 1,2 dihydroquinolin-3-yl acetic acid (Fig.2) and SMILES notation of 6 chloro 2 oxo 4 phenyl 1,2 dihydroquinolin-3-yl acetic acid was found in ChemSpider that was used to translate this 3D structure by SMILES translator online server, employing JAVA based structure applet in ChemSpider. Then the Universal Force Field (UFF) molecular mechanics was used to reframe the geometric and the final geometry was done by Semi empirical Quantum mechanics method (QM). Finally, 6 chloro 2 oxo 4 phenyl 1,2 dihydroquinolin-3-yl acetic acid structure was saved as MOL file for further docking procedure. Molecular Formula is $C_{17}H_{12}ClNO_3$ with an average mass of 313.735 Da and Monoisotopic mass of 313.050568 Da. The ChemSpider ID is 804714 (Ayers, 2012).

Molecular Docking

Molecular docking tool helps in computer-assisted design and the structural molecular biology to identify the binding compatibility of the ligand with a three-dimensional structure enabling this to find out the intermolecular complex formed between the two molecules (Lengauer and Rarey, 1996). Based on the binding strength and complex energy evaluation of the binding affinity between the protein and the ligand can be predicted using the scoring function by molecular docking (Kitchen et al., 2004) that may result in the inhibition or activation of the enzyme. It is used as the most prevalent method in drug design because it can predict the accuracy and the conformation of small molecule ligands with the target binding site (Meng et al., 2011). Based on the affinity of the ligand-receptor complex, molecular docking algorithms also execute strings of the docked compounds (Huang and Zou, 2010). For designing the drug of the therapeutic target, a detailed understanding of the principles gives the outline of different interactions for which, a variety of docking methods are employed. Scoring function dictates the optimal binding. A pose is generated during the computational docking, scored, and compared with the previous pose, which is decided based on the score whether to consider or delete the previous pose. This scoring iteration has to be continued until it achieves an endpoint and identifies the perfect fit.

ArgusLab software packages are used to implement different molecular docking algorithms according to various search methods, (Thompson, 2004). In this investigation, molecular docking was performed using ArgusLab and SwissDock software. ArgusLab is a genetic algorithm-based molecular modeling, graphics, and drug designing program implemented with exhaustive search methods, the AScore scoring function and Argus Dock docking engine able to perform molecular structure visualization and molecular geometry calculations (Wang et al., 2004). ArgusLab is consisting of a grid resolution of 20Å. The ligand and binding site groups were created using ChemSpider with ID 804714. Then the 3D structure of 6 chloro 2 oxo 4 phenyl 1,2 dihydroquinolin-3-yl acetic acid was opened in the ArgusLab that was then introduced and docking calculation was allowed to run using shape-based search algorithm and scoring function. Docking was performed with HIV-1 IN protein.

Results and Discussion

Protein Sequence

>3LPT:A|PDBID|CHAIN|SEQUENCE
GSHMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIE
AEVIPAETGQETAYFLLKLAGRWPVKTVHTDNGSNFTSTT
VKAACWWAGIKQEFGIPYNPQSQGVIESMNKELKKIIGQV
RDQAEHLKTAVQMAVFIHNKKRKGGIGGYSAGERIVDIIA
TDIQTKE

Amino acid Composition

Length = 166 amino acids
Molecular Weight = 18170.78 Daltons
Theoretical pI: 7.21

Ala (A)	14	8.4%
Arg (R)	4	2.4%
Asn (N)	5	3.0%
Asp (D)	6	3.6%
Cys (C)	3	1.8%
Gln (Q)	10	6.0%
Glu (E)	11	6.6%
Gly (G)	17	10.2%
His (H)	7	4.2%
Ile (I)	15	9.0%
Leu (L)	8	4.8%
Lys (K)	13	7.8%
Met (M)	3	1.8%
Phe (F)	4	2.4%
Pro (P)	5	3.0%
Ser (S)	8	4.8%
Thr (T)	11	6.6%
Trp (W)	4	2.4%
Tyr (Y)	4	2.4%
Val (V)	14	8.4%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 17
Total number of positively charged residues (Arg + Lys): 17

Atomic composition

Carbon	C	810
Hydrogen	H	1278
Nitrogen	N	224
Oxygen	O	239
Sulfur	S	6

Formula: $C_{810}H_{1278}N_{224}O_{239}S_6$
Total number of atoms: 2557
Ext. coefficient 28085
Abs 0.1% (=1 g/l) 1.546, assuming all pairs of Cys residues form cystines
Ext. coefficient 27960
Abs 0.1% (=1 g/l) 1.539, assuming all Cys residues are reduced
Estimated half-life:
The N-terminal of the sequence considered as G (Gly).
The estimated half-life is >10 h (*Escherichia coli*, *in vivo*), >20 h (yeast, *in vivo*), 30 h (mammalian reticulocytes, *in vitro*).
Instability index:
The instability index (II) is computed to be 36.41. This classifies the protein as stable.
Aliphatic index: 86.93
GRAVY (Grand average of hydropathicity): -0.208

SOPMA

The parameters of secondary structure by SOPMA are
Similarity threshold: 8
Window width: 17
Number of states: 4
The secondary structure analysis was conducted by various tools such as SOPMA. It contains various types of secondary structures like 72 Alpha helices, 24 extended strands, 13 Beta turns and 57 Random coils (Fig 3).

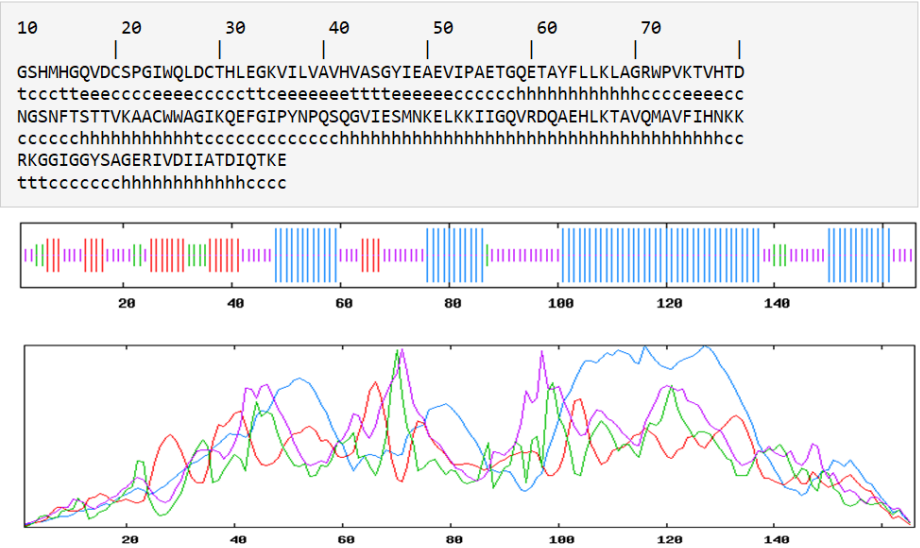

Fig. 3: SOPMA Analysis

Virtual Docking Results

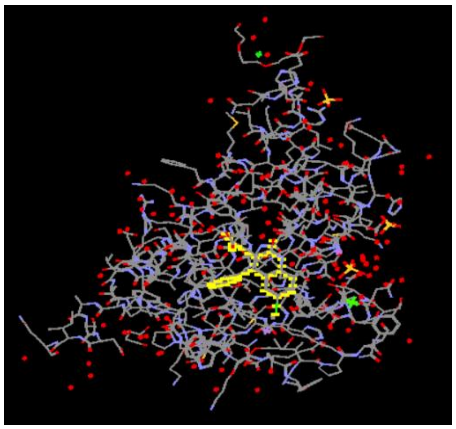

Fig.4: Ligand docked onto HIV-1 Integrase 3LPT

Docking was carried out to evaluate the drug-protein interactions. It was reported that the binding of the drug molecule to the HIV Integrase protein would lead to loss or no activity, which can be assessed based on the minimization of energy run for 200 cycles and the final Self-Consistent Field (SCF) energy status was -135.2363727469 au or -84862.1817 kcal/mol (Thompson and Schenter, 1995; Thompson, 1996) and the final geometrical energy was 0.0624080103 au or 39.1617 kcal/mol (Thompson and Schenter, 1995; Thompson, 1996). The bonded topology parameters between 6 chloro 2 oxo 4 phenyl 1,2 dihydroquinolin-3-yl acetic acid-HIV-1IN indicated that the total number of bonds was 24 with 34 bond angles and the number of dihedral angles was 45 and the important torsions were 8. The number of candidate poses found was 8. three final unique configurations were considered by clustering the final poses. The best ligand pose energy was -1.96224 kcal/mol.

This is related to the real effect on the patients and serves the best model to evaluate the effect of mutation in changing the protein folding and its effect on the inhibition.

Conclusion

It was concluded that drug-protein interaction of 6 chloro 2 oxo 4 phenyl 1,2 dihydroquinolin-3-yl acetic acid-HIN IN 1 was stable and therefore the inhibition of HIV-1 IN activity can be maximized.

Acknowledgments

The authors thank the management of Koneru Lakshmaiah Education Foundation for providing us the infrastructural facilities.

Conflict of Interest

This statement is to certify that all authors have seen and approved the manuscript being submitted. We warrant that the article is the Authors' original work. We warrant that the article has not received prior publication and is not under consideration for publication elsewhere. On behalf of all Co-Authors, the corresponding Author shall bear full responsibility for the submission.

References

Alsamarrai, A. S., Abdulla, N. H., & Aldoori, M. K. (2018). Synthesis and Characterization of 2-((4R, 4aR, 5aS, 6S)-1, 3-dioxo-3, 3a, 4, 4a, 5, 5a, 6, 6a-octahydro-4, 6-ethenocyclopropa [f] isoindol-2 (1H)-yl)-N-(Substituted Phenyl) Acetamides Derivatives Anticipated to Inhibit HIV-1 Activity. INTERNATIONAL JOURNAL OF PHARMACEUTICAL AND PHYTOPHARMACOLOGICAL RESEARCH, 8(5), 7-11.

Amasha, R. H. (2018). Use Of Nmr To Determine Compatible Solutes In Halophilic Bacteria Isolated From Highly Saline Areas. Pharmacophore, 9(5), 50-60.

Asante-Appiah, E., & Skalka, A. M. (1997). Molecular mechanisms in retrovirus DNA integration. Antiviral research, 36(3), 139-156.

Ayers, M, ChemSpider: The Free Chemical Database, Reference Reviews. 2012, 26(7):45-46).

Bénard, C., Zouhiri, F., Normand-Bayle, M., Danet, M., Desmaële, D., Leh, H., ... & Le Bret, M. (2004). Linker-modified quinoline derivatives targeting HIV-1 integrase: synthesis and biological activity. Bioorganic & medicinal chemistry letters, 14(10), 2473-2476.

Brown, P. O., Coffin, J., Hughes, S., & Varmus, H. (1997). Integration. Retroviruses. Cold Spring Harbor Laboratory Press, Cold spring Harbor, 161-204.

Craigie, R. (2001). HIV integrase, a brief overview from chemistry to therapeutics. Journal of Biological Chemistry, 276(26), 23213-23216.

Embaby, M., Shaker, O., Abd El Aziz, G., Rashad, A., & Yousri, A. (2018). Toll like receptor 7 & 8 gene variations and sustained virological response in Hepatitis C Virus Patients treated with interferon. Journal of Advanced Pharmacy Education & Research| Jul-Sep, 8(3).

Engelman, A. L. A. N., & Craigie, R. O. B. E. R. T. (1992). Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. Journal of virology, 66(11), 6361-6369.

Engelman, A., Bushman, F. D., & Craigie, R. (1993). Identification of discrete functional domains of HIV‐1 integrase and their organization within an active multimeric complex. The EMBO Journal, 12(8), 3269-3275.

Espeseth, A. S., Felock, P., Wolfe, A., Witmer, M., Grobler, J., Anthony, N., ... & Cole, J. L. (2000). HIV-1 integrase inhibitors that compete with the target DNA substrate define a unique strand transfer conformation for integrase. Proceedings of the National Academy of Sciences, 97(21), 11244-11249.

Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In The proteomics protocols handbook (pp. 571-607). Humana press.

Geourjon, C., & Deleage, G. (1995). SOPMA: Significant improvement in protein secondary structure prediction by c prediction from alignments and joint prediction. CABIOS, 11, 681-684.

Goldgur, Y., Craigie, R., Cohen, G. H., Fujiwara, T., Yoshinaga, T., Fujishita, T., ... & Davies, D. R. (1999). Structure of the HIV-1 integrase catalytic domain complexed with an inhibitor: a platform for antiviral drug design. Proceedings of the National Academy of Sciences, 96(23), 13040-13043.

Guex, N., & Peitsch, M. C. (1997). SWISS‐MODEL and the Swiss‐Pdb Viewer: an environment for comparative protein modeling. electrophoresis, 18(15), 2714-2723.

Hazuda, D. J., Felock, P., Witmer, M., Wolfe, A., Stillmock, K., Grobler, J. A., ... & Miller, M. D. (2000). Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. Science, 287(5453), 646-650.

Hindmarsh, P., & Leis, J. (1999). Retroviral DNA integration. Microbiol. Mol. Biol. Rev., 63(4), 836-843.

Huang, S. Y., & Zou, X. (2010). Advances and challenges in protein-ligand docking. International journal of molecular sciences, 11(8), 3016-3034.

Kitchen, D. B., Decornez, H., Furr, J. R., & Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. Nature reviews Drug discovery, 3(11), 935.

Lengauer, T., & Rarey, M. (1996). Computational methods for biomolecular docking. Current opinion in structural biology, 6(3), 402-406.

Levin, J. M., Robson, B., & Garnier, J. (1986). An algorithm for secondary structure determination in proteins based on sequence similarity. FEBS letters, 205(2), 303-308.

Meng, X. Y., Zhang, H. X., Mezei, M., & Cui, M. (2011). Molecular docking: a powerful approach for structure-based drug discovery. Current computer-aided drug design, 7(2), 146-157.

Mohy, A. A., Aljanaby, A. A. J., & Al-Hadraawy, S. K. (2018). Serum concentrations of CD4+ and CD8+ in patients infected with scabies caused by Sarcoptes scabiei. Journal of Advanced Pharmacy Education & Research| Jan-Mar, 8(1).

Thompson, M. A. (1996). QM/MMpol: A consistent model for solute/solvent polarization. Application to the aqueous solvation and spectroscopy of formaldehyde, acetaldehyde, and acetone. The Journal of Physical Chemistry, 100(34), 14492-14507.

Thompson, M. A. (2004, March). Molecular docking using ArgusLab, an efficient shape-based search algorithm and the AScore scoring function. In ACS meeting, Philadelphia (Vol. 172, p. 42).

Thompson, M. A., & Schenter, G. K. (1995). Excited states of the bacteriochlorophyll b dimer of Rhodopseudomonas viridis: a QM/MM study of the photosynthetic reaction center that includes MM polarization. The Journal of Physical Chemistry, 99(17), 6374-6386.

van Gent, D. C., Vink, C., Groeneger, A. A., & Plasterk, R. H. (1993). Complementation between HIV integrase proteins mutated in different domains. The EMBO Journal, 12(8), 3261-3267.

Wang, R., Fang, X., Lu, Y., & Wang, S. (2004). The PDBbind database: Collection of binding affinities for protein− ligand complexes with known three-dimensional structures. Journal of medicinal chemistry, 47(12), 2977-2980.