

Identification of STAT5a Inhibitors for Breast Cancer Treatment Through In silico Approach

Bavya Chandrasekhar^{1†}, Dona Samuel Karen¹ and Veena Jaganivasan¹

^{1†}Department of Genetic Engineering, Computational Biology Lab,
School of Bioengineering, SRM Institute of Science and Technology, Kattankulathur-603203,
Chengalpattu district, Tamil Nadu, India

Abstract

Female breast cancer is the fifth highest cause of mortality. Breast cancer is the most prevalent type of cancer in women globally, while it can also affect men. STAT5A plays a role in its development and progression. Given that activation of STAT5a is frequently linked to the growth and progression of tumors, STAT5a has been identified as a possible target for the therapy of several cancers. STAT5A, in particular, has proven to be overexpressed in various breast cancer cell lines and tumors, and it has been associated to the promotion of tumour cell proliferation and survival. STAT5A inhibition has been shown in vitro and in vivo to reduce the development of breast cancer cells. As a result, we have screened compounds from the FDA database that might serve as potential inhibitors of STAT5a through virtual screening, docking, DFT and MD simulation approaches. The drug Nilotinib has shown promising results inhibiting STAT5a. Further, in-vitro analysis will be carried forward to understand the anti-cancer activity.

Keywords: Breast cancer, STAT5a, inhibitors, cytotoxicity, chronic myelogenous leukemia

(Received December 13, 2023; Revised December 20, 2023; Accepted February 29, 2024)

1. Introduction

Breast cancer is a form of cancer that begins in the breast tissue. It is the most frequent cancer in women globally, while it can also occur in men. Breast cancer can occur in various regions of the breast, including the ducts, lobules, and connective tissue. Although the specific origin of breast cancer is unknown, variables such as age,

genetics, hormone changes, and lifestyle practices can all raise the risk of acquiring the disease^[1]. STAT5a a member of the STAT family is found to play a major role in the oncogenesis of breast cancer.

STAT5a has been reported to be activated in specific forms of breast cancer, such as hormone receptor-positive breast cancer. STAT5a signalling is required for normal mammary gland development and differentiation, but it can potentially contribute to the development and

[†]Corresponding author: bavyanair16@gmail.com

progression of breast cancer. STAT5a is triggered by growth factors and cytokines such as prolactin and growth hormone^[2], which are involved in the regulation of breast cell development and differentiation. Activation of STAT5a in breast cancer cells has been found to boost cell proliferation and survival, as well as resistance to chemotherapy and hormone therapy. STAT5a has been discovered as a chemoresistance inducer that modulates the expression of ABCB1 in breast cancer and can be targeted by pimozone, an FDA-approved psychiatric medication. STAT5a and ABCB1 are required for doxorubicin resistance in breast cancer cells, and STAT5a regulates ABCB1 transcription^[3]. Furthermore, high levels of STAT5a have been linked to a worse prognosis in breast cancer patients.

Drug repurposing can take many different forms, such as discovering new therapeutic uses for drugs that have already received approval for other uses, discovering creative drug combinations that could be used to treat a particular ailment, or modifying existing drugs to increase their efficacy or reduce their side effects. A great way to skip a few phases in the medication development process is through drug repurposing^[4]. Here, we have screened and shortlisted ten drugs from 1643 FDA approved drugs. Compounds that successfully bind to the target protein are found with the use of in silico screening. They are selected by means of their binding energies. Molecular docking is the primary strategy used in virtual screening. Based on how well they interact with the highlighted residues, drugs or small molecule compounds are selected to attach to the active site of target molecules. The different chemical

bonds formed and energies exchanged determine the binding's degree. To visualise the various docking poses, visualisation tools like PyMol and others are utilised. Following the separation of the desired molecules, they are further optimised through the use of Molecular Simulations, DFT optimisation, and other methods.

2. Materials and Method

2.1. Target Identification

Stat5a is a protein that is present in all human cells and plays an important role in many cellular processes, including cell proliferation, differentiation, and survival especially of the mammary gland^[5]. Additionally, overexpression of Stat5a has been observed in certain types of cancer, such as breast cancer and leukemia^[6]. In these cases, elevated levels of Stat5a may contribute to the development and progression of the cancer by promoting cell growth and survival. The PDB database was the key internet resource used in the study. The target proteins STAT5a (PDB ID: 7TVA) and the three - dimensional structures of the lead drug were obtained from the PubChem database.

2.2. Virtual screening

AutoDock Vina is a common molecular docking software that can be used for virtual screening^[7]. Virtual screening is a computational tool for identifying possible ligands that potentially bind to a certain receptor or protein target. The steps include preparing the protein and ligand structures

i.e. the 3D structures of the protein and ligand are downloaded in the PDB format, removing of water molecules from the protein, adding missing atoms or residues, assigning protonation states; setting up of the docking parameters i.e. setting up boundaries for the ligand to bind to the protein; running the docking simulation and finally analysing the results^[8].

2.3. DFT analysis using Gaussian

This experiment used conceptual DFT, which uses the electron structure to investigate the chemical behaviour. Gauss view was used to analyse the data in this experiment^[9]. The top two compounds were studied to learn more about the chemical bonds produced by the protein. Total energy, highest occupied molecular orbital (HOMO), lowest unoccupied molecular orbital (LUMO) absolute hardness, electronegativity, energy gap, electrophilicity index, global softness, molecular dipole moment and chemical potential were among the data calculated^[10].

2.4. MD simulation using Gromacs

Molecular Dynamics (MD) simulations are used to study the behavior of biomolecules, such as proteins and nucleic acids, at the atomic level. GROMACS is a popular software package for performing MD simulations, particularly for biomolecular systems. The basic steps for performing an MD simulation using GROMACS include system preparation; generation of a simulation box; defining of the simulation parameters i.e. Temperature - 300K , Pressure - 1-1.5Bar ; energy minimization of the system; equilibrating the system with ions i.e. Na and Cl; running the

production MD simulation and analyzing of the results^[11].

3. Results

3.1. Virtual screening and Molecular docking

Virtual screening often involves the use of molecular docking or other computer approaches to predict the binding affinity of small compounds for a target protein. The virtual screening procedure can be used to filter vast databases of chemicals, lowering the amount of compounds that need to be physically examined in a laboratory. We have shortlisted 10 drugs from a pool of 1643 FDA drugs on the basis of their binding affinities. The binding affinities were studied through docking and analysing with the help of Autodock 1.5.6. The drug with the highest binding affinity was found to be Nilotinib with B.A of -10.76(kcal/mol).

Also the molecular interactions of the lead hits i.e. the hydrogen and hydrophobic interactions were studied using Pymol software.

3.2. DFT Calculation

Density functional theory was calculated for hit compounds that have good binding affinity. Every compound has different optimized structures. The red and blue colors represent HOMO and LUMO states. The energy gap is calculated by using the difference between E HOMO and E LUMO. Nilotinib (0.35) have a low energy gap (ΔE). The low energy compounds are strong and stable. Nilotinib have low strength spacing,

resulting in a higher reactivity of the inhibitors to the protein. Electronegativity is a chemical effect that represents the course of molecules and draws electrons. The electronegativity value should be higher because only then the activity of the particular protein can be inhibited.

3.3. MD simulation

To understand the conformational changes in the protein with respect to the ligand molecules, we have carried out 100 ns molecular dynamics simulation for each designed molecules. The stability of the protein-ligand complex was evaluated using RMSD, RMSF and H-Bond plot.

The representation of RMSD plot indicates the STAT5a protein has undergone less conformational change during the simulation for all the complexes. We have also monitored the RMSF fluctuations of each residue in order to determine the effect of Ligand molecules on the protein residues dynamic. Hydrogen bond plays a very important role in stabilizing the protein-ligand complex. Hydrogen bond is responsible for drug specificity, metabolization, and adsorption in body. The number of hydrogen bonds between protein-ligand complexes which were formed during the 100 ns simulation.

the active site residues of STAT5a protein the leads drugs were short listed. The potency of these drugs was further validated by DFT and MD simulation studies. STAT5a plays an important role in the growth and regulation of mammalian cells - mammary gland development and lactogenesis and promotes tumour progression by regulating the expression of cell cycle and pro-inflammatory genes, which when overexpressed leads to mammary gland developmental defects and breast cancer. Identification of new lead compounds as STAT5a inhibitors may serve better treatment for breast cancer in future.

4. Discussion

Ten lead compounds for STAT5a inhibitor have been successfully identified through virtual screening and molecular docking approach. Based upon their interactions with

Reference

- [1] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.", *CA: A Cancer Journal for Clinicians*, Vol. 71, No. 3, pp. 209-249, 2021.
- [2] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.", *CA: A Cancer Journal for Clinicians*, Vol. 71, No. 3, pp. 209-249, 2021.
- [3] Li, Z., Chen, C., Chen, L., Hu, D., Yang, X., Zhuo, W., Chen, Y., Yang, J., Zhou, Y., Mao, M. and Zhang, X., "STAT5a confers doxorubicin resistance to breast cancer by regulating ABCB1.", *Frontiers in oncology*, p. 2761, 2021.
- [4] Pushpakom, S., Iorio, F., Eyers, P.A., Escott, K.J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., McNamee, C. and Norris, A., 2019. Drug repurposing: progress, challenges and recommendations. *Nature reviews Drug discovery*, Vol. 18, No. 1, pp.41-58, 2019.
- [5] Liu, X., Gallego, M.I., Smith, G.H., Robinson, G.W. and Hennighausen, L., "Functional rescue of Stat5a-null mammary tissue through the activation of compensating signals including Stat5b.", *Cell Growth and Differentiation-Publication American Association for Cancer Research*, Vol. 9, No. 9, pp. 795-804, 1998
- [6] Tan, S.H. and Nevalainen, M.T., "Signal transducer and activator of transcription 5A/B in prostate and breast cancers.", *Endocrine-Related Cancer*, Vol. 15, No. 2, p. 367, 2008.
- [7] Trott O., Olson A. J., "AutoDockVina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading.", *J. Comput. Chem.*, Vol. 31, No. 2, pp. 455-461, 2010. 10.1002/jcc.21334 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
- [8] Baba, N. and Akaho, E., "VSDK: Virtual screening of small molecules using AutoDock Vina on Windows platform.", *Bioinformatics*, Vol. 6, No. 10, p. 387, 2011.
- [9] Frisch M., Trucks G., Schlegel H., Scuseria G., Robb M., Cheeseman J., et al. (2016). Gaussian 16. (Version Revision B. 01)[Linux]. Wallingford CT: Gaussian. [Google Scholar]
- [10] Nagarajan, S.K., Babu, S., Sohn, H. and Madhavan, T., "Molecular-level understanding of the somatostatin receptor 1 (SSTR1)-ligand binding: a structural biology study based on computational methods.", *ACS omega*, Vol. 5, No. 33, pp. 21145-21161, 2020.
- [11] Beveridge, D.L., Barreiro, G., Byun, K.S., Case, D.A., Cheatham, T.E., Dixit, S.B., Giudice, E., Lankas, F., Lavery, R., Maddocks, J.H. and Osman, R., "Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides., I. Research design and results on d (CpG) steps.", *Biophysical journal*, Vol. 87, No. 6, pp.3799-3813, 2004.

Table. 1. FDA approved drugs and their interaction with STAT5a after virtually screening.

S.NO	COMPOUND	BINDING AFFINITY (kcal/mol)	HYDROGEN BONDING	HYDROPHOBIC INTERACTION
1	NILOTINIB	-10.76	Gly266, Gly265, Asn522, Ser521	Ala517, Lys516, Gly263, Asn486, Glu518, Val482, Tyr694
2	BICTEGRAVIR	-9.9	Lys460, Thr461, Leu462	Leu383, Val459, Lys256, Thr329, Arg493, Phe496
3	M_ENTRECTINIB	-9.8	Gly492, Arg493, Leu383	Ser463, Leu462, Pro495, Phe496, Lys256, Gln260
4	NALDEMEDINE	-9.8	Ser434, Arg560, Glu437	Ala427, Ser375, Lys-439, Val-435, Gln-377, Trp-555
5	DIHYDROERGOTAMINE	-9.7	Asn599, Ser543, Asn540, Gln535	Phe532, Gln602, Ser541, Asn539, Lys536, Asp-621
6	RIMEGEPANT	-9.7	Asp475, Val341, Gln340, Pro338, Gln-259	Thr346, Gly263, Gln337, Ala262, Lys514, Val482
7	VELPATASVIR	-9.7	Ser543, Gln535, Lys536, Trp566	Val531, Gln602, Asp621, Leu563, His544, Asn599
8	ACETYLDIGITOXIN	-9.6	Glu335, Gly263, Lys514, Gln340, Lys582, Asp475	Pro339, Val341, Lys336, Glu269, Thr479, Val482
9	ADAPALENE	-9.6	Asp621, Ser622, Arg618	Lys644, Leu643, Trp641, Thr628, Pro645, Ser620
10	AMCINONIDE	-9.6	Asn599, Gln535, Gly565	Lys536, Asn539, Ser541, Pro564, Gln602, Asn540

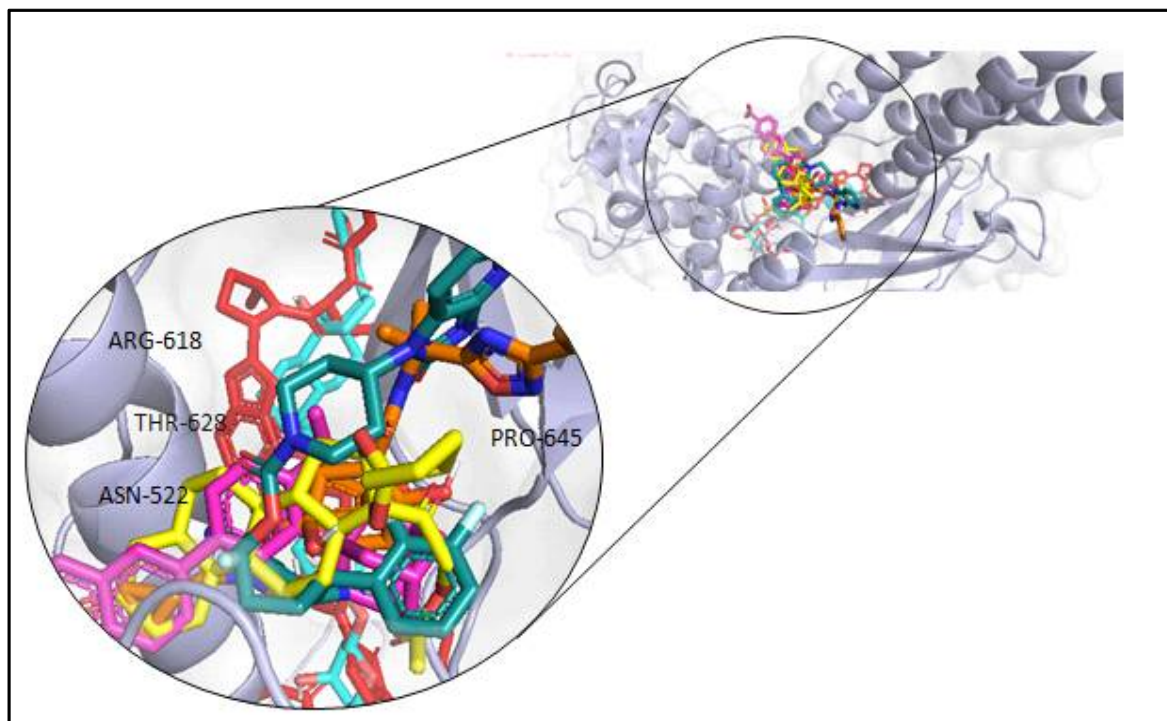
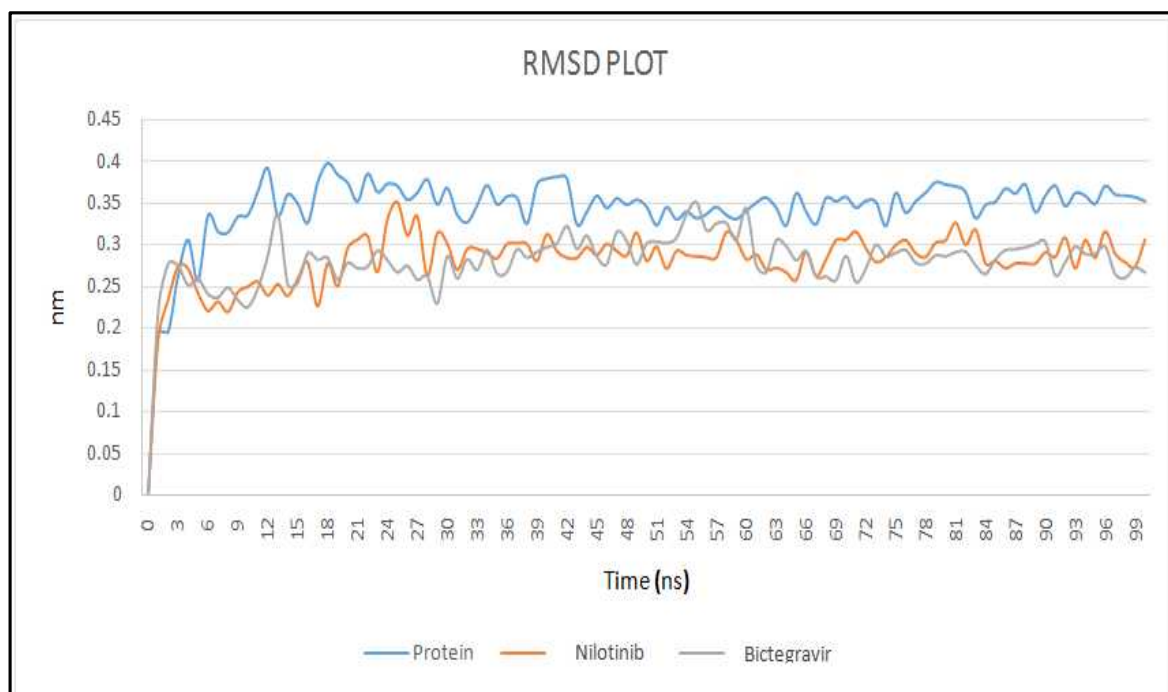
**Fig. 1.** Poses STAT5a with superimposed ligand structures generated using the pymol program.

Table. 2. Statistics of DFT based molecular descriptors of top hits to understand the chemical behaviour of the compounds

COMPOUND	(E γ) (in eV)	(Debye)	E _{HOMO}	E _{LUMO}	(ΔE)	(η)	(σ)	(γ)	(μ)	(ω)
NILOTINIB	-1826.6	6.43	-4.23	-3.88	0.35	0.17	2.85	-4.05	4.05	46.98
BICTEGRAVIR	-1390.13	3.89	-6.41	-2.74	3.67	1.83	0.27	-4.57	4.57	5.7
M_ENTRECTINIB	-51163.6	5.77	-5.97	-1.25	4.71	2.35	0.21	-3.61	3.61	4.6
NALDEMEDINE	-1480.6	10.98	-5.55	-5.1	0.45	0.22	2.22	-5.32	5.32	63.01
DIHYDROERGOTAMINE	-81337.2	10.39	-2.01	-1.68	0.33	0.16	2.98	-1.85	1.85	5.52
RIMEGEPANT	-53083.1	8.59	-6.04	-2.25	3.79	1.89	0.26	-4.14	4.14	1.09
VELPATASVIR	-49645.6	4.14	-5.63	-1.91	3.72	1.86	0.268	-3.77	3.77	3.82
ACETYLDIGITOXIN	-34695.1	5.107	-6.15	-1.15	5	2.5	0.2	-3.65	3.65	2.66
ADAPALENE	-35632.9	3.17	-5.71	-1.73	3.98	1.99	0.25	-3.72	3.72	3.47
AMCINONIDE	-39106.7	4.79	-5.99	-1.43	4.56	2.28	0.21	-3.71	3.71	3.01

**Fig. 2.** Change in RMSD values over 100ns to understand the stability of the protein- ligand complex

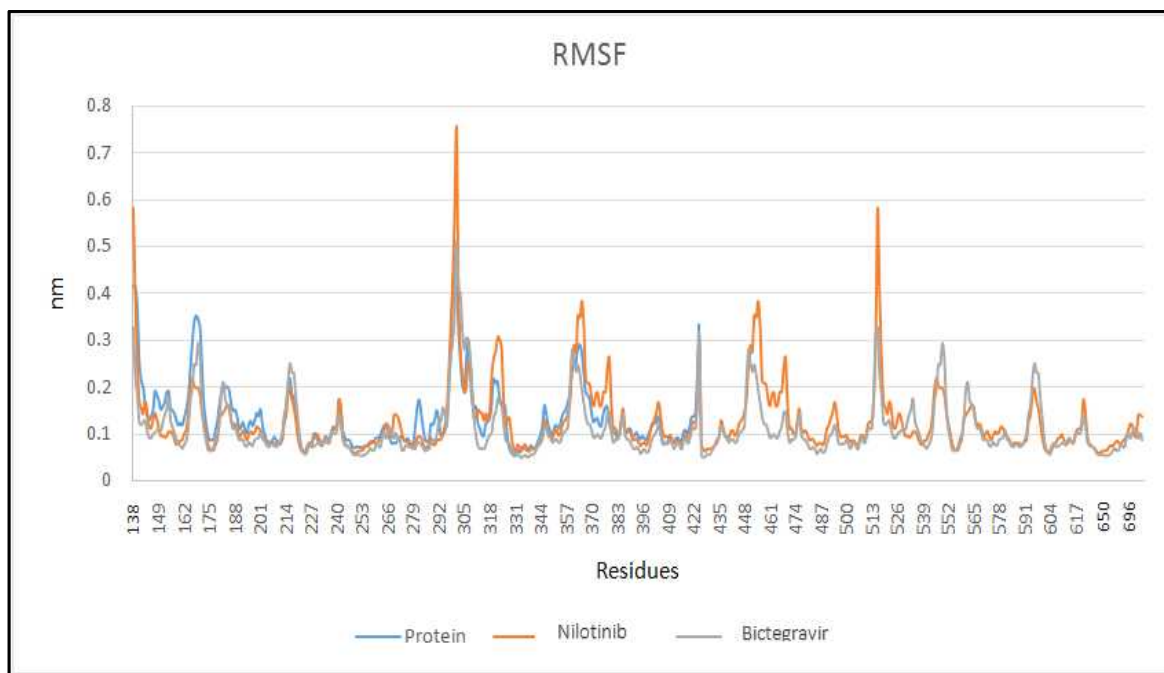


Fig. 3. Change in RMSF values of STAT5a residues over a period 100ns after binding with lead hits.

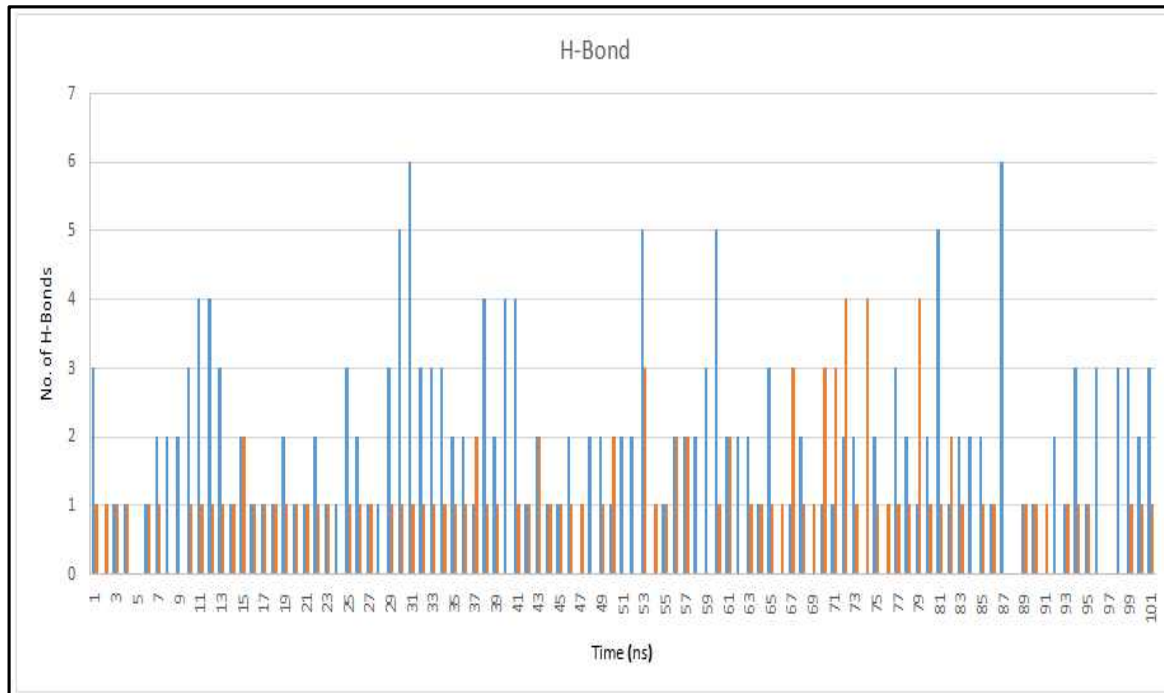


Fig. 4. No. of H-Bonds formed between the protein and ligand over 100ns to understand the interaction of the complex.