Research Article

Comparative Molecular Docking Studies Associated with a Series of Isatin-Oxime Derivates of RSV Fusion Inhibitors on STAT Protein

Saraswathi Jaggali[1], Roja Rani Anupalli[2]
[1]Department of Genetics, Osmania University, Hyderabad, India.
[2]Department of Genetics, Osmania University, Hyderabad, India.
***Corresponding author's E-mail:** Anupalli_rr@yahoo.com

Accepted on: 28-03-2013; **Finalized on:** 31-07-2013.

ABSTRACT

3D-QSAR analysis studies of substituted benzimidazole-isatin oximes inhibitors and along with COMFA and CoMSIA with Partial Least Squares (PLS) analysis were carried on SYBYL software. The atom and shape based root mean square alignment yielded the best predictive CoMFA model q^2cv= 0.783, r^2=0.969 with five components, standard error of estimate= 0.126 and while the CoMSIA model q^2cv=0.740, r^2=0.952 with six components, standard error of estimate= 0.160 respectively. Comparative molecular docking studies of Isatin-oxime were conducted on (RSV) and (STATS) proteins. The virtual mole dock values of Isatin-oximes derivatives were showed on PDB1G2C is -103.0967 and PDB1YVL -184.697. The most favorable docking interactions were observed at active site of RSV-Ser182 and STAT1-Glu 353, Gln 352, Glu 1352 and Leu1352. These Molecular docking approaches are commonly used in a modern drug design process to understand the drug-receptor interactions.

Keywords: Isatin-Oxime fusion Inhibitors, Molegro docking inhibitors, RSV and STAT Protein.

INTRODUCTION

Respiratory syncytial virus (RSV), a single-stranded RNA virus of negative genome polarity and is a member of the *Pneumo virus* genus of the *Para myxo virus* family. RSV was first shown to occur in humans in 1957, after being recovered from two infants hospitalized with severe lower Respiratory tract infections.[1] Today, RSV is recognized as the leading cause of virus-induced lower respiratory tract disease among infants and children's.[2] Most children's are infected with RSV before two years of age, re-infection is a common occurrence and morbidity due to Complications are high among premature infants and those with underlying cardiopulmonary problems. Moreover, Upper respiratory tract infection proceeds with severe nasal congestion and profuse rhinorrhea, advancing to a cough and pharyngitis. Progression to lower respiratory tract infection may follow, leading to pneumonia in the most serious cases.[3]

Stat Protein and Oncogenic

During the multistep process of tumor genesis, cells lose their normal ability to sense and repair DNA damage and to regulate cell cycle progression and directly to cell cycle checkpoint regulation or DNA repair, they contribute to tumor genesis through their intimate connection to growth factors like signaling, apoptosis and angiogenesis. In addition because these molecules play key roles in immune responses, defective STAT signaling can favor tumor development by compromising immune surveillance.[4] Signal transducer and activator of transcription (STAT) factors are implicated in programming gene expression in biological events as diverse as embryonic development, programmed cell death, organogenesis, innate immunity, adaptive immunity and cell growth regulation in organisms ranging from slime molds to insects to man.[5] Subversion of a cell's normal genetic program results in alterations in the expression patterns of genes involved in different facets of transformation, such as cell proliferation, anchorage-independent growth, survival, and morphological changes.[6] In contrast to normal signaling, in which STAT activation is rapid yet transient, constitutive signaling by STATs has been increasingly associated with malignant progression. Activation of STAT proteins requires tyrosine phosphorylation, and for some STAT family member's serine phosphorylation, identification of the kinases that are responsible for catalyzing STAT phosphorylation has yielded valuable insight into the molecular mechanisms involved in activation of STATs in human tumours.[6]

MATERIALS AND METHODS

Molecular structures and optimization

QSAR (Quantitative Structure -Activity Relationships)

The well-established approach to compare molecular properties and to determine their influence on the biological potency is QSAR. The QSAR and the Molecular Spreadsheet functionality in Sybyl software allow the chemists to establish models using 2D or 3D structures. The Molecular Spread sheet unites the results of applications from different molecular research components and is central to all kinds of molecular analyses capable of extracting relevant information from the bulk of molecular data. All statistical analysis methods including the PLS (Partial Least Squares) method, factor analysis or clustering techniques are applicable to any data set. Unlike 2D-QSAR, the CoMFA (Comparative Molecular Field Analysis) method directly takes the 3D molecular structures into account. COMSIA takes Steric and electrostatic fields are calculated for the

superimposed bioactive conformation of each molecule in the data set. The variance in these field data is used to explain the variance in the activity data by means of PLS and cross-validation statistics. The molecular regions that contribute to the activity variations can be identified and given as an indication of the pharmacophoric features. In addition, the predictive power of the QSAR correlations can be used to estimate the bioactivity of new compounds and to guide the modifications and refinements of existing structures to obtain a better activity profile.

Table 1: Side chains and R-Group of 76 compounds Isatin-Oxime etherderivatives[7]

S.No	Compound No	Side Chain	R-Group	S.No	Compound No	Side Chain	R-Group
1	11a	(CH2)2CH(Me)2	H	39	11an	–(CH2)3OMe	–(CH2)SO3H
2	11b	(CH2)2CH(Me)2	Me	40	11ap	–(CH2)3OMe	–CH(CO2H)CH2CO2H
3	11c	–(CH2)4OH	Me	41	11aq	–(CH2)2CHMe2	Ph
4	11d	–(CH2)4OAc	Me	42	11ar	–CH2)2CHMe2	4'-Ph-Br
5	11e	–(CH2)3OMe	Me	43	11as	–(CH2)2CHMe2	–CH2-Ph
6	11f	–(CH2)3OMe	Et	44	11at	–(CH2)2CHMe2	–CH2-Ph-4-CO2H
7	11g	–(CH2)3OMe	–CH2CH2F	45	11au	–(CH2)2NMe2	–CH -4'-Ph-CO2H
8	11i	–(CH2)3CN	–CH2CF3	46	11av	–(CH2)3OMe	–CH2-4'-Ph-CO2Me
9	11j	–(CH2)3OMe	i-Pr	47	11aw	–(CH2)2CHMe2	–CH2 -4'-Ph-CO2Me
10	11l	–(CH2)3OMe	–CH(CH2CH3)2	48	11ax	–(CH2)2NMe2	-CH2-4'-Ph-CO2Me
11	11m	–(CH2)3OMe	–CH2cPr	49	11ay	–(CH2)2NSO2Me	–CH2-4'-Ph-CO2Me
12	11n	–(CH2)3OMe	–CH2cBu	50	11az	–(CH2)3OMe	–CH2-4'-Ph-CONMe2
13	11o	–(CH2)3OMe	–4-Tetrahydro-2H-pyran	51	11ba	–(CH2)2CHMe2	–CH2-4'-Ph-CONMe
14	11p	–(CH2)3OMe	–Cyclo-hexyl	52	11bc	–(CH2)4OH	–CH2 -4'-Ph-SO2Me
15	11q	–(CH2)3OMe	–CH2-1- tetrahydrofuran	53	11bb	–(CH2)4OAc	–CH2-4'-Ph-SO2Me
16	11r	–(CH2)3OMe	–(CH2)2cHex	54	12j	–(CH2)3OMe	i-Pr
17	11s	–(CH2)3OMe	n-Pr	55	12i	–(CH2)3OMe	—CH(CH2CH3)2
18	11t	–(CH2)3OMe	–CH2CH2.CH2F	56	12p	–(CH2)3OMe	–cHex
19	11u	–(CH2)3OMe	n-Bu	57	18a	–(CH2)4–F	-H
20	11v	–(CH2)3OMe	n-Pentyl	58	18b	–(CH2)3–CN	-H
21	11w	–(CH2)3OMe	-CH2CH=CH2	59	18c	–(CH2)4 –OAc	-H
22	11x	–(CH2)3OMe	–(CH2)2CH=CH2	60	18d	–(CH2)4–F	–CH3
23	11y	–(CH2)3OMe	–(CH2)3 CH=CH2	61	18e	–(CH2)4–OH	–CH3
24	11z	–(CH2)3OMe	–(CH2)3C≡CH	62	18f	–(CH2)3–CN	–CH3
25	11aa	–(CH2)3OMe	–(CH2)3CN	63	18g	–(CH2)4–OH	-
26	11ab	–(CH2)3OMe	–CH2CH(OH)CH2CN	64	18h	–(CH2)4–F	–CH2CH2F
27	11ac	–(CH2)3OMe	–(CH2)4 OAc	65	18i	–(CH2)4–OH	–CH2CH2F
28	11ad	–(CH2)3OMe	CH2CONEt2	66	18j	–(CH2)4–OH	–CH2CH2F
29	11ae	–(CH2)3OMe	–CH2CONH2	67	18k	–(CH2)4–OAc	–CH2CH2F
30	11af	–(CH2)3CN	–CH2CONH2	68	18l	–(CH2)3 –CN	–CH2CH2F
31	11ag	–(CH2)3OMe	–(CH2)2NMe	69	18m	–(CH2)3–SO2Me	–CH2CH2F
32	11ah	–(CH2)3OMe	–(CH2)3NMe2	70	18n	–(CH2)3–SO2Me	–CH2CH2F
33	11ai	–(CH2)3OMe	–(CH2)2piperidine	71	18o	–(CH2)4 –OH	–CH2CF3
34	11aj	–(CH2)3OMe	–(CH2)3piperidine	72	18p	–(CH2)3 –CN	–CH2CF3
35	11ak	–(CH2)3OMe	–CH2-2-pyridine	73	18q	–(CH2)3–SO2Me	–CH2CH=CH2
36	11al	–(CH2)3OMe	CH2-3-pyridine	74	18r	–(CH2)3–SO2Me	–CH2-2-Pyridine
37	11m	–(CH2)3OMe	–CH2-4-pyridine	75	18t	–(CH2)3–SO2Me	–CH2-4'-Ph-SO2CH3
38	11ao	–(CH2)3OMe	–CH2CO2H	76	18u	–(CH2)3–CN	–CH2-4'-Ph-SO2CH3

Seventy six molecules selected (Table 1) for the present studies were taken from an earlier report with the structures of the compounds and their biological data. The IC50 values were converted to the corresponding pIC50 (-logIC50) and used as dependent variables in CoMFA and CoMSIA analysis. The pIC50 values span a range of 3-log units providing a broad and homogenous data set for 3D-QSAR study. The 3D QSAR models were generated using a training set of 56 molecules and predictive power of the resulting models was evaluated using a test set of 20 molecules. The test set compounds were selected manually such that the structural diversity and wide range of activity in the data set were included.

Alignment of Isatin–Oxime ether Derivatives

CoMFA results may be extremely sensitive to a number of factors such as alignment rules, overall orientation of the aligned compounds, lattice shifting step size and probe atom type. The accuracy of prediction of CoMFA models and the reliability of the contour models depend strongly on the structural alignment of the molecules. The molecular alignment was achieved by SYBYL routine database alignment method and selecting the common part. The most active compound (compound 11bb) was used as an alignment template and the rest of the molecules were aligned to it by using the common substructure. The aligned molecules were shown in the following (Figure 1). Molegro virtual docker is an integrated environment for studying and predicting how ligands interact with macromolecules. The identification of ligand binding modes is done by iteratively evaluating for a number cardiopulmonary disease.

RESULTS AND DISCUSSION

Molecular Docking

Docking studies were carried out using by Molegro virtual docker software and the Isatin –oxime derivaties 76 ligands given in (Table 1) were chem sketch on ACD lab software and which are all energy Minimized before loading into the docking wizard with protein PDB structures. The general information of two protein were shown in the (Table 2) the receptor- ligand interaction is modeled by a few special types of interactions were explained. These Are hydrogen bonds, metal - acceptor bonds and a few types of hydrophobic contacts. Interaction is modeled by an interaction

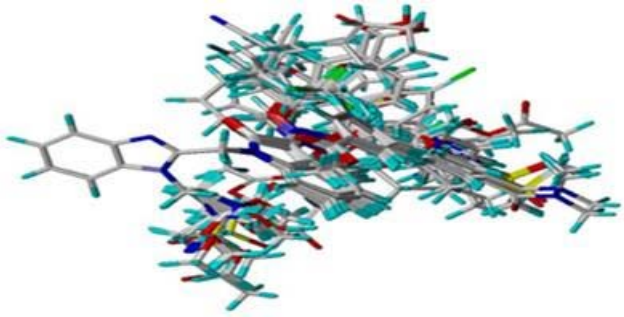

Figure 1: Alignment of Isatin-oxime compounds in Sybyl software

Center and an interaction surface located on a sphere around the center. molecular docking explaining about the Isatin–oxime ether derivatives which were obtained from the biologically active compounds tested with IC_{50} valves are taken for identification of the best inhibitor function on RSV protein which lead to best ligand design. The Protein Data Bank (PDB) is a repository for the 3-D structural data of large biological molecules, such as proteins, nucleic acids and X-ray crystallography. NMR Spectroscopy of respiratory syncytial virus Domain (PDB code 1G2C) and Signaling transducer and activator of transcription (PDB code 1YVL) is obtained from the protein data bank.

Table 2: Molecular Description RSV and STATS

Molecular Description		
Title	1G2C	1YVL
Classification	Viral Protein	Signaling Protein
Structural weight	126325.31	161965.14
Molecule	Fusion Protein (F)	Transcription
Chains	A, C, E, G, I, K, M, O, Q, S, U, W	A,B
Length	52	683
Experimental Description		
Method	X-Ray Diffraction	X-Ray
Resolution	0.233Å(obs.)	3
R-value	0.286	0.24

Active Site Residues Identification

Cavity detection method available in the Molegro Virtual Dockers was carried out using detects cavity option seen in (figure 2) and (figure 3). The maximum number of cavities that would return from this search was up to 5. The cavity detection based on expanded Van der Waals molecular surface option. The complete protein was scanned using the carbon probe with radius 1.2 A0. All the five cavities returned by MVD, the largest cavity

Steps in Methodology of Molegro Virtual Docker

1. Importing a protein PDB 1G2C and 1YVL File and chem sketched ligands file.
2. Protein preparation and detecting cavities of protein molecules are identifying for ligand Active site interaction given in the figure 2 &3.
3. Executing a docking set up through docking wizard panel.
4. Poses of protein-ligand complex is obtained after docking process with their specific Mol dock and Re rank scores displayed in output fit. The mol dock results of two docking protein were explained in the Fig 5 &6. The protein active site interacting with the Isatin –Oxime ether derivate showing in the (figure 7).

as 406.016area/m^3. The number of residues lining the cavity 1 is compared with the experimentally defined Active site residues from literature. It has been observed that nearly 50% of the residues were similar to the findings from literature. Hence it can be stated that the cavity detection algorithm of MVD is able to detect the active site.[8]

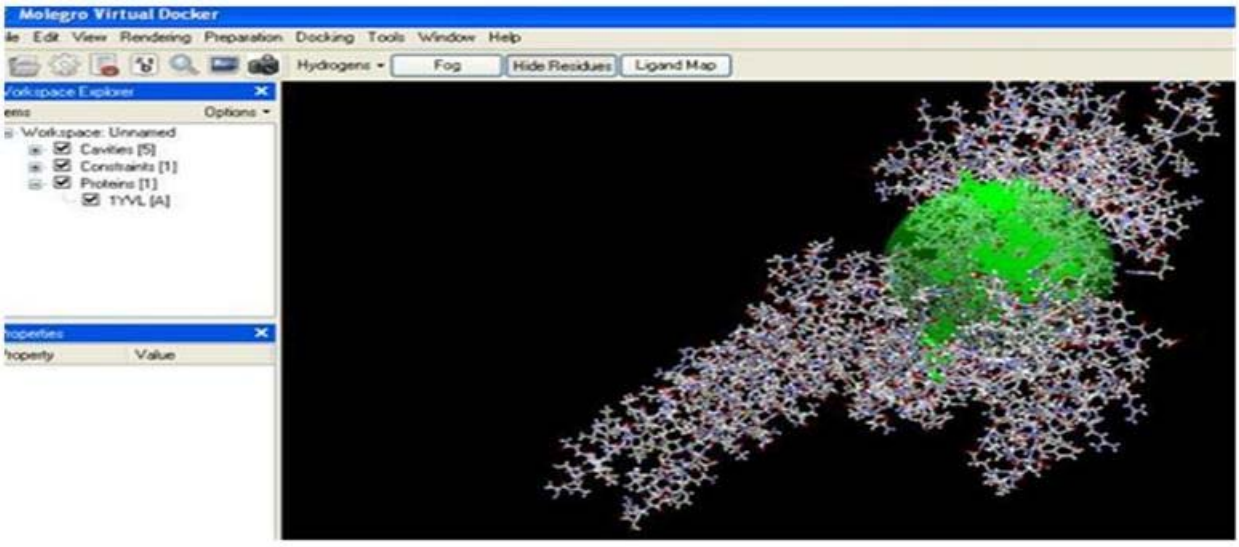

Figure 2: The Molegro vitual docker window showing upload of RSV PDB (1G2C) viral protein with constrain shows it target active place.

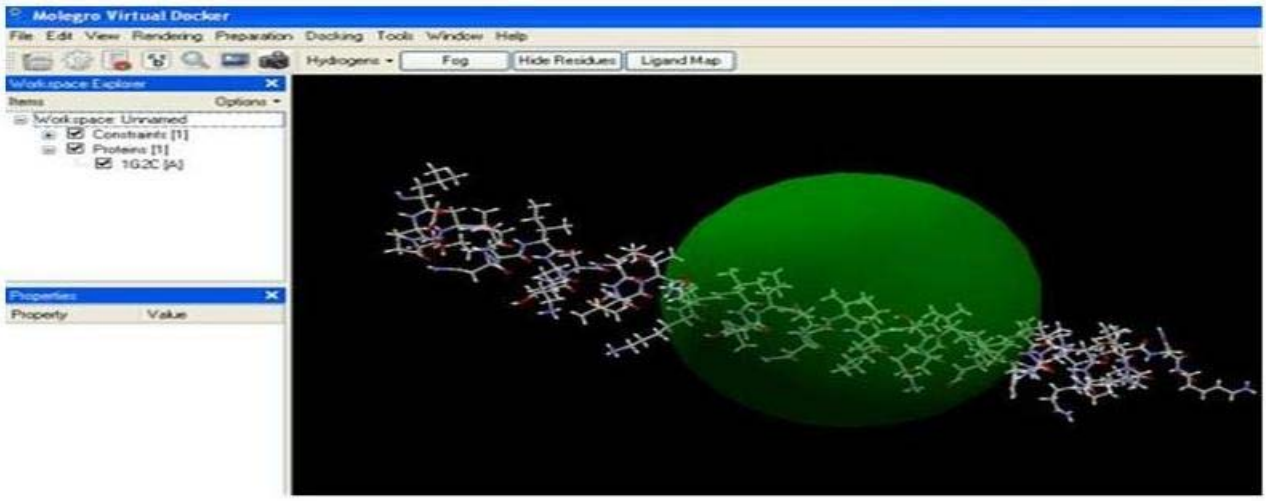

Figure 3: The Molegro vitual docker window showing upload of signaling protein STAT PDB (1YVL) with constrain shows its target active site region

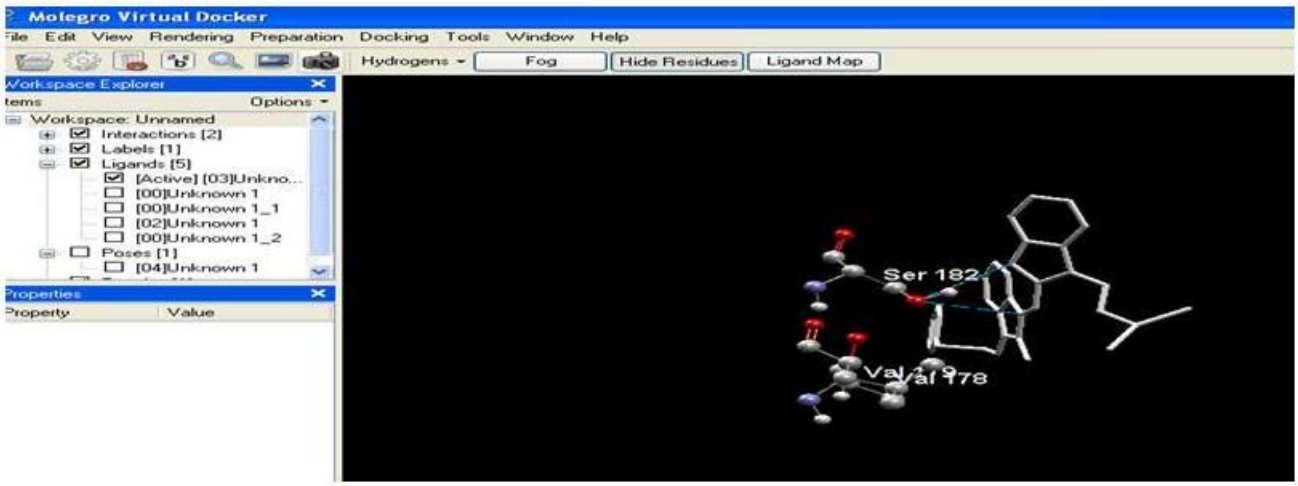

Figure 4: Molegro virtual docker window showing the most active two interactions of compound (11ai) with Ser -182 amino acid of viral protein (PDB 1G2C)

A) Ligand Molecule (11ai)	$C_{27}H_{33}N_5O_3$
Molecular weight	475.582
Mole dock score	-105.647
Rerank score	-68.236
LE1	-2.89678
LE3	-1.94963
HBond	-6.51476

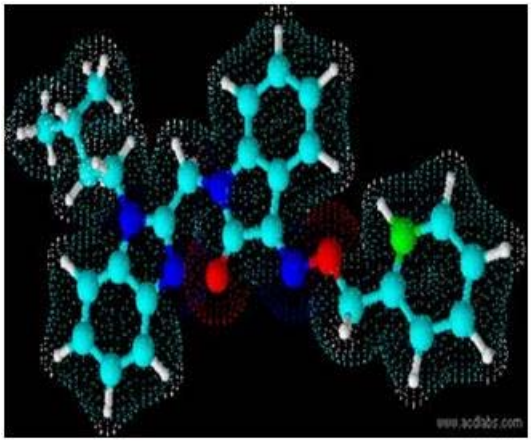

Figure 5: The Most active Compound (11ai) screened from Isatin- Oxime ether derviateis from which showing the best mole dock values with the protein 1G2C

Ligand 11i	$C_{22}H_{20}F_3N_5O_2$
MW	455.43324
Mol dock Scor	-184
Rerank score	-146.348
LE1	-4.25668
LE3	-1.13044
HBond	-3.6881

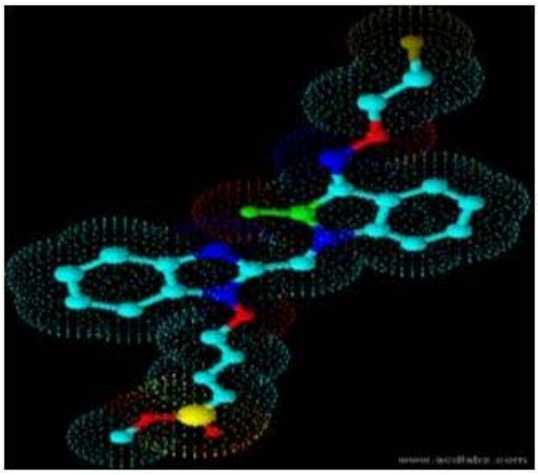

Figure 6: The Most active Compound (11i) screened from Isatin Oxime ethers showing best mole dock values with the protein 1YVL from 76 molecules.

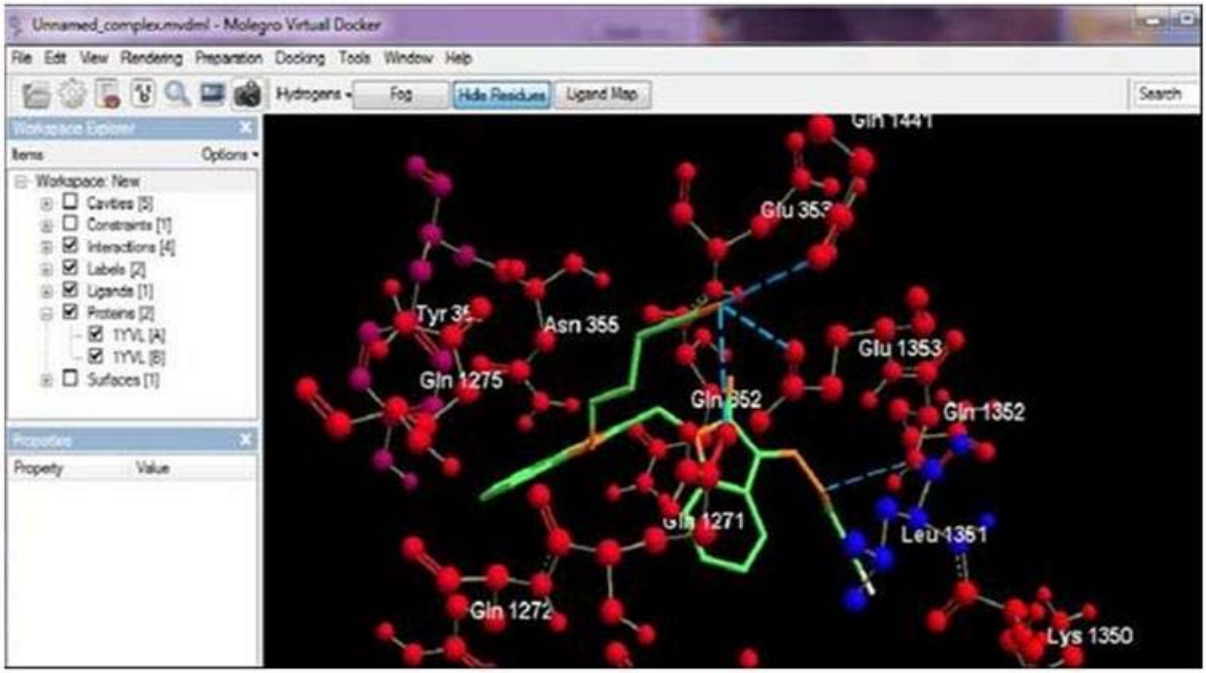

Figure 7: Illustration of Molegro virtual docker window showing the highest mol dock score with STAT protein .and the active site interactions with four amino acids Glu-353, Gln-352, Glu-1352 and leu-1352 of PDB 1YVL and 11i ligand showing the description of best drug activity on cancer protein when compare to respiratory diseases.

Validation of Mol docking

The ligand orientation and the position of interaction is obtained from the docking studies were likely to represent valid and reasonable binding modes of the inhibitors, Molegro virual docking results parameter are first validated on the crystal structure RSV fusion protein (PDB 1G2C) and then simultaneously with STAT protein (PDB 1YVL) with Isatin - oxime derivatives. The ligand present in the conformation found in the crystal structure selected and dock to the corresponding binding pocket to determine the ability to reproduce the orientation and position of the interaction of inhibitor observed in the crystal structure. The 1G2C PDB of RSV Isatin – oxime fusion inhibitor resulting highest mole dock score is -105.0967 (figure 5). Whereas the PDB 1YVL protein giving the high acceptable docking score is -184.697 (figure 6) at four amino acids Glu-353, Gln-352, Glu-1352 and leu-1352 (figure 7) and the zone of interaction between ligand –protein. This resulting the Isatin –Oxime derivaties ligand giving best drug able To the RSV.

CONCLUSION

In this work, we have utilized the structure and ligand-based approached of docking and 3D - QSAR to explore the intermolecular interactions of isatin-oxime of RSV fusion inhibitors on STAT protein. The docking study not only confirmed the essentials of the binding mode of the X-ray crystal structure but also provided the information on how lead optimization improved the activities. 3D-QSAR, CoMFA and CoMSIA models for Isatin-Oxime RSV fusion inhibitors show insight into the influence of various structural attributes for the biological activity. The molegro virtual docker software gives the best mol dock score and based on the score we can explore active ligand further to generate more effective and potential drug molecules through ligand based drug designing approaches and may be considered as a power-full tool in designing and forecasting more efficacious analogues, since they point towards the molecular sites that may be explored in order to maximize the bio profile and prevent toxicity protein.

REFERENCES

1. Chanock R, Roizman B, Myer R, Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). I. Isolation, properties and characterization, Am J Hyg, 66, 1957, 281–290.
2. Cianci C, Genovesi E.V, Lamb L, Medina I, Yang Z, Zadjura L, Yang H. D'Arienzo C, Sin N, Yu K.L, Combrink K, Li Z, Colonno R, Mean well N, Clark J, Krystal M, Oral efficacy of a respiratory syncytial virus inhibitor in rodent models of infection, Antimicrob Agents Chemother, 48, 2004, 2448-2454.
3. SigursN, GustafssonP.M., BjarnasonR, LundbergF, SchmidtS, Sigurbergsson F, Kjellman B, Severe respiratory syncytial virus bronchiolitis in infancy and asthma and allergy at age 13, Am J Respir Crit Care Med, 171, 2005, 137-141.
4. Jacqueline Bromberg *Stat proteins and Oncogenesis,* J Clin Invest, 109, 2002, 1139-1142.
5. Horvath C.M, STAT proteins and transcriptional responses to extracellular signal, TIBS, 2000, 25.
6. Tammy Bowman, Roy Garcia, James Turk son, Richard Jove' STATs in oncogenesis, Oncogene, 19, 2000, 2474-2488.
7. Ny Sin. Brian L, Venables. Keith D, Combrink H, Belgin Gulgeze, Kuo-Long Yu, Rita L, Civiello, Jan Thuring X, Alan Wang, Zheng Yang, Lisa Zadjura, Anthony Marino, Kathleen, F, Kadow, Christopher W, Cianci, Junius Clarke, Eugene V, Genovesi, Ivette Medina, Lucinda Lamb, Mark Krystal, Nicholas A. Mean well. Respiratory syncytial virus fusion inhibitors. Part 7: Structure–activity relationships associated with a series of isatin oximes that demonstrate antiviral activity *in vivo* Bioorganic & Medicinal Chemistry Letters, 19, 2009, 4857-4862.
8. Thomsen R, Christensen MH, MolDock A new technique for high- accuracy molecular docking, Med.chem, 49, 2006, 3315-3321.

Source of Support: Nil, **Conflict of Interest:** None.