ISSN 0975-6299 Vol.1/Issue-3/Jul-Sep.2010

International Journal of Pharma and Bio Sciences

STRUCTURE MODELING OF VRK1 PROTEIN AND ITS MOLECULAR DOCKING STUDY WITH RIBAVIRIN ANALOGS

SHANKARACHARYA*[1], BHAVANA SRIVASTAVA[2] AND AMBARISH S. VIDYARTHI [1]

[1] Department of Biotechnology, Birla Institute of Technology, MESRA, Ranchi – 835215
[2] Department of Computer science, Yagyavalkya Institute of Technology (YIT), Jaipur

**** Corresponding Author*** shankaracharya@bitmesra.ac.in, shankaracharya2222@gmail.com

ABSTRACT

In order to design better drug for the treatment of Hepatitis C Virus, a 3-D model of the Vaccine Related Kinase 1 protein was generated based on the crystal structure of the Vaccine Related Kinase 2 (PDB ID 2V62) having 57% sequence identity. The two sequences were also aligned with a third protein VRK 3 (2JII) to see the similarity between the three. Its 3-D structure was evaluated and validated using PROCHECK comprising 97.5% amino acid residues in favored and additional allowed region of Ramachandran plot. The model was also validated with What_check and Errat, which confirms that the model was of good quality. With this model, a docking study was performed with GLIDE using 93 analogs of Ribavirin and the results indicated that the ASN377, PRO190, SER136, LYS392, THR390 and SER376 form hydrogen bonds and have strong nonbonding interaction with the analog CID_196553. This finding may help to understand the nature of VRK1 and development of specific anti-HCV therapies. Our results may be helpful for further experimental investigations.

KEYWORDS

VRK1, molecular modeling, drug design, docking, Ribavirin

INTRODUCTION

Hepatitis C virus (HCV) has emerged as a major cause of human liver disease, with 3% of the world population persistently infected with the virus and more than 1 million new cases of infection reported annually [1]. In about 70% of the cases, HCV escapes the immune system and establishes a chronic infection. In the long term, these chronic carriers are at risk of developing life-threatening liver disease, including hepato-cellular carcinoma [1]. HCV is an enveloped virus that belongs to the *Hepacivirus* genus in the *Flaviviridae* family. Its genome consists of RNA of approximately 9.6 kb in length that contains a large open reading frame. Translated polyprotein is processed by cellular and viral proteases into at least 10 individual structural and nonstructural (NS) proteins.

Current therapy for hepatitis C involves treatment with a combination of ribavirin and interferon-α.

However, this is effective only in half of patients, but unsuitable for certain patient populations [2]. Thus, there is an intense effort to develop new and better treatments, mostly by targeting viral enzymes [3].

The sequencing of the human genome and availability of highly specific methods for gene inactivation now allow a systematic examination of the roles of individual human genes in HCV replication and the identification of new potential drug targets. Study showed a substantial reduction in cellular viral RNA and protein expression was observed when three kinases, Csk, Jak1 and Vrk1, were targeted with siRNAs [4].

Human vaccinia-related kinase1 (VRK1) is a novel serine-threonine kinase that regulates several transcription factors, nuclear envelope assembly, and chromatin condensation and is also required for cell cycle progression. The regulation of this kinase family is unknown. VRK is composed of three proteins, two of which are catalytically active, VRK1 and VRK2, and are mainly expressed in proliferating cells [5]. These proteins have a conserved kinase domain but differ in their regulatory region with little conservation among them or with any other protein. The VRK1 protein is mostly nuclear, although in some cell types it is also present in the cytosol; this subcellular localization is regulated in response to a specific signal.

Vaccinia-related kinase1 (VRK1) is the mammalian homolog of Nucleosomal histone kinase 1 (NHK1). Sequence similarities between NHK1 and VRK1 are evident in the kinase domain (approximately 40% identity), and the carboxyl termini contain a characteristic basic-acidic-basic motif. VRK1, identified from the screening of novel genes involved in cell cycle regulation from fetal liver is designated on the basis of 40% sequence identity with vaccinia virus B1 kinase, which plays a critical role in viral DNA replication [4].

Ribavirin is a small molecule type of drug. Ribavirin is an anti-viral drug active against a number of DNA and RNA viruses. Its Chemical IUPAC Name is 1-[(2R, 3R, 4S, 5R)-3,4-dihydroxy-5-(hydroxymethyl)oxolan-2-yl]-1,2,4-triazole-3-carboxamide and has the Chemical formula $C_8H_{12}N_4O_5$. It is a member of the nucleoside antimetabolite drugs that interfere with duplication of the viral genetic material. The drug inhibits the activity of the enzyme RNA dependent RNA polymerase, due to it's resemblence to building blocks of the RNA molecules. The oral form is used in the treatment of hepatitis C, in combination with interferon drugs. The primary serious adverse effect of ribavirin is hemolytic anemia, which may worsen preexisting cardiac disease. Ribavirin is also incorporated into the viral genome causing lethal mutagenesis and a subsequent decrease in specific viral infectivity [6]. In the present study we have modeled the structure of VRK 1 protein using modeler 9v7. Further after the validation of the structural model, molecular docking was performed with 93 analogs of Ribavirin. The aim was to find out the best interacting analog of ribavirin which may help in the designing better drug for HCV.

MATERIALS AND METHODS

Search and retrieval of target protein sequence

Information about protein sequence of VRK1 was retrieved from NCBI (http://www.ncbi.nlm.nih.gov/protein/BAA19108.1).

Selection of template: Template was selected by homology search of query protein (VRK1) sequence against the databases available on PDB (http://ww.rcsb.org). Homologous structure of sequence having the lowest E-value, 50% and above identity, lower resolution was selected as template.

Homology modeling: Homology modeling was done using Modeler 9v7 [7, 8]. This requires one sequence of known 3D structure with significant similarity with the target sequence and Python 2.5 script files containing Modeler commands.

The co-ordinate file of template from PDB was used as such.

Evaluation and Refinement of predicted models: All predicted 5 models were evaluated by Procheck [9], What_check [10] and Errat [11]. Ramachandran plot statistics was used to evaluate the stability of the model. Gnuplot was finally used to plot the profiles generated by Modeler (http://www.gnuplot.info) for substantial validation of the structure. Refinement of the structure was done by molprobity server[12].

Virtual screening of flavopiridol analogues through molecular docking: Ribavirin and its analogs were taken from NCBI Pub-chem in SDF format and converts to 3-D structure using weblab viewer lite program. The 3-D structure of VRK1 and flavopiridol 3-D analogs was used for molecular docking using GLIDE[13].

Protein structure accession number: The refined homology model of 3D structure of VRK1 of human was submitted to PMDB (http://mi.caspur.it/PMDB/)[14] and the same was assigned the identifier **PM0076358**.

RESULTS AND DISCUSSIONS

Search for template on Protein Data Bank through BLAST generated 51 homologous structures. Most probable homologous proteins are listed in Table 1. Amongst them 2V62 was selected on the basis of lowest resolution and E-Value with highest identity (Table 1). The homology of VRK1 sequence showed 57% identity with crystal structure of VRK3 (PDB ID-2V62).The protein sequences of VRK1, VRK2 (PDB ID-2V62) and VRK3 (PDB ID-2JII) were aligned and shown in figure 1. The asterisk showed the identity of amino acids present in the three protein sequences.

Table 1.

RCSB Blast result

Sl. No.	Templates	Length	Resolution (Å)	E-Value	% Identity
1.	**2V62**	**303**	**1.70**	**1.459**	**57%**
2.	2JII	316	2.50	1.169	36%
3.	1CKI	317	2.30	6.57	29%
4.	2C47	294	2.40	1.850	29%
5.	2CHL	299	1.95	1.348	28%

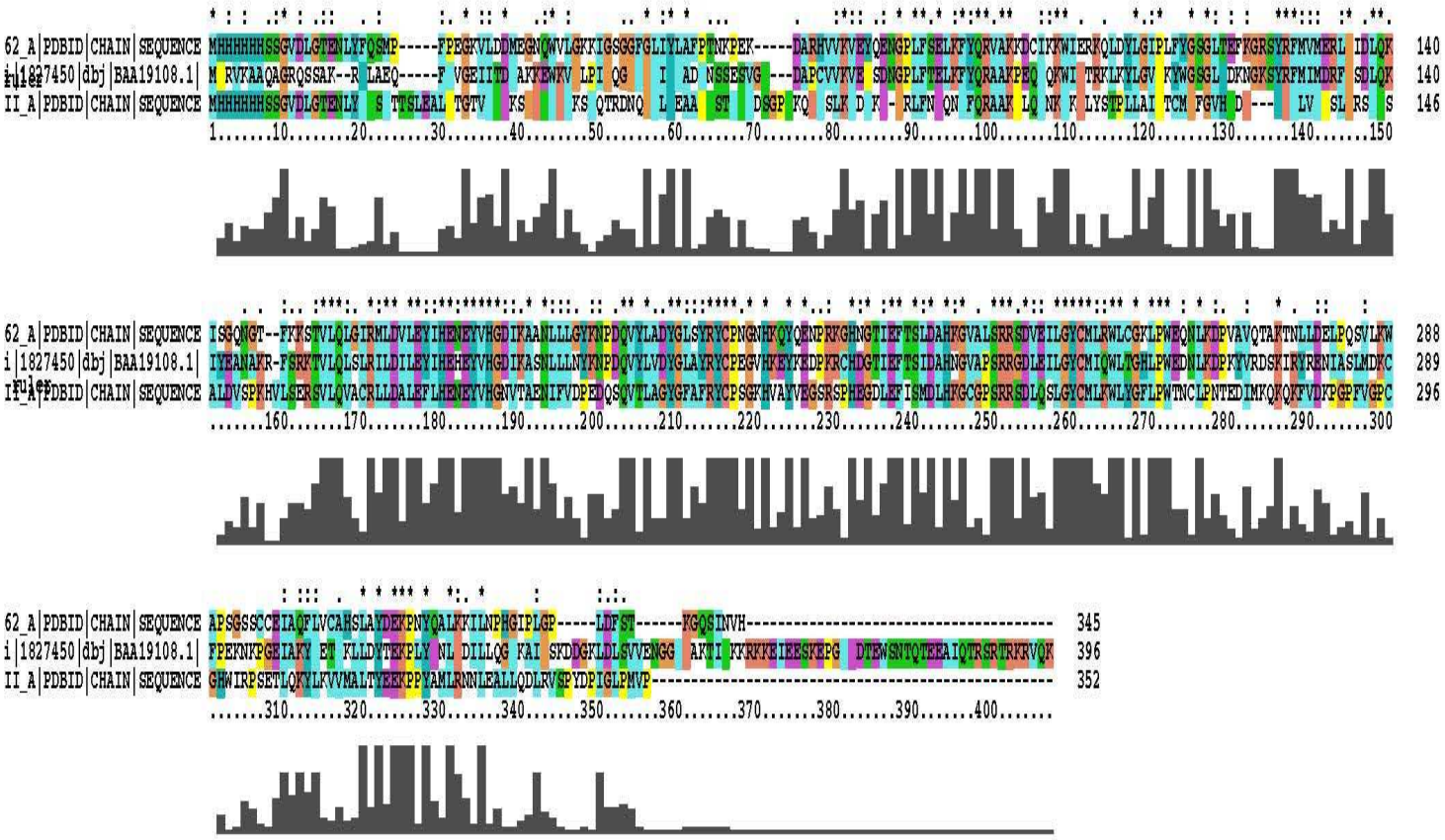

Figure 1
MSA Result of VRK1, VRK2 and VRK3

VRK1 sequence was used to generate the 3-Dstructure using known crystal structure 2V62. Homology modeling was carried out through Modeller 9v7 using Basic Modelling. Total 5 models were generated. Dope scores of the generated models were calculated using the command model-single.py. The model having minimum dope score was considered as the best model of protein VRK1 (Table 2). The selected best model vrk.B99990002.pdb is depicted in figure 2.

Table 2
Comparison of Dope score, H- bonds, strands and turns of the models

Sl. No.	Model	Dope score	H-bonds	Strands	Turns
1.	vrk.B99990001.pdb	-35024.30469	221	17	37
2.	**vrk.B99990002.pdb**	**-35662.75391**	**224**	**19**	**39**
3.	vrk.B99990003.pdb	-35568.48828	213	17	38
4.	vrk.B99990004.pdb	-35660.68359	217	17	35
5.	vrk.B99990005.pdb	-35560.71875	215	17	44

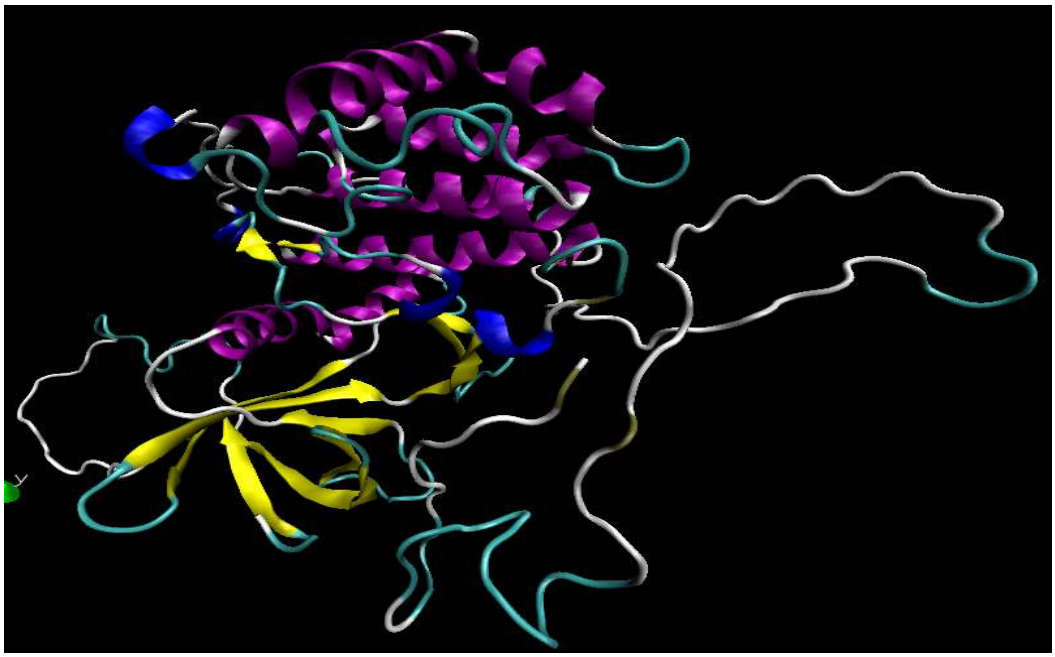

Figure 2

3 D structure of model protein vrk.B99990002.pdb

Further evaluation and validation was performed using a series of programs like Gnuplot, Procheck[9] and Errat[11]. Free energy of 3-D structures of target and template was evaluated. The free energy of VRK1 was almost similar with the template 2V62. Maximum number of H-bonds, strands and turns shows the compactness of the best model (vrk.B99990002.pdb) in comparison to other models generated (Table 2).

Further validation program, Procheck[9] was used to perform full geometric analysis as well as stereochemical quality of a protein structure by analyzing residue-by-residue geometry and overall structure geometry. After running Procheck, Ramachandran plot (Figure 4) shoes that for the model vrk.B99990002**,** 85.1% residues were in favored region, 12.4% in the additional allowed region, 2.3% in the generously allowed region and 0.3% of the residues in the disallowed region, which made this model more acceptable as compared to other predicted models.

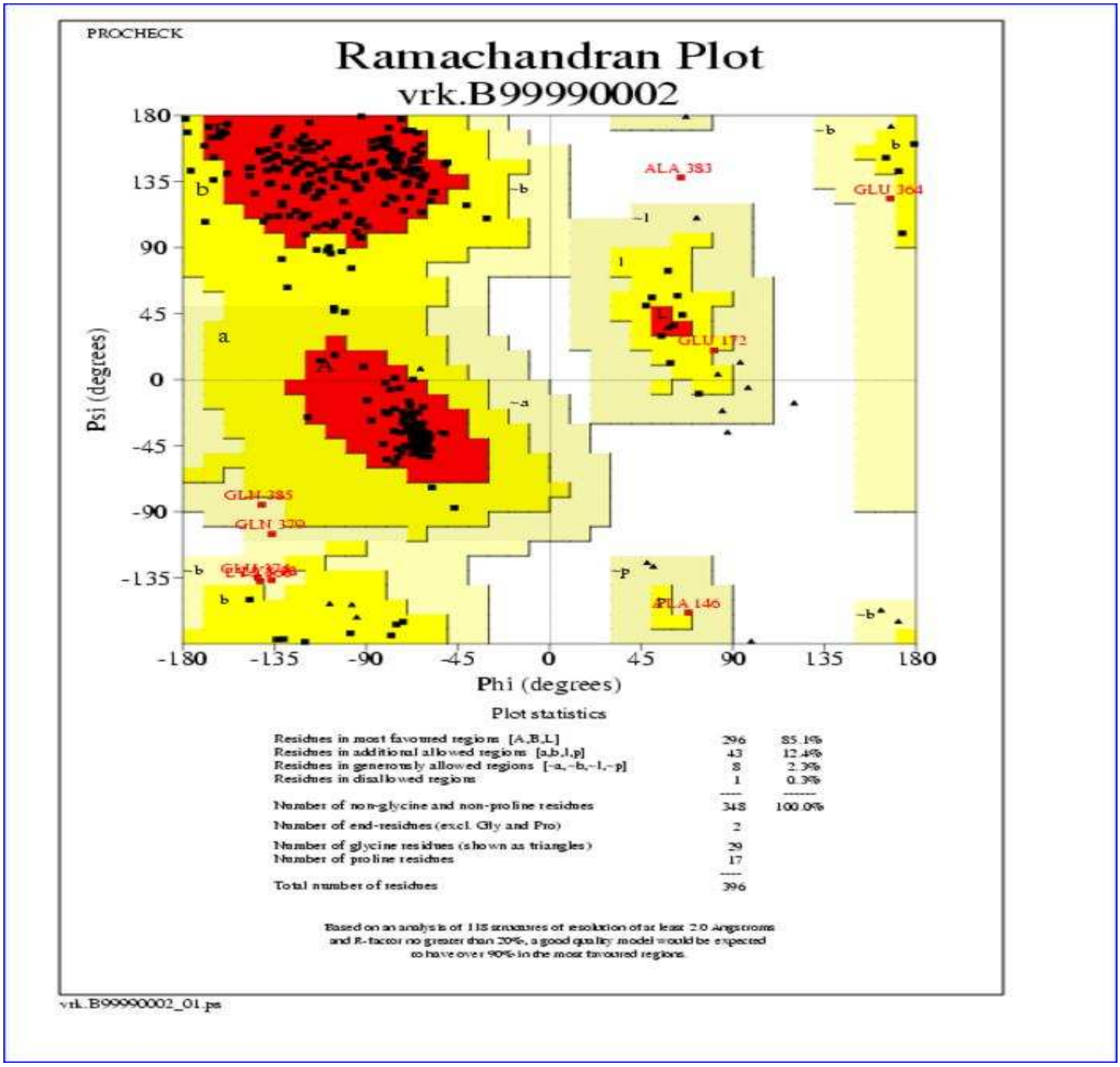

Figure 3.

Ramchandran Plot for vrk.B99990002.pdb

Errat program was used to analyze the statistics of non-bonded interactions between different atom types and provide the overall quality factor of the predicted models. For model vrk.B99990002.pdb the overall quality factor was found to be 61.813% (Table 3), which indicates that the model is acceptable.

Table 3

Ramchandran Statistics of the predicted models

SL. No.	Predicted Models	No. of residues in (Regions)				Overall Quality Factor
		Most favored region	Addional allowed region	Generously allowed region	Disallowed region	
1.	vrk.B99990001.pdb	87.1%	10.9%	1.4%	0.6%	62.865
2.	**vrk.B99990002.pdb**	**85.1%**	**12.4%**	**2.3%**	**0.3%**	**61.813**
3.	vrk.B99990003.pdb	87.9%	9.5%	2.0%	0.6%	62.908
4.	vrk.B99990004.pdb	84.8%	11.8%	2.3%	1.1%	64.972
5.	vrk.B99990005.pdb	87.9%	10.1%	1.1%	0.9%	52.101

Comparison of the Ramachandran Plot statistics of the modeled protein vrk.B99990002.pdb and the template (Table 4) showed that our selected model's residues geometry lies near to the template.

Table 4

Comparison between vrk1 and 2v62

Sl. No.	Properties	Vrk1 (Model)	2v62 (Template)
1	Most favored Region	85.1%	88.4%
2	Allowed Region	12.4%	11.2%
3	Generously Allowed Region	2.3%	0.4%
4	Disallowed Region	0.3%	0.0%

Z-value or Z-score is the normality of a score which is the number of standard deviations that the score deviates from the expected value. A property of Z-values is the root-mean-square of a group of Z-values (the RMS Z-value) which is expected to be 1.0. Z-values above 4.0 and below -4.0 are quite uncommon. In our result the RMS Z-score for all improper dihedrals in the structure is within normal ranges and the distribution of residue types over the inside and the outside of the protein is normal. The different Z-scores of our best model vrk.B99990002.pdb are presented in Table 5.

Table 5

Z-scores as obtained from WHAT_CHECK program

Sl. No.	Residues Property	Z-scores
1.	Bond lengths	0.955
2.	Bond angles	1.323
3.	Inside/Outside distribution	1.139
4.	Omega angle	0.818
5.	Side chain planarity	0.401
6.	Improper dihedral distribution	0.913

In one study, the possible targets for drug development against Hepatitis C Virus; the 3-D NS3 protease has been generated using modelling [15]. Based on this study Singh et al described the 3-D structure of NS3 helicase and was used to search theoretical binding affinity with ribavirin and its analogs. The molecular docking was used for searching of better affinity of drugs with NS3 helicase. Five drugs were selected for docking with NS3 helicase. Out of five, three drugs viz. Levovirin, Ribamidine and Ribavirin were showed strong binding affinity with helicase. The docking energy of Levovirin, Ribamidine and Ribavirin and NS3 complexes were −13.61, −12.5 and −14.25 kcal/mol respectively.[16]

In order to make more comprehensive study of Ribavirin as more potent drug for HCV docking was performed with its 93 analogues and target VRK1. For this the sitemap was generated (Data not shown). In the docking process, energy all ligands in the dataset were minimized and docked with the target. The conformation with the lowest Glide score was chosen as the optimally docked ligand. For Ribavirin, ligand **CID_196553** had the minimum glide score of -7.14 with cumulative van der waals force as -67.3 (Table 6).

Table 6

Docking result of analogs of rebavirin with the modeled protein vrk.B99990002.pdb

Rank	Analogs (CID)	GScore	Rewards	vdW	Coul	Emodel	CvdW	Lipo
1	196553	-7.14	-0.4	-42.2	-25.1	-95.2	-67.3	-0.3
2	25245443	-7.10	-2.6	-24.6	-19.1	-88.4	-43.7	0.0
3	124970	-6.65	-0.9	-27.8	-26.1	-74.5	-53.9	-0.3
4	129235	-6.65	-1.3	-24.1	-20.6	-61.2	-44.7	-0.4
5	122108	-6.67	-0.3	-36.1	-25.4	-86.5	-61.5	-0.6
6	10220469	-6.54	-2.0	-14.8	-20.5	-51.1	-35.2	-0.4
7	100252	-6.43	-1.4	-25.0	-23.1	-65.9	-48.1	-0.2
8	9794820	-6.37	-1.9	-15.7	-19.4	-48.9	-35.2	-0.2
9	196691	-6.33	-1.6	-34.2	-21.1	-72.8	-48.1	-0.1
10	10111813	-6.29	-0.7	-27.2	-26.1	-74.2	-35.0	-0.1
11	5064	-6.24	-2.0	-18.3	-20.1	-53.2	-55.3	-0.3
12	503875	-6.23	-0.3	-41.0	-24.2	-86.6	-38.4	-0.2
13	503876	-6.15	-0.0	-44.9	-23.8	-91.5	-65.3	-0.1
14	44341303	-6.08	-1.9	-14.3	-21.9	-50.0	-68.7	-0.2
15	6604395	-6.04	-2.0	-18.8	-19.3	-51.6	-36.2	-0.2
16	451448	-6.01	-1.8	-10.5	-21.6	-44.5	-38.1	0.0
17	6713992	-6.01	-1.8	-19.4	-19.4	-51.8	-32.1	-0.4
18	3348128	-5.92	-0.1	-37.9	-37.9	-85.4	-38.8	-0.3
19	37542	-5.85	-2.0	-18.4	-19.1	-51.1	-62.8	-0.1
20	490511	-5.85	-2.0	-18.4	-19.1	-51.1	-37.5	-0.1

The Active site residues in the target protein VRK1 of human includes ASN377, SER136, PRO190, LYS392, THR390, TYR 187and ARG 331. Six hydrogen bonds were formed between the Ribavirin analog **CID_196553** and VRK1 residues of ASN 377, LEU 185, SER 136, LYS 392, THR 390, SER 376 having the length of 2.276 Å, 2.036 Å, 3.112 Å, 1.185 Å, 1.965 Å and 2.332 Å, respectively. Presence of large amount of hydrogen bonds between the protein and analog shows the strong binding affinity between them (Figure 4).

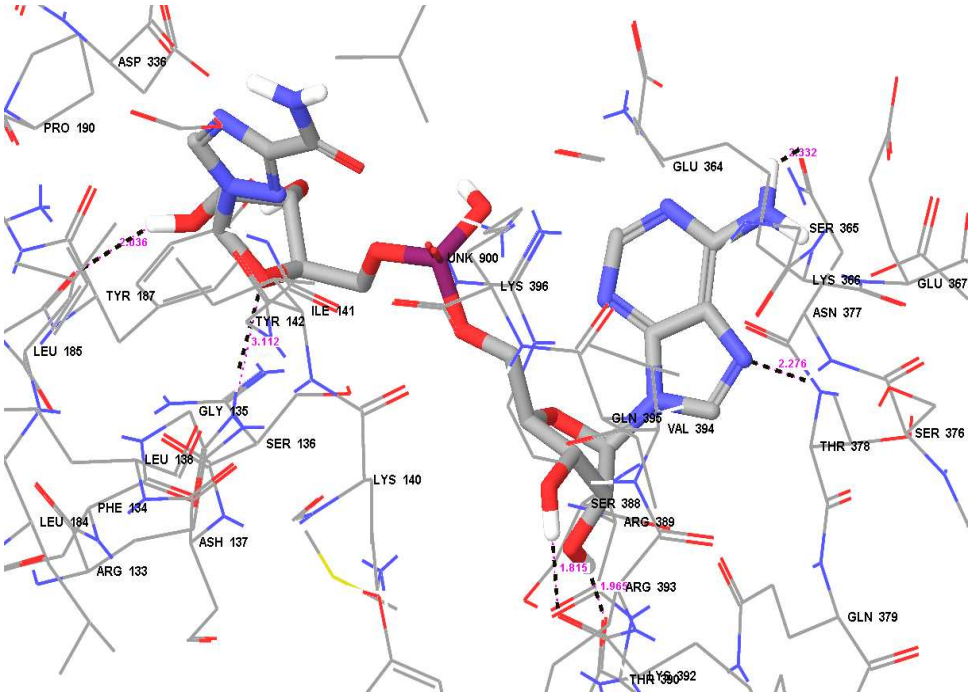

Figure 6

Docked result showing the interaction between active site residues of modeled protein and Ribavirin analog CID_196553

Hence the selected analog CID_196553 can be used as a replacement of the existing drug Ribavirin and may be subject to clinical trail.

CONCLUSION

The present study was undertaken to construct the 3-D structure of VRK1 protein of human. The virtual screening of drug was performed to search suitable analog among 93 analogs of drug Ribavirin. CID_196553 was found more potent one comprising highest binding affinity with modeled VRK1 protein. The present investigation may provide new insight to control superfluous use of drugs *in vitro*. The 3-D structure of VRK1 of human and Ribavirin analog CID_196553 would needed to be verified experimentally. Further various variants of this analog can be designed and docked to find out more potent drug than the suggested Ribavirin (CID_196553).

ACKNOWLEDGEMENT

The authors acknowledge BTISnet, Department of Biotechnology, Government of India, New Delhi for providing Infrastructure facility for Bioinformatics. Authors also acknowledge Departments of Biotechnology and Pharmaceutical Sciences, B. I. T. MESRA, Ranchi for providing Docking software.

REFERENCES

1. Wasley A and Alter MJ, Epidemiology of hepatitis C: geographic differences and temporal trends, Semin. Liver Dis. 20:1–16, (2000).
2. Fried MW, Shiffman ML, Reddy K R, Smith C, Marinos G, Goncales FL, Jr, Haussinger D, Diago M, Carosi G, Dhumeaux D, Craxi A, Lin A, Hoffman J, and Yu J, Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection, N. Engl. J. Med., 347:975–982, (2002).
3. Lin K, Perni RB, Kwong AD and Lin C. VX-950, a Novel Hepatitis C Virus (HCV) NS3-4A Protease Inhibitor, Exhibits Potent Antiviral Activities in HCV Replicon Cells, Antimicrob. Agents Chemother., 50: 1813-1822, (2006).
4. Supekova L, Supek F, Lee J, Chen S, Gray N, Pezacki JP, Schlapbach A and Schultz PG. Identification of Human Kinases Involved in Hepatitis C Virus Replication by Small Interference RNA Library Screening, J. Biol. Chem., 283: 29-36, (2008).
5. Nezu J, Oku A, Jones MH, Shimane M. Identification of two novel human putative serine/threonine kinases, VRK1 and VRK2, with structural similarity to vaccinia virus B1R kinase, Genomics, 45 (2): 327–31, (1998).
6. http://www.drugbank.ca/drugs/DB00811.
7. Fiser A and Sali A. Modeller: generation and refinement of homology-based protein structure models. Methods Enzymol., 374: 461-469, (2003).
8. Sali A and Blundell TL Comparative protein modelling by satisfaction of spatial restraints, J Mol Biol., 234: 779-815, (1993).
9. Laskoswki RA, MacArthur MW, Moss DS and Thorton JM. PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Cryst., 26:283-291, (1993).
10. G. Vriend. Parameter relation rows: a query system for protein structure function relationships, Protein Eng., 4: 221-223, (1990).
11. Colovos C and Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions, Protein Sci. 2:1511-1519, (1993).
12. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall III WB, Snoeyink J, Richardson JS and Richardson DC. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Research, 35:W375-W383, (2007).
13. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS. Glide: a new approach for rapid, accurate docking and scoring. Method and assessment of docking accuracy, J. Med. Chem., 47 (7): 1739–49, (2004).
14. Castrignano T, De Meo PD, Cozzetto D, Talamo IG, and Tramontano A. The PMDB Protein Model Database, Nucleic Acids Res., 34(1): D306 - D309, (2006).
15. Brinkworth RI, Fairlie DP, Leung D, Young PR, Homology model of the Dengue 2 virus NS3 protease: putative interactions with both substrate and NS2 cofactor, J Gen Virol., 80:1167–1177, (1999).
16. Singh V and Somvanshi P. Structural Modeling of the NS 3 Helicase of Tick-borne Encephalitis Virus and Their Virtual Screening of Potent Drugs Using Molecular Docking Interdiscip Sci Comput Life Sci., 1: 168–172, (2009).