

# Exploring potential inhibitors against Kyasanur forest disease by utilizing molecular dynamics simulations and ensemble docking

---

Kandagalla, Shivananda; Novak, Jurica; Shekarappa, Sharath Belenahalli; Grishina, Maria A; Potemkin, Vladimir A; Kumbar, Bhimanagoud

Source / Izvornik: **Journal of biomolecular structure & dynamics, 2021**

Journal article, Accepted version

Rad u časopisu, Završna verzija rukopisa prihvaćena za objavljivanje (postprint)

<https://doi.org/10.1080/07391102.2021.1990131>

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:193:878070>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-01-08**

Repository / Repozitorij:



[Repository of the University of Rijeka, Faculty of Biotechnology and Drug Development - BIOTECHRI Repository](#)



# Exploring potential inhibitors against Kyasanur forest disease by utilizing molecular dynamics simulations and ensemble docking

Shivananda Kandagalla<sup>1\*</sup>, Jurica Novak<sup>1\*</sup>, Sharath Belenahalli Shekarappa<sup>2</sup>, Maria A. Grishina<sup>1</sup>, Vladimir A. Potemkin<sup>1</sup>,  
Bhimanagoud Kumbar<sup>2,3\*</sup>

<sup>1</sup>Laboratory of Computational Modeling of Drugs, Higher Medical and Biological School, South Ural State University,  
Chelyabinsk, Russia

<sup>2</sup>Department of PG Studies and Research in Biotechnology and Bioinformatics, Kuvempu University, Jnana Sahyadri,  
Shankaraghatta, Shivamogga, 577451, Karnataka, India

<sup>3</sup>ICAR- National Institute of Veterinary Epidemiology and Disease Informatics, Yelahanka, Bengaluru, 560064, Karnataka,  
India

\*Corresponding authors

E-mail: kandagallas@gmail.com; kandagallas@susu.ru (S.K); novaki@susu.ru (J.N); kumbar.bhimanagoud@gmail.com  
(B.K.)

ORCID numbers:

Shivananda Kandagalla: 0000-0003-4540-3500

Jurica Novak: 0000-0001-5744-6677

Sharath Belenahalli Shekarappa: 0000-0003-2695-6484

Maria A. Grishina: 0000-0002-2573-4831

Vladimir A. Potemkin: 0000-0002-5244-8718

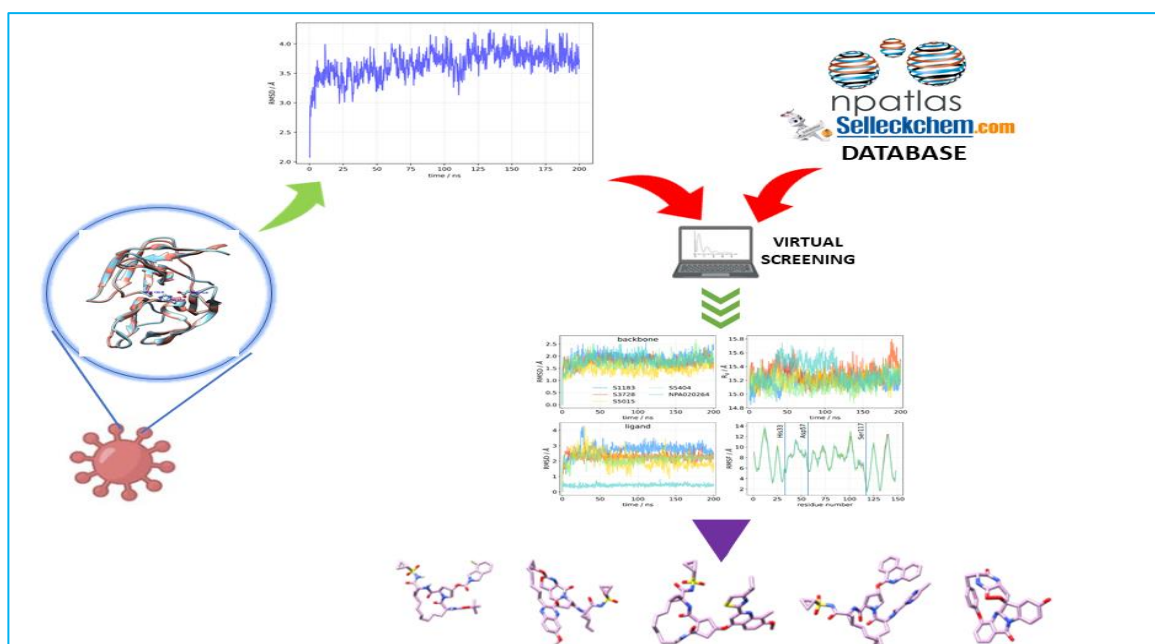
Bhimanagoud Kumbar: 0000-0002-9980-3246

## Abstract:

Kyasanur forest disease (KFD) is a tick-borne, neglected tropical disease, caused by KFD virus (KFDV) which belongs to *Flavivirus* (Flaviviridae family). This emerging viral disease is a major threat to humans. Currently, vaccination is the only controlling method against the KFDV, and its effectiveness is very low. An effective control strategy is required to combat this emerging tropical disease using the existing resources. In this regard, *in silico* drug repurposing method offers an effective strategy to find suitable antiviral drugs against KFDV proteins. Drug repurposing is an effective strategy to identify new use for approved or investigational drugs that are outside the scope of their initial usage and the repurposed drugs have lower risk and higher safety compared to *de novo* developed drugs, because their toxicity and safety issues are profoundly investigated during the preclinical trials in human/other models. In the present work, we evaluated the effectiveness of the FDA approved and natural compounds against KFDV proteins using *in silico* molecular docking and molecular simulations. At present, no experimentally solved 3D structures for the KFD viral proteins are available in Protein Data Bank and hence their homology model was developed and used for the

analysis. The present analysis successfully developed the reliable homology model of NS3 of KFDV, in terms of geometry and energy contour. Further, *in silico* molecular docking and molecular dynamics simulations successfully presented four FDA approved drugs and one natural compound against the NS3 homology model of KFDV.

**Keywords:** Kyasanur forest disease, NS3, molecular dynamics, Aspertryptanthrin C, Grazoprevir



## 1. Introduction

Kyasanur Forest Disease (KFD) is an emerging zoonotic disease caused by the Kyasanur Forest Disease Virus (KFDV). The disease is associated with sudden onset of high-grade fever, prostration, nausea, vomiting, diarrhea, and occasionally neurological and hemorrhagic manifestations (Sreenivasan et al., 1986; Upadhyaya, Murthy, & Anderson, 1975; WORK et al., 1959). The virus belongs to the genus *Flavivirus* of the *Flaviviridae* family. KFDV was first identified and isolated in Kyasanur forest area, Shivamogga district of Karnataka, India, in the year 1957. The disease name was derived from the forest range where the virus was first isolated. Initially, the disease was endemic in few districts of Karnataka along Western Ghats of India, but subsequently it has spread to other states, like Maharashtra, Goa, and Tamil Nadu. The districts Shivamogga, Chikamagaluru, Uttara Kannada, Dakshin Kannada, and Udupi from Karnataka have reported outbreaks due to the KFD virus (Eldridge et al., 2004; WORK et al., 1957). Districts bordering Karnataka state like Nilgiris from Tamil Nadu, Wayanad and Malappuram from Kerala have reported monkey deaths and human cases in 2013-2014 (Mourya et al., 2013). Also, bordering states North Goa district in Goa (in 2015) and Sindhadurga district in Maharashtra (in 2016) have reported outbreaks (Sreenivasan et al., 1986).

Many forests dwelling small mammals like rodents, shrews, insectivorous bats, and many birds maintain the natural enzootic cycle of the virus in the forest ecosystem (Eldridge et al., 2004). The wild primates, black-faced langurs (*Semnopithecus entellus*), and red-faced bonnet monkeys (*Macaca radiata*) get the virus infection by tick bite (Pattnaik, 2006). Ticks maintain the transovarial and transstadial transmission of the KFDV in the ecosystem from larva to nymph to adult (Murhekar et al., 2015). The incubation period of the KFDV is between 3-8 days after the bite of an infected tick. The infection causes severe febrile illness in most of the monkeys. The “hot spot” of the disease will generate when the KFDV infected monkey death occurs, and the infected ticks will drop from the body, which leads to the further transmission of disease. Humans get infected from an infected tick bite or contact with the infected dead monkeys or movement to nearby place, where the recent monkey death is reported (Eldridge et al., 2004; Pattnaik, 2006). Since the disease is associated with monkey deaths, it is also known as monkey disease or monkey fever. There are no human-to-human transmission reports (Murhekar et al., 2015; Upadhyaya, Murthy, & Murthy, 1975).

The KFDV genome spans approximately 11 kb in length, it is a positive-sense, single-stranded RNA genome, that encodes a single polyprotein (3416 aa length). The encoded polyprotein is cleaved post-translationally into a total of three structural and seven non-structural proteins (Dodd et al., 2011). The three structural proteins are capsid C, pre-membrane (prM) or transmembrane protein M, and envelope E protein. The seven non-structural (NS) proteins are NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5. The role of these structural and non-structural proteins are similar to the ones in other viruses of the *Flavivirus* genus. The structural proteins are mainly involved in the formation of the viral particles and are involved in viral fusion with host cells (Z. Li et al., 2017). The detailed mechanism of structural proteins in *Flavivirus* can be found in references (Blazevic et al., 2016; Z. Li et al., 2017). The majority of the NS proteins in *Flavivirus* are multifunctional and are involved in RNA replication, virion assembly, and evasion of innate immune responses. The main functions of NS proteins include NS1, a large glycoprotein, necessary for the negative-strand RNA synthesis (B D Lindenbach & Rice, 1997; Brett D. Lindenbach & Rice, 1999; Muylaert et al., 1997). NS2A is involved in the generation of virus-induced membranes during virus assembly (Kümmerer & Rice, 2002; Leung et al., 2008). NS2B is a cofactor required for the activity of NS3 (Arias et al., 1993; Chambers et al., 1993), a large multifunctional protein having a serine protease activity (with NS2B as a cofactor), 5' RNA triphosphatase activity, nucleoside triphosphatase activity, and helicase activity (H. Li et al., 1999; Warrener et al., 1993; Wengler & Wengler, 1991). NS4A is a membrane protein involved in membrane rearrangement and is involved in the formation of the viral replication complex (S. Miller et al., 2007; Roosendaal et al., 2006). NS4B protein modulates viral RNA synthesis and inhibits host cells type I interferon response (Grant et al., 2011; Muñoz-Jordán et al., 2005; Umareddy et al., 2006). Among the NS proteins, NS5 is the largest *flaviviral* protein, with multiple enzymatic activities, namely the RNA-dependent RNA polymerase (Ackermann & Padmanabhan, 2001; Guyatt

et al., 2001), the RNA guanylyl transferase (Issur et al., 2009), and the *N*-7 guanine and 2'-O ribose methyltransferase (Dong et al., 2012; Egloff et al., 2002; Koonin, 1993; Ray et al., 2006). Currently, there is no specific anti-viral drug therapy for KFD and existing treatment includes the maintenance of proper hydration and circulation by transfusion of intravenous fluids, colloids, or blood products depending upon the patient's condition (*Kyasanur Forest Disease (KFD) | CDC*, n.d.; Munivenkatappa et al., 2018). Recent cases and death reports on the KFD show the severity of the disease and the present vaccine was found to be moderately effective against the KFDV (Kasabi et al., 2013). Genomic analysis of KFDV collected from the different zones shows the mutations in structural and non-structural proteins (Yadav et al., 2020). These reports warrant to development of alternative prevention and control strategies for KFD to avoid large scale outbreaks in other states of India and transmission to other countries. To prevent this disease from the transmission to other states or other countries, new antiviral therapy needs to be developed. With the objective of developing alternatives to tick control and vaccine, the present investigation is focused on developing new antiviral inhibitors against KFDV proteins by using *in silico* drug repurposing methods. The *in silico* drug-repurposing methods were found to be a reliable and fast approach in drug discovery, allowing screening of large existing drugs against a target of interest. Drug repurposing is an effective strategy to identify new use for approved or investigational drugs that are outside the scope of their initial usage, and this approach includes various data-driven and experimental procedures (Ashburn & Thor, 2004; Pushpakom et al., 2018). The repurposed drugs have lower risk and higher safety compared to *de novo* developed drugs, because their toxicity and safety issues are profoundly investigated during the preclinical trials in human/other models. The drug repurposing method was also successfully applied to find potential drugs against other *Flavivirus* genus virus such as Dengue virus (DENV) (Pathak et al., 2017), Zika virus (ZIKV) (Santos et al., 2020; Shiryaev et al., 2017), etc. At present, no experimentally solved 3D structures for the KFD viral proteins are available or stored in Protein Data Bank (Berman et al., 2000), and one of the main criteria in computational drug discovery pipeline is the availability of the reliable 3D structure of a target or protein of interest. However, the non-availability of the experimentally solved structure can be fulfilled by homology modelling. Homology modelling is an important method of choice for obtaining the 3D structure of a protein that does not have any prior experimental structural information and these methods use the fact that evolutionarily connected proteins share a similar structure (Vitkup et al., 2001; Vyas et al., 2012). To the best of our knowledge until the writing of this manuscript, there are no attempts to find antiviral drugs against the KFDV. This motivated us to identify a suitable target among KFDV proteins and design effective antiviral agents against it, using a combination of state-of-the-art bioinformatics and drug discovery methods. With this objective, by employing homology modelling, we built the 3D structures of KFDV proteins and validated them by molecular dynamic simulations. Further, virtual screening of FDA approved drugs and natural products against the selected target was performed. Subsequently, for the

complexes of the target protein and potential hit molecules, molecular dynamic simulations were performed, to estimate the binding free energies.

## **2. Materials and method**

### **2.1 Sequence retrieval and homology modelling**

All annotated protein sequences of the Kyasanur Forest Disease virus were collected from the NCBI gene bank (accession id AFF18434) (Dodd et al., 2011). Since no experimentally solved structures are available for the KFD viral proteins, we employed the template-based modeling approach to develop a three-dimensional protein model. In this regard, a sequence similarity search was performed for all collected protein sequences of KFDV proteins against the sequences with available structural information in the protein data bank (PDB) (Berman et al., 2000) using the BLASTP (Protein BLAST program) available through NIH NCBI. The high similarity sequence or templates with valid structural information was collected and used as an input for the homology modeling. The SWISS-MODEL server (Waterhouse et al., 2018) was used for the homology modelling. Based on the GMQE (Cardoso et al., 2018) (Global Model Quality Estimation) and QMEAN (Benkert et al., 2009) (Qualitative Model Energy Analysis) values, the most reliable homology models were selected. The QMEAN score below 4.0 indicates the high reliability of the predicted structures. The GMQE score ranges between 0 and 1, and the models with a score near 1 are expected to be highly reliable (Patel et al., 2019).

### **2.2 Structural validation**

The structural validation of the homology models was done using the PROCHECK (A. L. Morris et al., 1992) and the MolProbity (V. B. Chen et al., 2010) servers. The quality of the homology model was analyzed using ERRAT program (Colovos & Yeates, 1993). The consistency of the homology model with the template was checked in the ProSA server (Wiederstein & Sippl, 2007). Further, the model reliability was evaluated based on the RMSD (root mean square deviation) calculation between the template and the homology model using UCSF Chimera (Pettersen et al., 2004). Lastly, the structural stability of the model protein was explored by molecular dynamics simulations of free, unbound, protein using AMBER package (Case et al., 2016).

### **2.3 Molecular dynamic (MD) simulations of the homology model**

Initial minimization of the NS3 homology model was executed to discard high energy unfavorable intramolecular interaction using AMBER ff14SB (Maier et al., 2015) force field. All MD simulations were performed with the AMBER16 program (Case et al., 2016). The simulation was carried out by following the protocol provided in the AMBER16 manual and in our previous manuscripts with few modifications. (Kandagalla et al., 2020; Novak et al., 2021) It consists a series of minimization steps, followed by equilibration and production steps.

In brief, the homology model of NS3 was solvated in a truncated octahedral box with 9288 TIP3P water molecules spanning a 12 Å thick buffer. Two sodium cations were added to neutralize the homology protein. Then geometry of the homology protein was minimized. For the first 2500 cycles of the steepest descent and the subsequent 2500 cycles of conjugate gradient optimization method, the modeled protein was restrained ( $k = 10.0 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ ) and only the positions of water molecules were optimized. In the second round of minimization (2500 + 2500 cycles), restraints were removed and the whole system was relaxed. Periodic boundary conditions in all directions were employed. The minimized system was gradually heated from 0 to 300 K and equilibrated using NVT conditions. Finally, productive MD simulation was run with 200 ns simulation time, a time step of 2 fs, and NPT ensemble (pressure was set to 1 atm and temperature to 300 K). The Langevin thermostat with a collision frequency of  $1 \text{ ps}^{-1}$  was used to control temperature. All constraints were removed and only the SHAKE algorithm (Andersen, 1983) was used to fix bond lengths with hydrogen atoms. The Particle Mesh Ewald method (Darden et al., 1993) was used to compute long-range electrostatic interactions. The threshold for non-bonded interactions was set at 11.0 Å. The resulting trajectory was processed and analyzed by the CPPTRAJ module of Amber tools (Roe & Cheatham, 2013).

#### **2.4 Collection of FDA approved drugs and natural products for virtual screening**

A database of approved drugs was collected from the Selleckchem Inc. (WA, USA) webpage, which contains information of approximately 3000 marketed drugs. Salts were excluded from the analysis. From drugs' SMILES 3D structures were generated and optimized using Chimera 1.14 (Pettersen et al., 2004), and saved as PDB files. The same procedure was performed for molecules extracted from the Natural Products Atlas database (version v.2019\_12) (Van Santen et al., 2019), a database of microbial natural products which hold the structure information of 25,523 molecules with names, source organisms, and other data related to the molecules.

#### **2.5 Molecular docking and hit selection**

For molecular docking studies, the collected approved drugs and natural molecules were prepared using AutoDockTools 4 script `prepare_ligand4.py` (G. M. Morris et al., 2009) and saved in the `pdbqt` format. Molecules with more than 15 rotatable bonds and molecular mass above 700 Da were excluded from subsequent calculations. Altogether, 1717 approved drugs and 19250 natural products were included in molecular docking studies.

The receptor, NS3, was prepared by adding Gasteiger charges to each atom, all nonpolar hydrogens were merged, atom types were determined and the prepared receptor was saved as `pdbqt` files using Chimera 1.14 (Pettersen et al., 2004). The active pocket information in the homology model was extracted from the template protein through sequence similarity (discussed later). Additional validation of the active pocket was done using the DoGSiteScorer program (Volkamer et al., 2012) using the default parameters available on the protein plus server (<https://proteins.plus/>). Based on the

active site information, the grid box was centered at Cartesian coordinates 25.6, 10.1, and 11.4, in the proximity of His33 residue, while the box size was set to  $20 \times 20 \times 20 \text{ \AA}^3$ . The number of modes and exhaustiveness were both set to 100. The plausibility of docked conformations for hit molecules was checked by visual inspection. The prepared ligands were docked to the active pocket of the NS3 receptor using the Autodock Vina (Trott & Olson, 2010). Validation of this approach was assessed by redocking. Docking experiments were performed on South Ural State University's supercomputer Tornado, based on Intel® Xeon X5680 processors @ 3.3 GHz. During the docking procedure, the NS3 protein was regarded as rigid while all ligand's single bonds could rotate freely.

As part of a selection of the hit molecules, several additional filters have been applied. First, the synthetic accessibility of all the top hit molecules was predicted using the SWISS ADME server (Daina et al., 2017). It predicts synthetic accessibility (SA) based on 1024 fragmental contributions (FP2) modulated by size and complexity penalties, trained on 12,782,590 molecules, and tested on 40 external molecules ( $r^2 = 0.94$ ). The SA score ranges from one to 10, which shows the complexity range from easy synthesis to difficult synthesis. Next, we evaluated the potential PAINS liabilities of the hit molecules using the servers SWISS ADME (Daina et al., 2017), CBLigand (L. Wang et al., 2013), and FAF-Drugs (Miteva et al., 2006). The colloidal aggregation potential of the hit molecules was assessed using the Aggregate Advisor server (advisor.bkslab.org) (Irwin et al., 2015). Further, the drug-likeness and physicochemical properties of the top hits were evaluated using the SWISS ADME server (Daina et al., 2017). Finally, promiscuity of the hit molecules was assessed using the Pubchem server (Kim et al., 2019). If any hit had been reported to have experimental activity against other targets, it had been defined as promiscuous.

The above-mentioned filters were applied to the molecules obtained from the Natural Products Atlas database because those are compounds of natural origin and in most cases with unknown ADMET properties, so their appropriateness as a drug has to be verified. The top three hit molecules satisfying the above criteria were selected for further investigation using molecular dynamics simulations. For approved drugs, ADMET properties were not calculated. But, as the most promising hits, we identify well known antiviral drugs, and the top four protease inhibitors were selected for subsequent MD simulations.

## **2.6 Molecular dynamics simulation of the NS3 protein complexed with hit ligands**

The initial structures for MD simulations of the NS3 protein complexed with hit ligands bound into the binding site were extracted from docking studies. The protein parameterization was done in the AMBER ff14SB (Maier et al., 2015) force field. Ligands were parametrized in the AMBER 16 with the help of *antechamber* module (J. Wang et al., 2006) and the GAFF force field (J. Wang et al., 2004). Next, the complex was solvated in a truncated octahedral box with TIP3P (Jorgensen et al., 1983) water molecules spanning a  $12 \text{ \AA}$  thick buffer, and neutralized by two sodium ions. The



geometry of the complex was minimized following the same procedure as described above for the unbound protein. Then, followed by productive and unrestrained MD simulations for 200 ns with constant temperature (300 K) and pressure (1 atm).

The analysis of the MD trajectories was performed using the CPPTRAJ (Roe & Cheatham, 2013) module of AmberTools. The Python written MMPSBA.py (B. R. Miller et al., 2012) program was used to calculate the binding free energies,  $\Delta G_{bind}$ , of the ligand-protein complexes using established molecular mechanics - generalized Born surface area (MM-GBSA) protocol (Genheden & Ryde, 2015; Hou et al., 2011):

$$\Delta G_{bind} = \Delta H - T\Delta S \approx \Delta E_{MM} + \Delta G_{sol} - T\Delta S \quad (1)$$

$$\Delta E_{MM} = \Delta E_{internal} + \Delta E_{electrostatic} + \Delta E_{vdW} \quad (2)$$

$$\Delta G_{sol} = \Delta G_{GB} + \Delta G_{SA} \quad (3)$$

$\Delta E_{MM}$ , the change of the gas phase MM energy, is a sum of the bond, angle and dihedral energy ( $\Delta E_{internal}$ ), and electrostatic ( $\Delta E_{electrostatic}$ ) and van der Waals ( $\Delta E_{vdW}$ ) energies.  $\Delta G_{sol}$  is a change of the solvation free energy. It is composed of polar component ( $\Delta G_{GB}$ , electrostatic solvation energy) and non-polar, non-electrostatic solvation contribution ( $\Delta G_{SA}$ ).  $-T\Delta S$  is conformational entropy upon binding. A total of 5000 snapshots were used for the MM-GBSA analysis, which were collected from the corresponding MD trajectories sampled at regular time steps starting from 100 ns to the end of the simulation. Entropic contributions ( $-T\Delta S$ ) were evaluated from 50 structures from the last 100 ns of the simulations. The calculated binding free energies from MM-GBSA were decomposed according to established procedure (Gohlke et al., 2003; Rastelli et al., 2010) into specific residue contribution as a per-residue basis.

## 2.7. Enzyme-ligand complementarity assessment

Enzyme-ligand complementarity assessment is based on the evaluation of the squared correlation coefficient of the linear relationship between the complementarity factor (CF) and the sum of distances SUMRLRE determined for each  $m$ th point of the intermolecular enzyme-ligand space in their complex without consideration of hydrogen atoms (Rimac et al., 2020):

$$CF = a + b \times SUMRLRE \quad (4)$$

where  $SUMRLRE = R_{ml} + R_{me}$ ,  $R_{ml}$  is the distance between the  $m$ th point and the  $l$ th atom of the ligand the contribution of which to the electron density is maximum at the point,  $R_{me}$  between the  $m$ th point and the  $e$ th atom of the enzyme the contribution of which to the electron density is maximum at the point.

$$CF = \log \left( \frac{\rho_E * \rho_{E(CNT)}}{N_e} \right) + \log \left( \frac{\rho_L * \rho_{L(CNT)}}{N_l} \right) \quad (5)$$

where  $\rho_E$  and  $\rho_L$  are the electron densities of the enzyme and the ligand at the  $m$ th point ( $e/\text{\AA}^3$ ),  $\rho_{E(CNT)}$  and  $\rho_{L(CNT)}$  are the electron densities at the centers of the  $e$ th enzyme atom and the  $l$ th ligand atom respectively,  $\rho_E$ ,  $\rho_L$ ,  $\rho_{E(CNT)}$  and  $\rho_{L(CNT)}$  are computed using AlteQ orbital-free quantum-chemical approach,  $N_e$  and  $N_l$  are the numbers of the outer electrons of the  $e$ th enzyme atom and the  $l$ th ligand atom respectively.

Recently, it has been shown that a squared correlation coefficient of the relationship (4) for experimental enzyme-ligand complexes exceeds 0.90 while the docked complexes with incorrect binding pose of the ligand and a high RMSD value are characterized by a low value of the squared correlation coefficient and a maximal value of CF ( $\text{maxCF}$ )  $\ll -5.0$  (Kandagalla et al., 2021; Rimac et al., 2020, 2021).

### 3. Results and discussion

#### 3.1 Sequence retrieval, homology modelling and structural validation

The protein sequences corresponding to protein C, protein M, protein E, and non-structural proteins NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5 were collected from the KFDV genome (accession id: AFF18434) and used for the sequence similarity analysis on the NCBI BLAST server against the sequences with available structural information in the Protein Data Bank. The KFDV protein sequences of protein M, protein E, NS1, NS3, NS4A, and NS5 showed high sequence similarity with proteins their structural information are available. Other KFDV protein sequences (protein C, NS2A, NS2B, and NS4B) were not showing any sequence similarity and hence were excluded from the analysis.

The homology models of protein M, protein E, NS1, NS3, NS4A, and NS5 were designed using the Swiss Modeller. As the most reliable model the protein sequence NS3 was pinpointed, based on the GMQE (0.79) and QMEAN (-0.10) scores. The other KFDV protein models did not show satisfactory scores, so are excluded from the analysis. The NS3 gene of KFDV codes the protease. Zika virus protease, with crystal structure deposited in Protein Data Bank under id 6JPW (Nitsche et al., 2019), was the template used to build the NS3 3D structure model. This protein was selected because it has 50 % sequence similarity with the NS3 of KFDV. In homology modeling, the minimum sequence similarity above 30% is essential for the successful template selection. Both KFDV and Zika virus belongs to the same *Flavivirus* family.

The coordinates of the refined NS3 homology model were submitted to the MolProbity and PROCHECK server to evaluate the integrity of the 3D structure (Table 1). Metrics assessed include Ramachandran backbone dihedral angle distributions, bond lengths, and angle distributions, atom-atom contact clashes, among other measures (Table 1). As expected, based on the high sequence similarity, the NS3 homology model shows improved scores relative to the template crystal structure

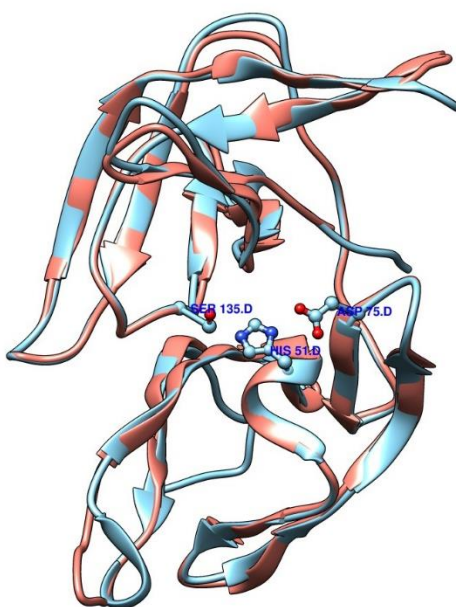
used to create the model. The MolProbity analysis corroborate those results, shows scores that are more favorable in 5 out of 6 categories, including the overall MolProbity score (Table 1, bold entries). Similarly, the program PROCHEK yields similar values in most of the eight categories, although the model outscored the template in only three of the eight categories. The ERRAT Server measures the structural error at each amino acid residue in the 3D structural model and gives a quality score. The overall quality of the NS3 homology model was 89.68% (Figure S11). Evaluation of NS3 homology model and template (PDB ID: 6JPW) with ProSA server revealed a compatible Z score value of -5.44 and -5.43, respectively. These results explain that the residue energies along with pair energy, combined energy, and surface energy were all negative and had comparable surface energy affinity with the template. The Z scores of modeled NS3 and 6JPW (template) proteins are shown in Figure SI 2. The RMSD between the NS3 homology model and the template was found to be 0.153 Å. The low RMSD (< 1 Å) indicates structural conservation, making it a good system for homology modeling (Umamaheswari et al., 2010). Overall, the homology model appears to be reliable, in terms of geometry and energy contours from a protein structure perspective, compared to the template.

**Table 1.** Structure validation for the NS3 homology model.

|                    |                  |                       | 6JPW template (NS3 protease of Zika virus) <sup>a</sup> | homology model (NS3 protease of KFD virus) <sup>a</sup> |
|--------------------|------------------|-----------------------|---------------------------------------------------------|---------------------------------------------------------|
| MolProbity         | atom contacts    | clash score           | 6.29 (94 <sup>th</sup> percentile)                      | <b>1.35 (99<sup>th</sup> percentile)</b>                |
|                    | protein geometry | poor rotamers         | 5 (0.84 %)                                              | <b>0</b>                                                |
|                    |                  | favored rotamers      | 560 (94.28 %)                                           | <b>114 (100.00 %)</b>                                   |
|                    |                  | Ramachandran outliers | <b>1 (0.13 %)</b>                                       | <b>2 (1.37 %)</b>                                       |
|                    |                  | Ramachandran favored  | <b>746 (97.77 %)</b>                                    | 137 (93.84 %)                                           |
|                    |                  | MolProbity score      | 1.39 (98 <sup>th</sup> percentile)                      | <b>1.27 (99<sup>th</sup> percentile)</b>                |
|                    |                  | PROCHECK              | Ramachandran plot                                       | most favored                                            |
| additional allowed | <b>11.7 %</b>    |                       |                                                         | 11.2 %                                                  |
| generously allowed | 0.3 %            |                       |                                                         | 2.6 %                                                   |

|                         |                       |            |                |                |
|-------------------------|-----------------------|------------|----------------|----------------|
|                         |                       | disallowed | 0.0 %          | 0.0 %          |
| residue stereochemistry | Morris classification |            | <b>1, 1, 2</b> | <b>1, 1, 2</b> |
|                         | bad contacts          |            | 46             | <b>0</b>       |
| overall properties      | G-factors             |            | <b>-0.01</b>   | -0.23          |
|                         | planar groups         |            | <b>99.6 %</b>  | 86.3 %         |

<sup>a</sup> Bolded entries indicate more favorable characteristics.

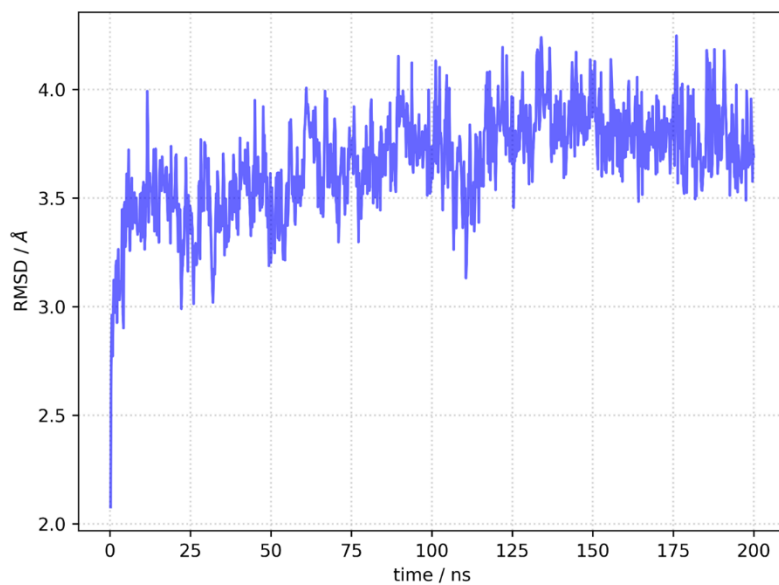


**Figure 1.** Overlaid structures of template protein (6PJW, sky blue) and the homology NS3 (salmon).

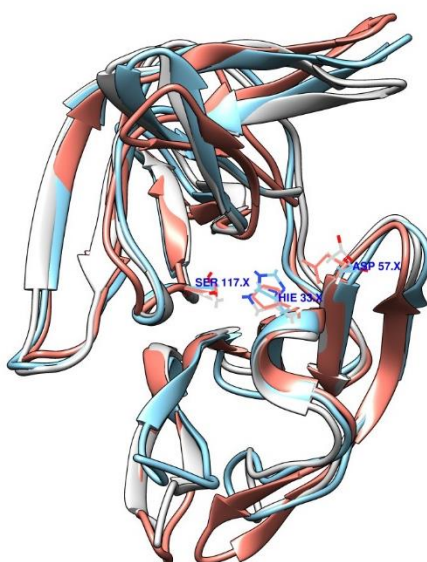
### 3.2 Molecular dynamic simulation of the homology model

The MD simulation was performed for the free, unbound NS3 homology model to gain insight into the stability and its dynamic properties. The simulation was carried out for 200 ns at 300 K temperature in NPT conditions. The evolution of the root mean square deviation (RMSD) of the backbone atoms of free NS3 protein is shown in Figure 2, indicating the overall structural fluctuation. It could be concluded that the NS3 model evolved to a stable state after approximately 120 ns, and it continues to explore conformational space around local minimum on the potential energy hypersurface. In that period, the RMSD compared to the initial structure varies in the range between 3.4 Å and 4.4 Å, with the mean value of 3.82 Å. Since RMSD of the backbone atoms at 120 ns and 200 ns of the simulation is only 1.33 Å, with conserved secondary structure and minor fluctuations of unstructured loops (Figure 2), the structure of the NS3 homology model was extracted from the

simulation and used as a receptor for the virtual screening. The structures obtained from the MD simulation at 120 ns and 200 ns time were superimposed to the initial input NS3 model to find fluctuating domains (Figure 3). The conserved residues of the catalytic pocket (Ser117-His33-Asp57) are not fluctuating significantly. As expected, unstructured parts of the protein, connecting  $\beta$  sheets have the highest conformational freedom and fluctuate the most.



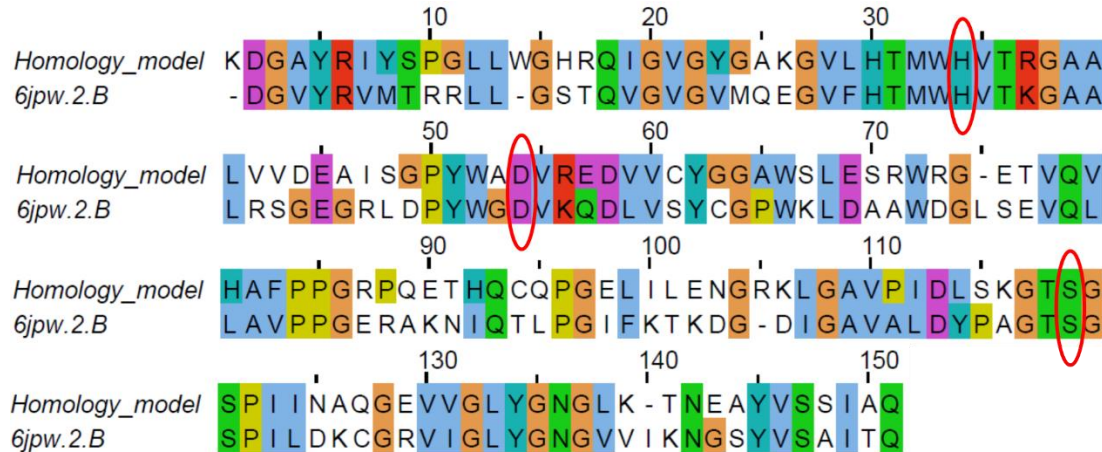
*Figure 2. Backbone RMSD of modelled NS3 protein for 200 ns molecular dynamics simulation.*



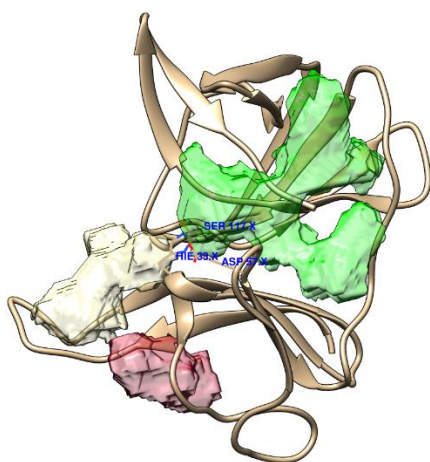
*Figure 3. Superimposed structures of NS3 protein model (salmon) and the structures at 120 ns (sky blue) and 200 ns of simulation (gray).*

### 3.3 Molecular docking analysis

The prepared database of the approved drugs from the Selleckchem Inc. (WA, USA) and the bacterial and fungal natural products from the Natural Products Atlas were used to screen against the putative active site of NS3 protein using molecular docking studies. The enzyme is mainly involved in the RNA replication process and in countering a host's innate immune system. The active site of the NS3 homology protein was identified by comparison with the active site of the template crystal structure (PDB ID: 6JPW). As the template, the NS3 protein of the Zika virus was used, with whom it shares the sequence homology of 50 % (Figure 4). Active site residues in both template and homology models are highly conserved. The NS3 of the Zika virus is a chymotrypsin type serine protease, with classical catalytic triad Ser135-His51-Asp75 necessary for its activity. KFDV, being in the same virus family of *Flavivirus* like Zika virus, has the same conserved Ser117-His33-Asp57 catalytic triad (Figure 1). To corroborate binding site around Ser117-His33-Asp57 residues as the potential target for drug binding, additional cavity search, and its characterization has been performed using the DoGSiteScorer web server. Three potential binding sites were identified (Figure 5), and pocket 1, with the catalytic triad in the heart of the pocket, has the highest drug score (0.75, compared to 0.59 and 0.44 for pockets 2 and 3, respectively). Additional geometrical parameters, like volume and surface, of all pockets, are provided in supplementary table SII. Successful identification of the active site of the KDFV NS3 protein, enable us to center grid box for subsequent docking studies near CE1 atom type of His33 residue.



**Figure 4.** Sequence alignment of the template (Zika virus NS3 protein, PDB ID: 6JPW) and the homology model of the NS3 protease of KFD virus. The conserved active pocket residues are highlighted by the red circles.

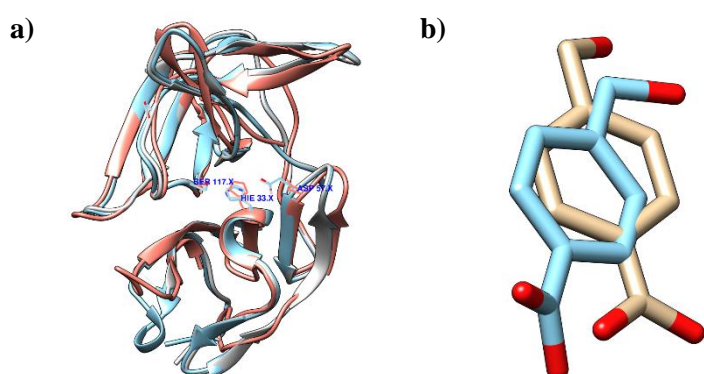


**Figure 5.** NS3 binding pockets identified by DoGSiteScorer: pocket 1 (green), pocket 2 (yellow), and pocket 3 (red), with the catalytic triad residues (Ser117-His33-Asp57) labelled in blue.

After docking, all compounds from the Natural Products Atlas database were ranked based on the docking score. Previous reports have shown the drawbacks of hit selection based on the binding energy without additional computational calculations (Y. C. Chen, 2015). Several computational techniques are available to filter out hit molecules (Jacobs et al., 2020). To further filter out the potential hit molecules, we implemented multiple filtering criteria such as synthetic accessibility, PAINS liabilities, colloidal aggregation, drug-likeness, promiscuity, and ADMET properties. Hit molecules should not be a cytochrome P450 inhibitor (predicted for CYP1A2, CYP2C9, CYP2C19, CYP2D6, and CYP3A4 enzymes), neither a Pgp substrate and no blood brain barrier (BBB) permeant. In a drug-likeness analysis, we evaluated Lipinski's rule of five. Violations in the molecular weight (< 500 Da) are allowed in selecting the hit molecules because several marketed drugs have a molecular weight above 500 Da (e.g., Atorvastatin, Paclitaxel, Romidepsin, Vincristine, etc.). The molecules violating two or more parameters of Lipinski's rule of five are considered not to be bioavailable orally (Ya'u Ibrahim et al., 2020). After employing all the above filtering criteria, top hit molecules with the best docking scores and the potential drug related properties were collected and are shown in Table 2 (the complete details of the top hits can be found in supplementary file 1). Among the natural molecules, the molecule NPA020264 (Aspertryptanthrin C) showed the highest binding energy and met all filtering criteria, followed by the molecule NPA008771 (Lolitre E), NPA015043 (Chaetominine), NPA014308 (Lolitre F), and NPA010113 (Macrosphelide B). The hit molecules NPA008771 and NPA014308 are having a molecular weight above 500 Da and their synthetic accessibility score was found above 7, which shows their complexity during the synthesis. Besides, their calculated LogP values were found to be above 3, which is in the range of many other known aggregators.

Additional validation of the docking method was done by the redocking approach. Redocking is usually used as a fast evaluation of the docking protocol before actual work with the potential ligands.

It includes removing ligand molecules from the targeted receptor and performing docking. Since no experimentally solved structures are available for the NS3 protein of the KFDV, we tested docking protocols on analogous NS3 protein from the Zika virus (PDB ID: 6L4Z) with 4-(hydroxymethyl)benzoic acid in the catalytic pocket. Both viruses belong to the same viral family and NS3 proteins share 50 % sequence similarity, as discussed above in the homology modelling section. First, to check the structure similarity between 6L4Z (redocking) and 6JPW (homology modelling) the RMSD of backbone atoms for two crystal structures was found to be 1.477 Å, indicating structural conservation (Figure 6a). After redocking of 4-(hydroxymethyl) benzoic acid to the 6L4Z, RMSD between redocked and experimental ligand structure was calculated (Figure 6b) using Chimera rmsd function. Since RMSD is low (1.132 Å) and visual pose inspection suggested similar ligand orientation within the pocket, the docking procedure is considered to be reliable and structures obtained by docking could be used as initial structures for MD simulations.

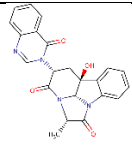
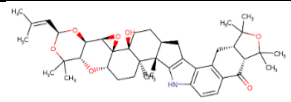
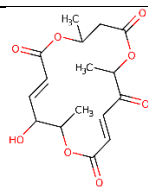


**Figure 6.** a) Overlaid structures NS3 protein of the Zika virus used in homology modelling (6JPW, sky blue) and redocking (6L4Z, gray) with the NS3 homology model of KFD virus (salmon). b) Comparison of the crystallized (tan) and redocked pose of the 4-(hydroxymethyl) benzoic acid (light blue) in the active site of the NS3 protein.

**Table 2.** Binding energy and selected properties of top hits molecules form the Natural Product Alert database.

| ATLAS<br>(alternative<br>name)         | ID | 2D structure | BE<br>a  | MW <sup>b</sup> | RB<br>c | RO5<br>d | ADME<br>e | PAI<br>N <sup>f</sup> | Ag<br>g <sup>g</sup> | Pr<br>o <sup>h</sup> | SA<br>i |
|----------------------------------------|----|--------------|----------|-----------------|---------|----------|-----------|-----------------------|----------------------|----------------------|---------|
| NPA020264<br>(Aspertryptanthr<br>in C) |    |              | -<br>8.6 | 498.<br>5       | 0       | NO       | NO        | NO                    | NO                   | N<br>O               | 6.2     |
| NPA008771<br>(Lolitre E)               |    |              | -<br>8.2 | 687.<br>9       | 4       | MW       | NO        | NO                    | YE<br>S              | N<br>O               | 7.7     |

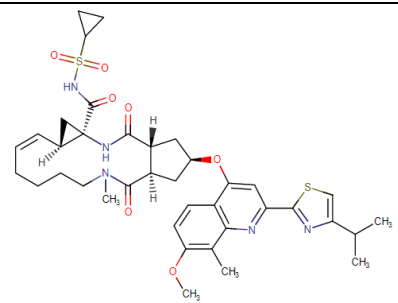


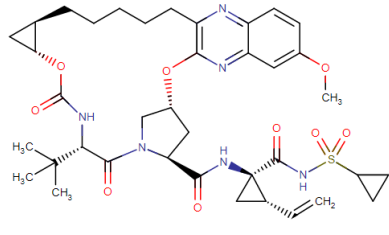
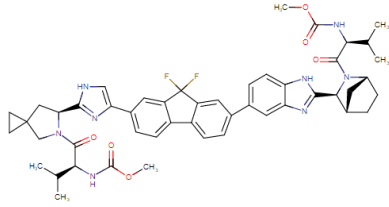
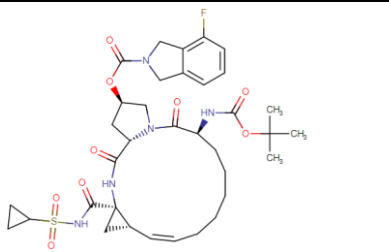
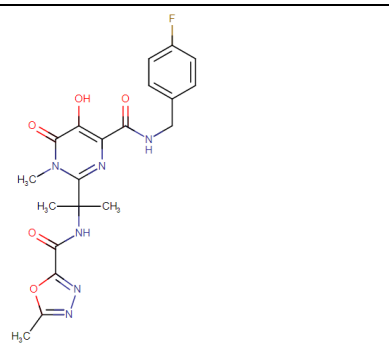
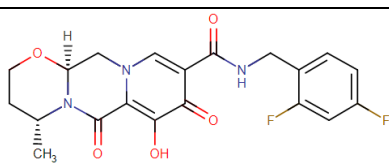
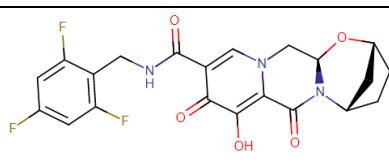
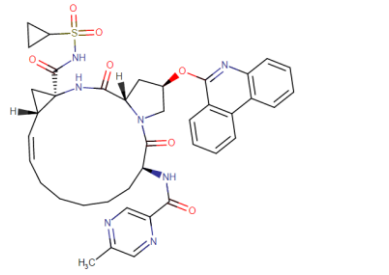
|                                |                                                                                   |          |           |   |    |    |    |         |        |     |
|--------------------------------|-----------------------------------------------------------------------------------|----------|-----------|---|----|----|----|---------|--------|-----|
| NPA015043<br>(Chaetominine)    |  | -<br>8.2 | 402.<br>4 | 1 | NO | NO | NO | NO      | N<br>O | 4.2 |
| NPA014308<br>(Lolitrein F)     |  | -<br>8.0 | 685.<br>9 | 1 | MW | NO | NO | YE<br>S | N<br>O | 7.8 |
| NPA010113<br>(Macrosphelide B) |  | -<br>7.9 | 340.<br>3 | 0 | NO | NO | NO | NO      | N<br>O | 5.0 |

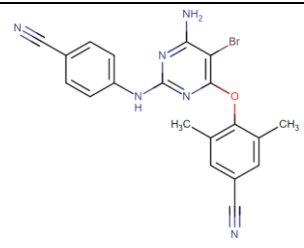
<sup>a</sup>BE = binding energy (kcal/mol), <sup>b</sup>MW = molecular weight (g/mol), <sup>c</sup>RB = number of rotatable bonds, <sup>d</sup>RO5 = Lipinski's rule of 5 violations, <sup>e</sup>ADME = violations of ADME properties (cytochrome P450 enzymes inhibitors, blood brain barrier penetration, Pgp substrate), <sup>f</sup>PAIN = PAINS liabilities, <sup>g</sup>AGG = colloidal aggregation, <sup>h</sup>Pro = promiscuity. <sup>i</sup>SA = synthetic accessibility.

No filters were used for short-listing hits from the approved drug molecules, as their properties were investigated in detail prior to make it available in the market and used to treat different diseases. Previous reports have reviewed the possible strategies of drug repurposing (Jarada et al., 2020; Pushpakom et al., 2018) and in our work for finding the hits, we used drugs' biological activities. In this concern, as hits the antiviral drugs were selected, since they were shown to have specific activity against the viral proteins. The binding energy of the top ten viral hits, together with their current usage, are shown in the Table 3 and the results for other drugs are provided in supplementary file 2. Among the hits, the drug Simeprevir showed the highest binding energy, followed by Grazoprevir, Ledipasvir, Danoprevir, Raltegravir, Dolutegravir, Bictegravir, Paritaprevir, and Etravirine. It is interesting to note here that the drugs Simeprevir, Grazoprevir, Danoprevir, and Paritaprevir are the serine protease inhibitor of the hepatitis C virus (HCV).

**Table 3.** The binding energy of top hits molecules form the list of the approved antiviral drugs.

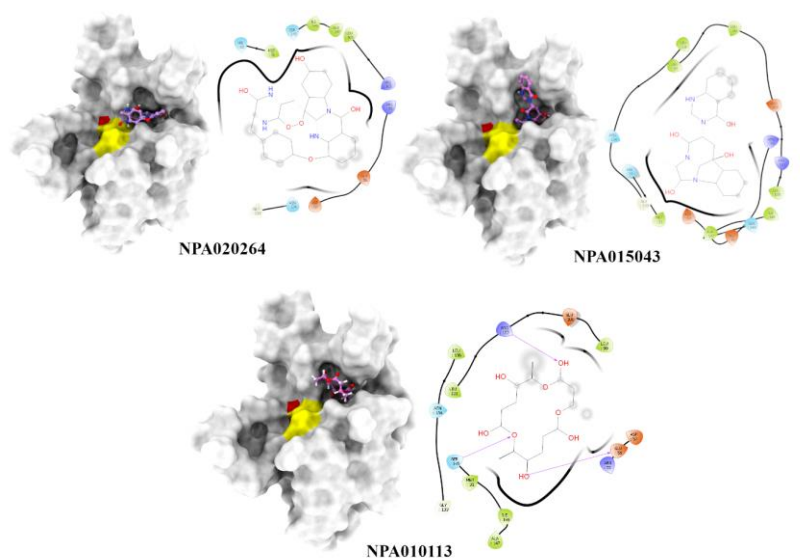
| Drug ID (Drug name) | 2D structure                                                                        | BE <sup>a</sup> | Target            | Mechanism                        |
|---------------------|-------------------------------------------------------------------------------------|-----------------|-------------------|----------------------------------|
| S5015(Simeprevir)   |  | -8.5            | hepatitis C virus | NS3/4A serine protease inhibitor |

|                     |                                                                                                                                                                                                                                                                                                                                           |      |                   |                                           |
|---------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-------------------|-------------------------------------------|
| S3728 (Grazoprevir) |  The structure of Grazoprevir features a central benzimidazole core. It is substituted with a long-chain cyclopropyl ester, a methoxy group, and a complex side chain containing a piperidine ring, a cyclopropylmethyl group, and a sulfonamide moiety. | -7.5 |                   | NS3/4A serine protease inhibitor          |
| S7579 (Ledipasvir)  |  Ledipasvir consists of a central benzimidazole core with two piperidine rings. It is substituted with a difluorophenyl group, a methyl group, and a methyl ester group.                                                                                 | -7.5 | hepatitis C virus | NS5A inhibitor                            |
| S1183(Danoprevir)   |  Danoprevir features a large macrocyclic ring system. It is substituted with a fluorophenyl group, a tert-butyl group, and a cyclopropylmethyl group.                                                                                                    | -7.5 | hepatitis C virus | NS3/4A protease inhibitor                 |
| S2005(Raltegravir)  |  Raltegravir has a central pyrimidine ring substituted with a hydroxyl group, a methyl group, and a side chain containing a piperazine ring and a 4-fluorophenyl group.                                                                                 | -7.5 | HIV               | HIV-1 integrase inhibitor                 |
| S2667(Dolutegravir) |  Dolutegravir features a central pyrimidine ring substituted with a methyl group, a hydroxyl group, and a side chain containing a piperazine ring and a 2,4-difluorophenyl group.                                                                      | -7.3 | HIV               | HIV-1 intergrase inhibitor                |
| S5911(Bictegravir)  |  Bictegravir consists of a central pyrimidine ring substituted with a hydroxyl group, a methyl group, and a side chain containing a piperazine ring and a 2,4,6-trifluorophenyl group.                                                                 | -7.1 | HIV               | HIV-1 integrase strand transfer inhibitor |
| S5404(Paritaprevir) |  Paritaprevir features a large macrocyclic ring system. It is substituted with a sulfonamide group, a piperidine ring, and a pyridine ring.                                                                                                            | -6.9 | hepatitis C virus | NS3/4A serine protease inhibitor          |

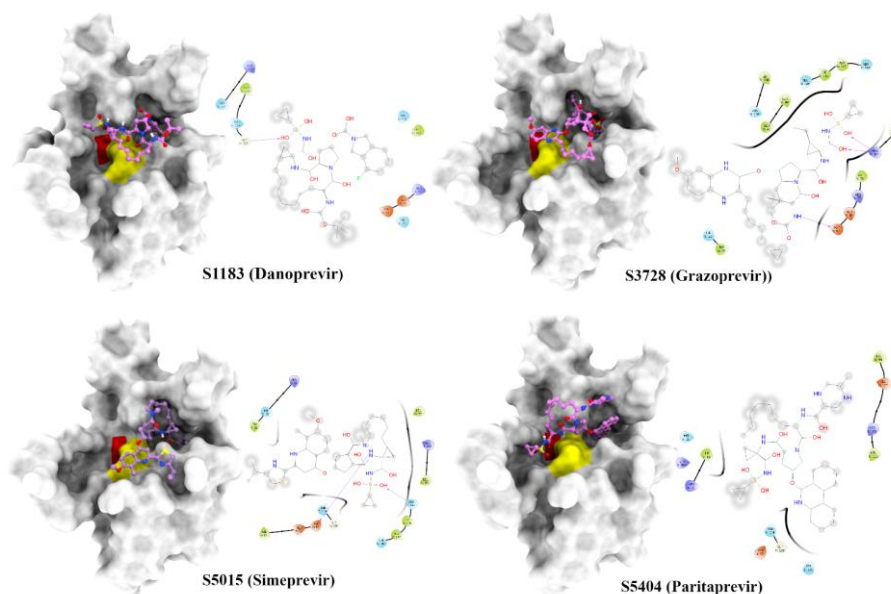
|                    |                                                                                   |      |     |                                                |
|--------------------|-----------------------------------------------------------------------------------|------|-----|------------------------------------------------|
| S3080 (Etravirine) |  | -6.8 | HIV | non-nucleoside reverse transcriptase inhibitor |
|--------------------|-----------------------------------------------------------------------------------|------|-----|------------------------------------------------|

<sup>a</sup>BE = binding energy (kcal/mol)

Together, we can observe the relative difference in the binding energies between approved antiviral drugs and natural products and the binding energies of natural products found high against the NS3 protein. Further stability of natural products and approved drugs against NS3 protein were evaluated using molecular dynamic simulations. For simulations, we selected natural molecules NPA020264, NPA015043, and NPA010113, since they passed all our filtering criteria (Figure 7). In approved drugs, we selected S5015 (Simeprevir), S3728 (Grazoprevir), S1183 (Danoprevir), and S5404 (Paritaprevir) as these drugs are known to have pharmacokinetic interactions with serine protease of HCV (Figure 8).



**Figure 7.** Molecular interactions of the selected natural molecules with the catalytic dyad (colored in yellow and red) of the NS3 homology model.



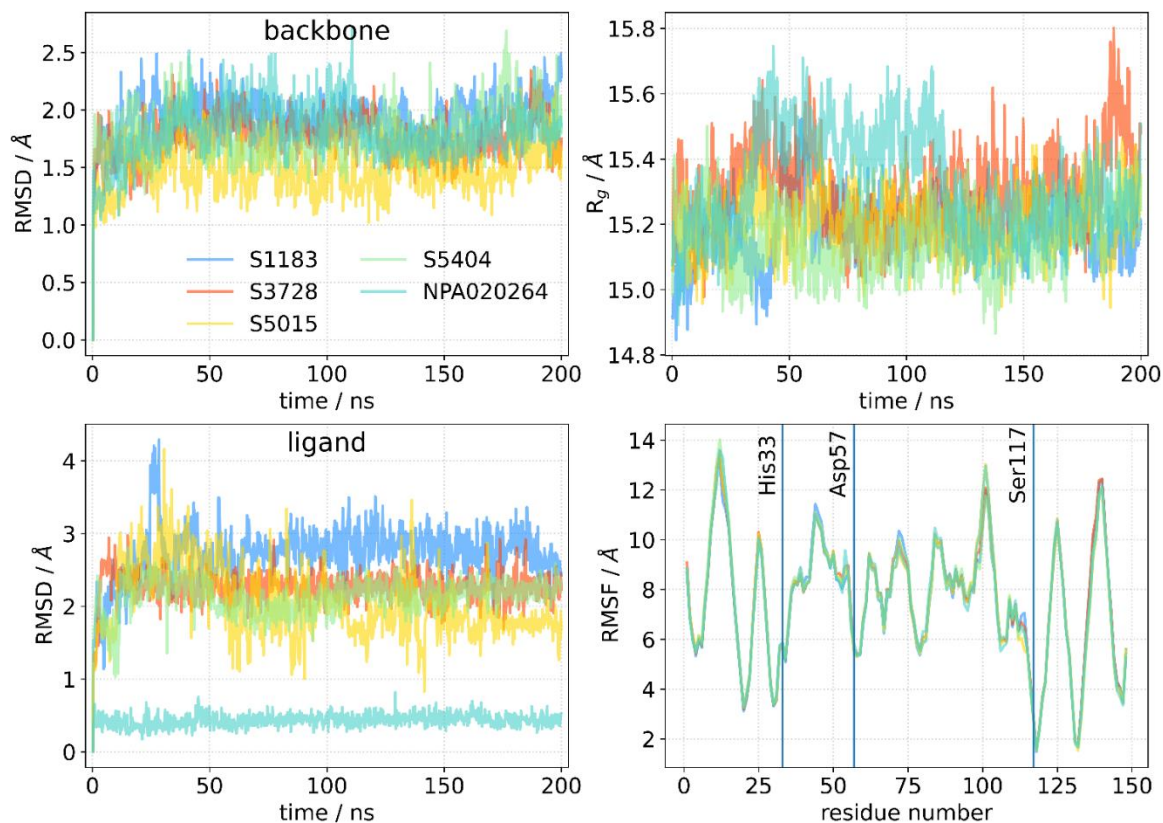
**Figure 8.** Molecular interactions of the selected FDA approved drug molecules with the catalytic dyad (colored in yellow and red) of the NS3 homology model.

### 3.4 Molecular dynamic simulation of complexes of KDFV NS3 protein

Molecular dynamics simulations of KDFV NS3 protein complexes with selected natural products or antiviral drugs were performed to investigate the stability of the complexes, to evaluate binding energies and to select, based on those criteria, the compound with the highest probability to become an inhibitor of KDFV NS3 enzyme. Unfortunately, a closer analysis of the 200 ns trajectories of complexes with two natural products (NPA010113 and NPA015043) showed that they were moving on the surface of the enzyme, completely leaving the catalytic cavity and breaking initial interactions with His33 and Asp57, residues from catalytic triad (Figure SI3). Since those compounds are not likely to be good inhibitors of KDFV, they are excluded from further analysis.

On the other hand, the radius of gyration ( $R_g$ ), root mean square deviation (RMSD), and root mean square fluctuations (RMSF) obtained from the 200 ns molecular dynamics simulations indicate that drugs S5404, S5014, S3728, and S1183 and natural compound NPA020264 may form stable complexes with KDFV NS3 protein (Figure 9). The  $R_g$  may be considered as an indicator of the size and the compactness of the protein. It is defined as the mass weighted root mean square distance of selected atoms from their center of mass. Usually, globular proteins are more compact and have smaller variations in the  $R_g$ , compared to expand and unstructured proteins. (Aouidate et al., 2020) Figure 9 shows the time variation of the radius of gyration of complexes of KDFV NS3 protein with selected antiviral drugs and natural products during 200 ns of molecular dynamics simulation. The  $R_g$  curves for complexes of S1138, S5015, and S5404 follow a similar pattern, with the mean values (standard deviations) being 15.18 (0.10) Å, 15.20 (0.09) Å and 15.16 (0.10) Å, respectively. S3728

and NPA020264 make protein slightly less compact, with mean values of  $R_g$  being 15.31 Å and 15.32 Å, respectively. Also, they experience greater fluctuations, especially in Leu11-Leu12-Trp13 loop region. Since the distance between minimum and maximum mean values of  $R_g$  is 0.16 Å (approximately 1 % of mean  $R_g$ ), all complexes might be considered stable.



**Figure 9.** Analysis of the MD trajectories of KFDV NS3 protein and selected potential inhibitors. RMSD (left), radius of gyration and RMSF (right).

The RMSD, a structural parameter measuring the difference between structures, was computed for each trajectory (Figure 9). In the present study, the geometries of complexes at regular time steps were translated and rotated with the aim to overlay with reference structure (initial geometry, at time  $t = 0$ ) in a way to minimize the RMSD value. By tracking the evolution of the RMSD over the whole trajectory it is possible to get an idea about the stability of the complexes and detect possible conformational changes. We monitored the RMSD of the protein's backbone and of the ligands. The lowest mean RMSD has complex with S5015 ( $1.48 \pm 0.21$  Å), while the complex with S1183 has the highest RMSD value ( $1.94 \pm 0.18$  Å). The results of the analysis of protein's backbone RMSD are in line with a radius of gyration results, demonstrating stability and absence of conformational changes. Besides the dynamics of the target protein, we were interested in the dynamics of the ligands bound to the catalytic pocket. All atom RMSD of the ligands is presented on the left lower panel of Figure 9. While antiviral drugs have mean RMSD above 2 Å, the mean RMSD along 200 ns trajectory of natural product NPA020264 is only  $0.43 \pm 0.09$  Å. This can be explained by the more rigid structure

of NPA020264, an indole diketopiperazine alkaloid, which has only two rotatable bonds. On the other hand, antivirals are much more flexible, which is reflected by higher RMSD values and a higher number of rotatable bonds (7 for S3728 and S5404, and 8 for S1183 and S5015).

The third structural parameter that was monitored is a root mean square fluctuation (RMSF), a property revealing flexibility of each individual residue. Usually, the RMSF per residue versus residue labels (residue number in primary structure) are plotted (Figure 9). RMSF curves follow the same pattern independent from the bound ligand, with only minor differences connected with perturbations introduced by ligands. From the graph is evident that residues Leu11-Leu12-Trp13, a part of a loop region connecting two  $\beta$  sheets, have the highest flexibility. The other two flexible regions include Thr138-Asn139-Glu140-Ala141 and Glu100-Asn101-Gly102 loops. The least flexible residues are Gly118 and Tyr132, both being in close contact with Ser117. It is interesting to note that all three amino acids forming the catalytic triad are showing a certain level of rigidity, especially Ser117, being in the vicinity of (local) minima of the RMSF curve. This might be an indication that during evolution the structure of the enzyme has been optimized to minimize fluctuations of residues having an important catalytic role. (Hedstrom, 2002)

MM-GBSA method is a tested and validated method for estimation of the binding free energies,  $\Delta G_{\text{bind}}$ . The solvation free energies were evaluated by solving the generalized Born equation, and the total binding free energies are a sum of various contributions (equations 1-3). Selected contributions are listed in Table 4. The dominant role in binding have van der Waals interactions, since net polar contribution is positive for all complexes. S5015 and NPA020264 have the lowest estimated binding free energy (without the unfavorable entropy contribution (Sun et al., 2018)) among studied compounds. With entropic component included, NPA020264 has the lowest binding free energy,  $-8.9 \text{ kcal mol}^{-1}$ , followed by S3728 ( $-4.9 \text{ kcal mol}^{-1}$ ). S5015 has the most favorable van der Waals interaction energy, while nonpolar contribution to the solvation free energy is almost twice the solvation free energy for NPA020264. Based on binding free energy results, NPA020264 might be the best inhibitor of KDFV NS3 protein.

**Table 4.** Energy contributions to the binding free energy (in  $\text{kcal mol}^{-1}$ ) for KDFV NS3 protein and potential inhibitors obtained by MM-GBSA approach. The entropic contribution was estimated for  $T = 298.15 \text{ K}$ .

|                                                   | S1183  | S3728  | S5015  | S5404  | NPA020264 |
|---------------------------------------------------|--------|--------|--------|--------|-----------|
| $\Delta G_{\text{bind}}$                          | -2.04  | -4.90  | -3.73  | -2.99  | -8.87     |
| $\Delta G_{\text{bind}} (\text{w/o } -T\Delta S)$ | -25.14 | -19.20 | -32.24 | -25.89 | -31.94    |
| $\Delta E_{\text{vdW}}$                           | -35.96 | -33.12 | -43.03 | -42.58 | -39.90    |
| $\Delta E_{\text{electrostatic}}$                 | -49.78 | -9.38  | -64.52 | -22.25 | -28.41    |
| $\Delta G_{\text{GB}}$                            | 65.42  | 27.35  | 81.18  | 43.68  | 41.42     |
| $\Delta G_{\text{SA}}$                            | -4.82  | -4.05  | -5.88  | -4.75  | -5.03     |

|      |       |       |       |       |       |
|------|-------|-------|-------|-------|-------|
| -TAS | 23.10 | 14.30 | 28.51 | 22.90 | 23.07 |
|------|-------|-------|-------|-------|-------|

MM-GBSA binding free energy decomposition enabled us to identify key residues with dominant contributions to the protein-ligand binding. We set a threshold of  $-1.5 \text{ kcal mol}^{-1}$  for single residue binding free energy to classify it as key residue with a dominant contribution. For KDFV NS3 protein complexed with S5015, there are three key residues - His33 ( $-1.54 \text{ kcal mol}^{-1}$ ), Asn134 ( $-1.90 \text{ kcal mol}^{-1}$ ) and Arg103 ( $-3.11 \text{ kcal mol}^{-1}$ ). His33, as a member of the catalytic triad, and Asn134 are in contact with the oxygen atoms from aminosulfonyl moiety at the beginning of the simulation. Then, during the dynamics, ligand reorients and both His33 and Asn134 establish hydrogen bonds with carbonyl oxygen from the macrocycle. When KFDV NS3 is complexed with NPA020264, Leu150 ( $-1.64 \text{ kcal mol}^{-1}$ ), Arg55 ( $-1.66 \text{ kcal mol}^{-1}$ ), Leu136 ( $-1.94 \text{ kcal mol}^{-1}$ ), Arg103 ( $-2.05 \text{ kcal mol}^{-1}$ ), Leu99 ( $-3.48 \text{ kcal mol}^{-1}$ ) and Glu56 ( $-3.64 \text{ kcal mol}^{-1}$ ) are found to be key residues. Glu56 and Arg103 are forming hydrogen bonds with the ligand, while hydrophobic residues Leu99 and Leu136 interact with the aromatic phenyl moiety of the natural compound. Except for His33 in complexes with S1183 and S5015, no residues from the conserved catalytic triad appear to have a crucial role in ligand binding.

#### 4. Discussion

*Flaviviruses* are vector-borne RNA viruses, they cause a broad spectrum of potentially severe diseases and their recent outbreaks highlight their transmission potential. Presently, *Flaviviruses* are distributed globally, and nearly 400 million people are annually infected with different kinds of *Flavivirus* diseases. The changing epidemiology of *Flaviviruses* raises serious concern for the large-scale emergence of epidemics and the possibility of future pandemics. KFD is one of the important emerging *Flavivirus* diseases in India and in the last decade, the spread of the KFD disease rapidly increasing to newer geographical areas. KFDV is categorized as a risk IV pathogen, according to the Center for disease control and prevention (CDC), every year 400-500 KFDV human cases have been reported with a fatality rate of 3-10% (Shah et al., 2018). Current antiviral therapy is not available against KFD, and the existing vaccine effectiveness was only 62% and 83% after first and second booster doses respectively (Kasabi et al., 2013). with this background, the present work was designed to find suitable small molecule inhibitors against the KFDV through *in silico* drug repurposing method. Since there is no experimentally solved structure of KFDV proteins available, we designed the protein models based on sequence similarity and homology modeling methods. Among the KFDV proteins, a reliable homology model for the NS3 protein was developed, with the amino acid region between 1512 to 1660 in the KFDV polyprotein (accession id: AFF18434). The amino acid stretch between 1512 to 1660 within the KFDV polyprotein codes for the serine protease and it shares high sequence similarity (50%) with Zika virus NS3 protease (PDB: 6JPW) (Nitsche et al., 2019). We have not found consistent models for other structural and non-structural proteins of KFDV proteins. The protein structure of the generated NS3 model is found to have more stable geometry and energy

contours compared with the template Zika virus NS3 protease. The quality of the NS3 model is found more satisfactory, as seen in its quality scores measured in different online servers.

As discussed in the results KFDV NS3 protein (149 amino acids sequence) is compared with other *Flaviviruses* for the multiple sequence alignment and phylogenetic analysis with the Alkhumra hemorrhagic fever virus (AHFV), Langat virus (LGTV), tick-borne encephalitis virus (TBEV), Omsk hemorrhagic fever virus (OHFV), and Zika virus. The sequence alignment and phylogenetic analysis show the divergence among the species.

It is important to note here that for the activity of NS3 protease the flexible NS2B cofactor is essentially in *Flavivirus* and without NS2B cofactor the NS3 protease does not fold properly to the functional active chymotrypsin-like conformation (Tomar et al., 2017). In the Zika virus, two distinct conformations (open and close) of protease were observed and it is based on the orientation of NS2B cofactor around the NS3 protein (Kang et al., 2017). In the open conformation, the flexible C-terminal region of the NS2B region does not orient towards the active site and as a result, crucial subpockets and cavity are completely missing in the active site (open conformation can be observed in crystal structure such as 5T1V, 5TFN, and 5TFO). In the closed conformation, a  $\beta$ -hairpin of the NS2B cofactor completing the active site and thereby it facilitating substrate/inhibitor binding (it can be observed in crystal structures such as 5GJ4, 5LC0, and 6JPW). The closed conformation ideal model for the inhibitor discovery (Pathak et al., 2020) and hence we modeled the NS3 protein of KFDV using the closed conformation of NS3 protease of Zika virus (PDB: 6JPW). Along with this, we attempted to mimic the interaction of NS3 protein with NS2B cofactor in KFDV through protein-protein docking, but we did not find a suitable experimentally solved crystal template against NS2B for homology modelling (the results are not shown). Hence, we considered the homology model of the NS3 protein of KFDV for molecular docking and MD simulations. During docking, we centered the grid on the catalytic triad (Ser135-His51-Asp75) and avoided the docking to the NS2B binding region of the NS3 protein.

The molecular docking analysis revealed that the molecules such as NPA020264 (Aspertryptanthrin C), NPA008771 (Lolitre E), NPA015043 (Chaetominine), NPA014308 (Lolitre F), and NPA010113 (Macrosphelide B) from the natural products atlas database showed the highest binding energy with the NS3 homology protein. These natural products originated from the fungus, Aspertryptanthrin C & Chaetominine synthesized from *Aspergillus* sp. Further Lolitre E & Lolitre F from *Epichloë festucae* and Macrosphelide B from *Microsphaeropsis* sp. FO-5050. These natural products also have good drug likeness properties. In the FDA approved drugs, Simeprevir, Grazoprevir, Ledipasvir, Danoprevir, Raltegravir, Dolutegravir, Bictegravir, Paritaprevir and Etravirine showed the highest binding energy with the NS3 homology protein. Molecular interactions of these top hits in the catalytic triad (Ser135-His51-Asp75) of NS3 protein are shown in figure 5 and



6, these interactions revealed that the natural products bind slightly away from the catalytic triad (colored in yellow and red) and that region of the NS3 protein is mainly involved in the interaction with NS2B. Hence, these natural products have the potential to inhibit the interaction of the NS3 protein with the NS2B cofactor. Whereas all the drugs showed their interactions with the catalytic triad (Ser135-His51-Asp75) of the NS3 protein. The present analysis gives the suitable small molecules, which have the highest binding potential against the NS3 homology of KFDV in the absence of the NS2B cofactor, and their binding pattern in the presence of NS2B cofactor needs to be studied further *in vitro*. The hit molecule from the natural origin i.e., NPA020264 (Aspertryptanthrin C) has the capacity to block the interaction between the NS3 protein and the NS2B cofactor binding to the region of the NS3 protein which is mainly involved in the interaction with NS2B cofactor and further evaluation is essential to confirm the interaction of NPA020264 against the NS3 protein of KFDV.

Further, we performed a complementarity assessment for the docked complexes. Most complexes demonstrated good enzyme-ligand complementarity, squared correlation coefficients of the complexes exceed 0.90 (Supplementary file 3), Altogether, the present analysis successfully presented four FDA approved drugs and one natural compound against the NS3 homology model of KFDV in the absence of the NS2B cofactor. The study leads towards antivirals for KFD which is an important and emerging zoonotic disease. Further, the identified antivirals need to be tested in a laboratory for their efficacy.

### **Acknowledgements**

S.K, J.N., M.A.G, and V.A.P, acknowledges the Ministry of Science and Higher Education of Russia (Grant FENU-2020-0019).

### **Conflicts of interest**

The authors declare no conflict of interest.

**Author Contributions:** S.K and B.K. devised the project, the main conceptual ideas and proof outline; S.K., S.B.S and J.N. conducted the computational studies; S.K., J.N. and B.K. wrote the manuscript, M.A.G. and V.A.P. conducted complementary principle analysis; S.K., J.N. and B.K. edited and approved the final version of manuscript. All authors have read and agreed to the published version of the manuscript.

### **5. References**

Ackermann, M., & Padmanabhan, R. (2001). De Novo Synthesis of RNA by the Dengue Virus RNA-dependent RNA Polymerase Exhibits Temperature Dependence at the Initiation but Not Elongation Phase. *Journal of Biological Chemistry*, 276(43), 39926–39937.

<https://doi.org/10.1074/jbc.M104248200>

- Andersen, H. C. (1983). Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics*, 52(1), 24–34. [https://doi.org/10.1016/0021-9991\(83\)90014-1](https://doi.org/10.1016/0021-9991(83)90014-1)
- Aouidate, A., Ghaleb, A., Chtita, S., Aarjane, M., Ousaa, A., Maghat, H., Sbai, A., Choukrad, M., Bouachrine, M., & Lakhlifi, T. (2020). Identification of a novel dual-target scaffold for 3CLpro and RdRp proteins of SARS-CoV-2 using 3D-similarity search, molecular docking, molecular dynamics and ADMET evaluation. *Journal of Biomolecular Structure and Dynamics*, 0(0), 1–14. <https://doi.org/10.1080/07391102.2020.1779130>
- Arias, C. F., Preugschat, F., & Strass, J. H. (1993). Dengue 2 virus ns2b and ns3 form a stable complex that can cleave ns3 within the helicase domain. *Virology*, 193(2), 888–899. <https://doi.org/10.1006/viro.1993.1198>
- Ashburn, T. T., & Thor, K. B. (2004). Drug repositioning: Identifying and developing new uses for existing drugs. In *Nature Reviews Drug Discovery* (Vol. 3, Issue 8, pp. 673–683). Nature Publishing Group. <https://doi.org/10.1038/nrd1468>
- Benkert, P., Künzli, M., & Schwede, T. (2009). QMEAN server for protein model quality estimation. *Nucleic Acids Research*, 37(SUPPL. 2). <https://doi.org/10.1093/nar/gkp322>
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*, 28(1), 235–242. <https://doi.org/10.1093/nar/28.1.235>
- Blazevic, J., Rouha, H., Bradt, V., Heinz, F. X., & Stiasny, K. (2016). Membrane Anchors of the Structural Flavivirus Proteins and Their Role in Virus Assembly. *Journal of Virology*, 90(14), 6365–6378. <https://doi.org/10.1128/jvi.00447-16>
- Cardoso, J. M. S., Fonseca, L., Egas, C., & Abrantes, I. (2018). Cysteine proteases secreted by the pinewood nematode, *Bursaphelenchus xylophilus*: In silico analysis. *Computational Biology and Chemistry*, 77, 291–296. <https://doi.org/10.1016/j.compbiolchem.2018.10.011>
- Case, D. A., Betz, R. M., Cerutti, D. S., T.E., C. I., Darden, T. A., Duke, R. E., Giese, T. J., Gohlke, H., Goetz, A. W., Homeyer, N., Izadi, S., Janowski, P., Kaus, J., Kovalenko, A., Lee, T. S., LeGrand, S., Li, P., C.Lin, Luchko, T., ... Kollman, P. A. (2016). Amber 2016. *University of California, San Francisco*.
- Chambers, T. J., Nestorowicz, A., Amberg, S. M., & Rice, C. M. (1993). Mutagenesis of the yellow fever virus NS2B protein: effects on proteolytic processing, NS2B-NS3 complex formation, and

- viral replication. *Journal of Virology*, 67(11), 6797–6807.  
<https://doi.org/10.1128/jvi.67.11.6797-6807.1993>
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., & Richardson, D. C. (2010). MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, 66(1), 12–21. <https://doi.org/10.1107/S0907444909042073>
- Chen, Y. C. (2015). Beware of docking! In *Trends in Pharmacological Sciences* (Vol. 36, Issue 2, pp. 78–95). Elsevier Ltd. <https://doi.org/10.1016/j.tips.2014.12.001>
- Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Science*, 2(9), 1511–1519. <https://doi.org/10.1002/pro.5560020916>
- Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, 7(1), 1–13. <https://doi.org/10.1038/srep42717>
- Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12), 10089–10092.  
<https://doi.org/10.1063/1.464397>
- Dodd, K. A., Bird, B. H., Khristova, M. L., Albariño, C. G., Carroll, S. A., Comer, J. A., Erickson, B. R., Rollin, P. E., & Nichol, S. T. (2011). Ancient ancestry of KFDV and AHFV revealed by complete genome analyses of viruses isolated from ticks and Mammalian hosts. *PLoS Neglected Tropical Diseases*, 5(10). <https://doi.org/10.1371/journal.pntd.0001352>
- Dong, H., Chang, D. C., Hua, M. H. C., Lim, S. P., Chionh, Y. H., Hia, F., Lee, Y. H., Kukkaro, P., Lok, S. M., Dedon, P. C., & Shi, P. Y. (2012). 2'-O methylation of internal adenosine by flavivirus NS5 methyltransferase. *PLoS Pathogens*, 8(4).  
<https://doi.org/10.1371/journal.ppat.1002642>
- Egloff, M. P., Benarroch, D., Selisko, B., Romette, J. L., & Canard, B. (2002). An RNA cap (nucleoside-2'-O-)-methyltransferase in the flavivirus RNA polymerase NS5: Crystal structure and functional characterization. *EMBO Journal*, 21(11), 2757–2768.  
<https://doi.org/10.1093/emboj/21.11.2757>
- Eldridge, B. F., Scott, T. W., Day, J. F., & Tabachnick, W. J. (2004). Arbovirus Diseases. In *Medical Entomology* (pp. 415–460). Springer Netherlands. [https://doi.org/10.1007/978-94-007-1009-2\\_11](https://doi.org/10.1007/978-94-007-1009-2_11)
- Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-

- binding affinities. *Expert Opinion on Drug Discovery*, 10(5), 449–461.  
<https://doi.org/10.1517/17460441.2015.1032936>
- Gohlke, H., Kiel, C., & Case, D. A. (2003). Insights into Protein–Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras–Raf and Ras–RalGDS Complexes. *Journal of Molecular Biology*, 330(4), 891–913. [https://doi.org/10.1016/S0022-2836\(03\)00610-7](https://doi.org/10.1016/S0022-2836(03)00610-7)
- Grant, D., Tan, G. K., Qing, M., Ng, J. K. W., Yip, A., Zou, G., Xie, X., Yuan, Z., Schreiber, M. J., Schul, W., Shi, P.-Y., & Alonso, S. (2011). A Single Amino Acid in Nonstructural Protein NS4B Confers Virulence to Dengue Virus in AG129 Mice through Enhancement of Viral RNA Synthesis. *Journal of Virology*, 85(15), 7775–7787. <https://doi.org/10.1128/jvi.00665-11>
- Guyatt, K. J., Westaway, E. G., & Khromykh, A. A. (2001). Expression and purification of enzymatically active recombinant RNA-dependent RNA polymerase (NS5) of the flavivirus Kunjin. *Journal of Virological Methods*, 92(1), 37–44. [https://doi.org/10.1016/S0166-0934\(00\)00270-6](https://doi.org/10.1016/S0166-0934(00)00270-6)
- Hedstrom, L. (2002). Serine Protease Mechanism and Specificity. *Chemical Reviews*, 102(12), 4501–4524. <https://doi.org/10.1021/cr000033x>
- Hou, T., Wang, J., Li, Y., & Wang, W. (2011). Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of Chemical Information and Modeling*, 51(1), 69–82.  
<https://doi.org/10.1021/ci100275a>
- Irwin, J. J., Duan, D., Torosyan, H., Doak, A. K., Ziebart, K. T., Sterling, T., Tumanian, G., & Shoichet, B. K. (2015). An Aggregation Advisor for Ligand Discovery. *Journal of Medicinal Chemistry*, 58(17), 7076–7087. <https://doi.org/10.1021/acs.jmedchem.5b01105>
- Issur, M., Geiss, B. J., Bougie, I., Picard-Jean, F., Despins, S., Mayette, J., Hobdey, S. E., & Bisailon, M. (2009). The flavivirus NS5 protein is a true RNA guanylyltransferase that catalyzes a two-step reaction to form the RNA cap structure. *RNA*, 15(12), 2340–2350.  
<https://doi.org/10.1261/rna.1609709>
- Jacobs, A., Rizzo, R. C., Telehany, S. M., Humby, M. S., Dwight McGee, T., & Riley, S. P. (2020). Identification of zika virus inhibitors using homology modeling and similarity-based screening to target glycoprotein e. *Biochemistry*, 59(39), 3709–3724.  
<https://doi.org/10.1021/acs.biochem.0c00458>
- Jarada, T. N., Rokne, J. G., & Alhaji, R. (2020). A review of computational drug repositioning: Strategies, approaches, opportunities, challenges, and directions. In *Journal of Cheminformatics*

- (Vol. 12, Issue 1, pp. 1–23). BioMed Central. <https://doi.org/10.1186/s13321-020-00450-7>
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2), 926–935. <https://doi.org/10.1063/1.445869>
- Kandagalla, S., Rimac, H., Potemkin, V. A., & Grishina, M. A. (2021). Complementarity principle in terms of electron density for the study of EGFR complexes. *Future Medicinal Chemistry*. <https://doi.org/10.4155/fmc-2020-0265>
- Kandagalla, S., Shekarappa, S. B., Rimac, H., Grishina, M. A., Potemkin, V. A., & Hanumanthappa, M. (2020). Computational insights into the binding mode of curcumin analogues against EP300 HAT domain as potent acetyltransferase inhibitors. *Journal of Molecular Graphics and Modelling*, 101, 107756. <https://doi.org/10.1016/j.jmgm.2020.107756>
- Kang, C. B., Keller, T. H., & Luo, D. (2017). Zika Virus Protease: An Antiviral Drug Target. In *Trends in Microbiology* (Vol. 25, Issue 10, pp. 797–808). Elsevier Ltd. <https://doi.org/10.1016/j.tim.2017.07.001>
- Kasabi, G. S., Murhekar, M. V., Sandhya, V. K., Raghunandan, R., Kiran, S. K., Channabasappa, G. H., & Mehendale, S. M. (2013). Coverage and Effectiveness of Kyasanur Forest Disease (KFD) Vaccine in Karnataka, South India, 2005-10. *PLoS Neglected Tropical Diseases*, 7(1). <https://doi.org/10.1371/journal.pntd.0002025>
- Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B. A., Thiessen, P. A., Yu, B., Zaslavsky, L., Zhang, J., & Bolton, E. E. (2019). PubChem 2019 update: Improved access to chemical data. *Nucleic Acids Research*, 47(D1), D1102–D1109. <https://doi.org/10.1093/nar/gky1033>
- Koonin, E. V. (1993). Computer-assisted identification of a putative methyltransferase domain in NS5 protein of flaviviruses and  $\lambda 2$  protein of reovirus. *Journal of General Virology*, 74(4), 733–740. <https://doi.org/10.1099/0022-1317-74-4-733>
- Kümmerer, B. M., & Rice, C. M. (2002). Mutations in the Yellow Fever Virus Nonstructural Protein NS2A Selectively Block Production of Infectious Particles. *Journal of Virology*, 76(10), 4773–4784. <https://doi.org/10.1128/jvi.76.10.4773-4784.2002>
- Kyasanur Forest Disease (KFD) | CDC*. (n.d.). Retrieved July 16, 2021, from <https://www.cdc.gov/vhf/kyasanur/>
- Leung, J. Y., Pijlman, G. P., Kondratieva, N., Hyde, J., Mackenzie, J. M., & Khromykh, A. A. (2008). Role of Nonstructural Protein NS2A in Flavivirus Assembly. *Journal of Virology*, 82(10), 4731–

4741. <https://doi.org/10.1128/jvi.00002-08>

- Li, H., Clum, S., You, S., Ebner, K. E., & Padmanabhan, R. (1999). The Serine Protease and RNA-Stimulated Nucleoside Triphosphatase and RNA Helicase Functional Domains of Dengue Virus Type 2 NS3 Converge within a Region of 20 Amino Acids. *Journal of Virology*, *73*(4), 3108–3116. <https://doi.org/10.1128/jvi.73.4.3108-3116.1999>
- Li, Z., Zhang, J., & Li, H. (2017). Flavivirus NS2B/NS3 protease: Structure, function, and inhibition. In *Viral Proteases and Their Inhibitors* (pp. 163–188). Elsevier. <https://doi.org/10.1016/B978-0-12-809712-0.00007-1>
- Lindenbach, B D, & Rice, C. M. (1997). trans-Complementation of yellow fever virus NS1 reveals a role in early RNA replication. *Journal of Virology*, *71*(12).
- Lindenbach, Brett D., & Rice, C. M. (1999). Genetic Interaction of Flavivirus Nonstructural Proteins NS1 and NS4A as a Determinant of Replicase Function. *Journal of Virology*, *73*(6), 4611–4621. <https://doi.org/10.1128/jvi.73.6.4611-4621.1999>
- Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., & Simmerling, C. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical Theory and Computation*, *11*(8), 3696–3713. <https://doi.org/10.1021/acs.jctc.5b00255>
- Miller, B. R., McGee, T. D., Swails, J. M., Homeyer, N., Gohlke, H., & Roitberg, A. E. (2012). MMPBSA.py: An efficient program for end-state free energy calculations. *Journal of Chemical Theory and Computation*, *8*(9), 3314–3321. <https://doi.org/10.1021/ct300418h>
- Miller, S., Kastner, S., Krijnse-Locker, J., Bühler, S., & Bartenschlager, R. (2007). The non-structural protein 4A of dengue virus is an integral membrane protein inducing membrane alterations in a 2K-regulated manner. *Journal of Biological Chemistry*, *282*(12), 8873–8882. <https://doi.org/10.1074/jbc.M609919200>
- Miteva, M. A., Violas, S., Montes, M., Gomez, D., Tuffery, P., & Villoutreix, B. O. (2006). FAF-Drugs: Free ADME/tox filtering of compound collections. *Nucleic Acids Research*, *34*(WEB. SERV. ISS.). <https://doi.org/10.1093/nar/gkl065>
- Morris, A. L., MacArthur, M. W., Hutchinson, E. G., & Thornton, J. M. (1992). Stereochemical quality of protein structure coordinates. *Proteins: Structure, Function, and Bioinformatics*, *12*(4), 345–364. <https://doi.org/10.1002/prot.340120407>
- Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor

- flexibility. *Journal of Computational Chemistry*, 30(16), 2785–2791.  
<https://doi.org/10.1002/jcc.21256>
- Mourya, D. T., Yadav, P. D., Sandhya, V. K., & Reddy, S. (2013). Spread of Kyasanur Forest disease, Bandipur Tiger Reserve, India, 2012-2013. In *Emerging Infectious Diseases* (Vol. 19, Issue 9, pp. 1540–1541). Emerg Infect Dis. <https://doi.org/10.3201/eid1909.121884>
- Munivenkatappa, A., Sahay, R. R., Yadav, P. D., Viswanathan, R., & Mourya, D. T. (2018). Clinical & epidemiological significance of Kyasanur forest disease. In *Indian Journal of Medical Research* (Vol. 148, Issue 2, pp. 145–150). Indian Council of Medical Research.  
[https://doi.org/10.4103/ijmr.IJMR\\_688\\_17](https://doi.org/10.4103/ijmr.IJMR_688_17)
- Muñoz-Jordán, J. L., Laurent-Rolle, M., Ashour, J., Martínez-Sobrido, L., Ashok, M., Lipkin, W. I., & García-Sastre, A. (2005). Inhibition of Alpha/Beta Interferon Signaling by the NS4B Protein of Flaviviruses. *Journal of Virology*, 79(13), 8004–8013. <https://doi.org/10.1128/jvi.79.13.8004-8013.2005>
- Murhekar, M. V., Kasabi, G. S., Mehendale, S. M., Mourya, D. T., Yadav, P. D., & Tandale, B. V. (2015). On the transmission pattern of Kyasanur Forest disease (KFD) in India. *Infectious Diseases of Poverty*, 4(1), 37. <https://doi.org/10.1186/s40249-015-0066-9>
- Muylaert, I. R., Galler, R., & Rice, C. M. (1997). Genetic analysis of the yellow fever virus NS1 protein: identification of a temperature-sensitive mutation which blocks RNA accumulation. *Journal of Virology*, 71(1), 291–298. <https://doi.org/10.1128/jvi.71.1.291-298.1997>
- Nitsche, C., Onagi, H., Quek, J. P., Otting, G., Luo, D., & Huber, T. (2019). Biocompatible Macrocyclization between Cysteine and 2-Cyanopyridine Generates Stable Peptide Inhibitors. *Organic Letters*, 21(12), 4709–4712. <https://doi.org/10.1021/acs.orglett.9b01545>
- Novak, J., Rimac, H., Kandagalla, S., Grishina, M. A., & Potemkin, V. A. (2021). Can natural products stop the SARS-CoV-2 virus? A docking and molecular dynamics study of a natural product database. *Future Medicinal Chemistry*, 13(4), 363–378. <https://doi.org/10.4155/fmc-2020-0248>
- Patel, B., Singh, V., & Patel, D. (2019). Structural Bioinformatics. In *Essentials of Bioinformatics, Volume I* (pp. 169–199). Springer International Publishing. [https://doi.org/10.1007/978-3-030-02634-9\\_9](https://doi.org/10.1007/978-3-030-02634-9_9)
- Pathak, N., Kuo, Y. P., Chang, T. Y., Huang, C. T., Hung, H. C., Hsu, J. T. A., Yu, G. Y., & Yang, J. M. (2020). Zika Virus NS3 Protease Pharmacophore Anchor Model and Drug Discovery. *Scientific Reports*, 10(1), 1–17. <https://doi.org/10.1038/s41598-020-65489-w>

- Pathak, N., Lai, M. L., Chen, W. Y., Hsieh, B. W., Yu, G. Y., & Yang, J. M. (2017). Pharmacophore anchor models of flaviviral NS3 proteases lead to drug repurposing for DENV infection. *BMC Bioinformatics*, *18*(16), 39–51. <https://doi.org/10.1186/s12859-017-1957-5>
- Pattnaik, P. (2006). Kyasanur forest disease: An epidemiological view in India. In *Reviews in Medical Virology* (Vol. 16, Issue 3, pp. 151–165). John Wiley and Sons Ltd. <https://doi.org/10.1002/rmv.495>
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. <https://doi.org/10.1002/jcc.20084>
- Pushpakom, S., Iorio, F., Eyers, P. A., Escott, K. J., Hopper, S., Wells, A., Doig, A., Williams, T., Latimer, J., McNamee, C., Norris, A., Sanseau, P., Cavalla, D., & Pirmohamed, M. (2018). Drug repurposing: Progress, challenges and recommendations. In *Nature Reviews Drug Discovery* (Vol. 18, Issue 1, pp. 41–58). Nature Publishing Group. <https://doi.org/10.1038/nrd.2018.168>
- Rastelli, G., Del Rio, A., Degliesposti, G., & Sgobba, M. (2010). Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *Journal of Computational Chemistry*, *31*(4), 797–810. <https://doi.org/10.1002/jcc.21372>
- Ray, D., Shah, A., Tilgner, M., Guo, Y., Zhao, Y., Dong, H., Deas, T. S., Zhou, Y., Li, H., & Shi, P.-Y. (2006). West Nile Virus 5'-Cap Structure Is Formed by Sequential Guanine N-7 and Ribose 2'-O Methylations by Nonstructural Protein 5. *Journal of Virology*, *80*(17), 8362–8370. <https://doi.org/10.1128/jvi.00814-06>
- Rimac, H., Grishina, M. A., & Potemkin, V. A. (2020). Electron density analysis of CDK complexes using the AlteQ method. *Future Medicinal Chemistry*, *12*(15), 1387–1397. <https://doi.org/10.4155/fmc-2020-0076>
- Rimac, H., Grishina, M., & Potemkin, V. (2021). Use of the Complementarity Principle in Docking Procedures: A New Approach for Evaluating the Correctness of Binding Poses. *Journal of Chemical Information and Modeling*, *61*(4), 1801–1813. <https://doi.org/10.1021/acs.jcim.0c01382>
- Roe, D. R., & Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *Journal of Chemical Theory and Computation*, *9*(7), 3084–3095. <https://doi.org/10.1021/ct400341p>
- Roosendaal, J., Westaway, E. G., Khromykh, A., & Mackenzie, J. M. (2006). Regulated Cleavages at the West Nile Virus NS4A-2K-NS4B Junctions Play a Major Role in Rearranging Cytoplasmic



- Membranes and Golgi Trafficking of the NS4A Protein. *Journal of Virology*, 80(9), 4623–4632. <https://doi.org/10.1128/jvi.80.9.4623-4632.2006>
- Santos, F. R. S., Nunes, D. A. F., Lima, W. G., Davyt, D., Santos, L. L., Taranto, A. G., & Ferreira, M. S. J. (2020). Identification of Zika Virus NS2B-NS3 Protease Inhibitors by Structure-Based Virtual Screening and Drug Repurposing Approaches. *Journal of Chemical Information and Modeling*, 60(2), 731–737. <https://doi.org/10.1021/acs.jcim.9b00933>
- Shah, S. Z., Jabbar, B., Ahmed, N., Rehman, A., Nasir, H., Nadeem, S., Jabbar, I., ur Rahman, Z., & Azam, S. (2018). Epidemiology, pathogenesis, and control of a tick-borne disease- Kyasanur forest disease: Current status and future directions. In *Frontiers in Cellular and Infection Microbiology* (Vol. 8, Issue MAY). Frontiers Media S.A. <https://doi.org/10.3389/fcimb.2018.00149>
- Shiryaev, S. A., Mesci, P., Pinto, A., Fernandes, I., Sheets, N., Shresta, S., Farhy, C., Huang, C. T., Strongin, A. Y., Muotri, A. R., & Terskikh, A. V. (2017). Repurposing of the anti-malaria drug chloroquine for Zika Virus treatment and prophylaxis. *Scientific Reports*, 7(1), 1–9. <https://doi.org/10.1038/s41598-017-15467-6>
- Sreenivasan, M. A., Bhat, R., & Rajagopalan, P. K. (1986). The epizootics of kyasanur forest disease in wild monkeys during 1964 to 1973. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 80(5), 810–814. [https://doi.org/10.1016/0035-9203\(86\)90390-1](https://doi.org/10.1016/0035-9203(86)90390-1)
- Sun, H., Duan, L., Chen, F., Liu, H., Wang, Z., Pan, P., Zhu, F., Zhang, J. Z. H., & Hou, T. (2018). Assessing the performance of MM/PBSA and MM/GBSA methods. 7. Entropy effects on the performance of end-point binding free energy calculation approaches. *Physical Chemistry Chemical Physics*, 20(21), 14450–14460. <https://doi.org/10.1039/c7cp07623a>
- Tomar, S., Mudgal, R., & Fatma, B. (2017). Flavivirus protease: An antiviral target. In *Viral Proteases and Their Inhibitors* (pp. 137–161). Elsevier. <https://doi.org/10.1016/B978-0-12-809712-0.00006-X>
- Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2), 455–461. <https://doi.org/10.1002/jcc.21334>
- Umamaheswari, A., Pradhan, D., & Hemanthkumar, M. (2010). Virtual screening for potential inhibitors of homology modeled *Leptospira interrogans* MurD ligase. *Journal of Chemical Biology*, 3(4), 175–187. <https://doi.org/10.1007/s12154-010-0040-8>
- Umareddy, I., Chao, A., Sampath, A., Gu, F., & Vasudevan, S. G. (2006). Dengue virus NS4B interacts with NS3 and dissociates it from single-stranded RNA. *Journal of General Virology*,

87(9), 2605–2614. <https://doi.org/10.1099/vir.0.81844-0>

- Upadhyaya, S., Murthy, D. P. N., & Anderson, C. R. (1975). Kyasanur Forest disease in the human population of Shimoga District, Mysore State, 1959-1966. *Indian Journal of Medical Research*, 63(11), 1556–1563. <http://europepmc.org/article/med/1222964>
- Upadhyaya, S., Murthy, D. P. N., & Murthy, B. K. Y. (1975). Viraemia studies on the Kyasanur Forest disease human cases of 1966. *Indian Journal of Medical Research*, 63(7), 950–953. <https://europepmc.org/article/med/175006>
- Van Santen, J. A., Jacob, G., Singh, A. L., Aniebok, V., Balunas, M. J., Bunsko, D., Neto, F. C., Castaño-Espriu, L., Chang, C., Clark, T. N., Cleary Little, J. L., Delgadillo, D. A., Dorrestein, P. C., Duncan, K. R., Egan, J. M., Galey, M. M., Haeckl, F. P. J., Hua, A., Hughes, A. H., ... Linington, R. G. (2019). The Natural Products Atlas: An Open Access Knowledge Base for Microbial Natural Products Discovery. *ACS Central Science*, 5(11), 1824–1833. <https://doi.org/10.1021/acscentsci.9b00806>
- Vitkup, D., Melamud, E., Moulton, J., & Sander, C. (2001). Completeness in structural genomics. *Nature Structural Biology*, 8(6), 559–565. <https://doi.org/10.1038/88640>
- Volkamer, A., Kuhn, D., Grombacher, T., Rippmann, F., & Rarey, M. (2012). Combining Global and Local Measures for Structure-Based Druggability Predictions. *Journal of Chemical Information and Modeling*, 52(2), 360–372. <https://doi.org/10.1021/ci200454v>
- Vyas, V. K., Ukawala, R. D., Ghate, M., & Chintha, C. (2012). Homology modeling a fast tool for drug discovery: Current perspectives. In *Indian Journal of Pharmaceutical Sciences* (Vol. 74, Issue 1, pp. 1–17). Wolters Kluwer -- Medknow Publications. <https://doi.org/10.4103/0250-474X.102537>
- Wang, J., Wang, W., Kollman, P. A., & Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling*, 25(2), 247–260. <https://doi.org/10.1016/j.jmgm.2005.12.005>
- Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and testing of a general Amber force field. *Journal of Computational Chemistry*, 25(9), 1157–1174. <https://doi.org/10.1002/jcc.20035>
- Wang, L., Ma, C., Wipf, P., Liu, H., Su, W., & Xie, X. Q. (2013). Targethunter: An in silico target identification tool for predicting therapeutic potential of small organic molecules based on chemogenomic database. *AAPS Journal*, 15(2), 395–406. <https://doi.org/10.1208/s12248-012-9449-z>

- Warrener, P., Tamura, J. K., & Collett, M. S. (1993). RNA-stimulated NTPase activity associated with yellow fever virus NS3 protein expressed in bacteria. *Journal of Virology*, *67*(2), 989–996. <https://doi.org/10.1128/jvi.67.2.989-996.1993>
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., De Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, *46*(W1), W296–W303. <https://doi.org/10.1093/nar/gky427>
- Wengler, G., & Wengler, G. (1991). The carboxy-terminal part of the NS 3 protein of the West Nile Flavivirus can be isolated as a soluble protein after proteolytic cleavage and represents an RNA-stimulated NTPase. *Virology*, *184*(2), 707–715. [https://doi.org/10.1016/0042-6822\(91\)90440-M](https://doi.org/10.1016/0042-6822(91)90440-M)
- Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, *35*(SUPPL.2). <https://doi.org/10.1093/nar/gkm290>
- WORK, T. H., RODERIGUEZ, F. R., & BHATT, P. N. (1959). Virological epidemiology of the 1958 epidemic of Kyasanur Forest disease. *American Journal of Public Health*, *49*(7), 869–874. <https://doi.org/10.2105/ajph.49.7.869>
- WORK, T. H., TRAPIDO, H., MURTHY, D. P., RAO, R. L., BHATT, P. N., & KULKARNI, K. G. (1957). Kyasanur forest disease. III. A preliminary report on the nature of the infection and clinical manifestations in human beings. *Indian Journal of Medical Sciences*, *11*(8), 619–645. <https://europepmc.org/article/med/13474777>
- Ya'u Ibrahim, Z., Uzairu, A., Shallangwa, G., & Abechi, S. (2020). Molecular docking studies, drug-likeness and in-silico ADMET prediction of some novel  $\beta$ -Amino alcohol grafted 1,4,5-trisubstituted 1,2,3-triazoles derivatives as elevators of p53 protein levels. *Scientific African*, *10*, e00570. <https://doi.org/10.1016/j.sciaf.2020.e00570>
- Yadav, P. D., Patil, S., Jadhav, S. M., Nyayanit, D. A., Kumar, V., Jain, S., Sampath, J., Mourya, D. T., & Cherian, S. S. (2020). Phylogeography of Kyasanur Forest Disease virus in India (1957–2017) reveals evolution and spread in the Western Ghats region. *Scientific Reports*, *10*(1), 1–12. <https://doi.org/10.1038/s41598-020-58242-w>