Identification of dengue viral RNA-dependent RNA polymerase inhibitor using computational fragment-based approaches and molecular dynamics study

Shanmugam Anusuya**, Devadasan Velmurugan** and M. Michael Gromiha**

*Department of Biotechnology, Indian Institute of Technology Madras, Chennai 600036, Tamil Nadu, India; **Centre of Advanced Study in Crystallography and Biophysics, University of Madras, Guindy Campus, Chennai 600 025, India; ***Bioinformatics Infrastructure Facility (BIF), University of Madras, Guindy Campus, Chennai 600 025, India

Communicated by Ramaswamy H. Sarma

(Received 1 July 2015; accepted 6 August 2015)

Dengue is a major public health concern in tropical and subtropical countries of the world. There are no specific drugs available to treat dengue. Even though several candidates targeted both viral and host proteins to overcome dengue infection, they have not yet entered into the later stages of clinical trials. In order to design a drug for dengue fever, newly emerged fragment-based drug designing technique was applied. RNA-dependent RNA polymerase, which is essential for dengue viral replication is chosen as a drug target for dengue drug discovery. A cascade of methods, fragment screening, fragment growing, and fragment linking revealed the compound [2-(4-carbamoylpiperidin-1-yl)-2-oxoethyl]-8-(1,3-benzothiazol-2-yl)naphthalene-1-carboxylate as a potent dengue viral polymerase inhibitor. Both strain energy and binding free energy calculations predicted that this could be a better inhibitor than the existing ones. Molecular dynamics simulation studies showed that the dengue polymerase–lead complex is stable and their interactions are consistent throughout the simulation. The hydrogen-bonded interactions formed by the residues Arg792, Thr794, Ser796, and Asn405 are the primary contributors for the stability and the rigidity of the polymerase–lead complex. This might keep the polymerase in closed conformation and thus inhibits viral replication. Hence, this might be a promising lead molecule for dengue drug designing. Further optimization of this lead molecule would result in a potent drug for dengue.

Keywords: dengue; RNA-dependent RNA polymerase; fragment-based drug designing; fragment growing; fragment linking; lead molecule; substructure search; neighborhood search; fragment library; molecular dynamics simulation

1. Introduction

Dengue is an important mosquito-borne viral infection found in tropical and subtropical countries of the world. In recent decades, the incidence of dengue has grown dramatically around the globe. Today, over 2.5 billion people are at risk from dengue, (Whitehorn & Farrar, 2010). Four closely related but antigenically distinct serotypes of dengue virus (DV 1-4) causes dengue infection (Weaver & Vasilakis, 2009). DV belongs to the Flavivirus genus and Flaviviridae family. Humans are the most susceptible hosts of dengue viruses. DV is transmitted to human via the infected mosquitoes of the genus Aedes, especially by Aedes aegypti (Vitek, Gutierrez, & Dirrgl, 2014).

The primary infection by any one of the dengue viral serotype results in mild febrile illness known as dengue fever, whereas the subsequent infections by different serotypes causes the severe form of the disease known as dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS). DHF and DSS are the leading causes of hospitalization and death (Gubler, 2006). The World Health Organization recently estimated 50–100 million dengue infections and 20,000 deaths due to DHF and DSS worldwide every year (Gubler, 2006). No specific antiviral medications are currently available to treat dengue, and the current treatment is only supportive care (Tsai et al., 2013). Fluid replacement therapy is suggested for DHF and DSS (Hung, 2012). There are no licensed vaccines for dengue; however, few are under clinical trials (Dayan et al., 2014; Osorio et al., 2014; Watanaveeradej et al., 2014; Wei, Chen, & An, 2014). So far, numerous candidates have been reported to target the viral proteins (Lim et al., 2011, 2013; Niyomrattanakit et al., 2010; Tomlinson et al., 2009; Velmurugan, Mythily, & Rao, 2014; Velmurugan et al., 2012; Wang et al., 2009; Xie et al., 2011; Yin et al., 2006). But none of them have entered into the later stages of clinical trials. This enforces the need for the development of safe and effective anti-viral therapy for dengue.

DV has an enveloped single-stranded positive-sense RNA genome, which is approximately 11 KB in size. This encodes a single polyprotein that is subsequently cleaved by viral and host proteases into three structural (C-PrM-E) and seven non-structural proteins.
Non-structural proteins participate in genome replication and virion assembly (Back & Lundkvist, 2013).

The non-structural protein NS5 is a multifunctional enzyme, consists of 900 amino acids. The N-terminal methyltransferase (MTase) and C-terminal RNA-dependent RNA polymerase (RdRp) domains of NS5 are involved in RNA capping and genome replication, respectively (Potisopon et al., 2015). Replication is a process which involves the generation of negative strand RNA, which acts as a template for the synthesis of multiple strands of positive strand viral RNA genome. RdRp is involved in the generation of both minus and plus-strand RNAs. RdRp is the most conserved protein among the members of the Flavivirus genus (Bollati et al., 2010). In dengue, more than 65% of the amino acid sequence is conserved across the four different dengue viral serotypes (Yap et al., 2007). As humans do not have RdRp and its homologous enzymes, the antiviral agents that target RdRp would not produce any toxic effects in human host. Hence, RdRp was considered as a promising target for antiviral drug development.

Several candidates have already been reported as RdRp inhibitors. These RdRp inhibitors are broadly divided into nucleoside/nucleotide inhibitors (NI) and non-nucleoside inhibitors (NNI) (Nguyen et al., 2013; Niyomrattanakit et al., 2010; Noble et al., 2013; Yin, Chen, Kondreddi, et al., 2009, Yin, Chen, Schul, et al., 2009). NI mimic rNTP substrates and bind competitively in the active site of an enzyme. When these NI are incorporated into growing RNA chain, it results in chain termination. Two such nucleoside analogs, NITD-008 and Balapiravir, which target DV RdRp, successfully entered into preclinical and clinical trials. However, due to the toxic effects of NITD-008 and lack of potency of Balapiravir, both were terminated (Nguyen et al., 2013; Yin, Chen, Kondreddi, et al., 2009, Yin, Chen, Schul, et al., 2009).

NNI bind non-competitively to allosteric pockets located outside the active site of an enzyme. This changes the conformation and by which impairs the enzymatic activity of the protein [27]. NITD-2 and NITD107 were identified as the NNI of DV RdRp (Niyomrattanakit et al., 2010; Noble et al., 2013). The 3D structure of DV RdRp resembles a right hand with a characteristic finger, palm, and thumb subdomains. DV RdRp has two different conformations, open and closed. During initiation of RNA polymerization, the polymerase switches from a closed conformation to an open conformation to accommodate the viral RNA template. Both the NNIs selectively bind with DV RdRp domain and lock the polymerase in closed conformation, and thus, prevent viral RNA synthesis. However, they exhibit only weak inhibitory and binding activities (Noble et al., 2013).

In addition to targeting viral proteins, antiviral research also focuses on the development of host protein inhibitors to overcome dengue infection (Low et al., 2014; Wang, Bushell, et al., 2011, Wang, Kondreddi, et al., 2011). NITD-451 targets a host factor involved in translation (Wang, Kondreddi, et al., 2011). NITD-982 inhibits pyrimidine biosynthesis through dihydroorotate dehydrogenase (DHODH) (Wang, Bushell, et al., 2011). Celgosivir is a host alpha-glucosidase inhibitor, and in clinical studies, it was found to be a safe and well-tolerated drug. However, it did not show any antiviral activity (Low et al., 2014). Even though NI and host enzyme inhibitors of DV RdRp successfully entered into clinical studies, they failed to establish themselves as a potent antiviral therapeutics for dengue. This enforces the need for designing a better therapeutics for dengue.

Fragment-based drug design (FBDD) has emerged as an efficient method in drug discovery and design. This method identifies small molecular fragments that make high-quality interactions with the drug target. These fragments are developed into lead molecules with high affinity and selectivity. The availability of suitable high-throughput screening approaches, sensitive bio-physical techniques such as nuclear magnetic resonance and X-ray crystallography and best optimization strategies makes FBDD a powerful tool for finding good quality hits for a wide range of drug targets (Kumar, Voet, & Jiang, 2012). Several lead molecules developed by FBDD have already progressed into later stages in clinical development (Turnbull & Boyd, 2012). An anticancer drug Zelboraf (Vemurafenib) approved by FDA is a milestone in the success history of FBDD (Bollag et al., 2012). In this study, we applied computational FBDD method and designed a potent non-nucleoside inhibitor for DV RdRp enzyme.

2. Materials and methods

All the molecular modeling calculations such as protein preparation, ligand preparation, docking, and ADME/Toxicity predictions were carried out using the Schrödinger software suite. For molecular visualization and to produce high-quality images, PyMOL v1.3 was used. The hydrogen bonds, π–π interactions, cation–π interactions, and hydrophobic interactions were analyzed using Maestro v9.9.

2.1. Protein preparation

The 3D structure of the target protein, DV RdRp, was retrieved from the Protein Data Bank (PDB ID: 3VWS) (Noble et al., 2013) into the Maestro workspace. The 3D structures downloaded from PDB could not be used as
such in Schrodinger. So, it was prepared using protein preparation wizard (Sastry, Adzhigirey, Day, Annabhimouju, & Sherman, 2013). The protein preparation consists of preprocessing, analyzing, modifying, and refining the structure.

During preprocessing, the bond orders were assigned. Metals were assigned zero-order bonds. Hydrogen atoms were added. The missing atoms of Gln351 and Val480 were added using Prime. Similarly, the missing loops (residues 314–317, 344–350 and 456–464) in the active site region, were filled and refined using Prime. However, the loops (residues 266–271 and 884–900) far from the active site were ignored.

The water molecules which were beyond 5 Å from the ligands were removed. The structural analysis showed that the residues Arg382, Glu648, Glu834, Arg842, and Asp881 have alternate positions. The current average occupancy values of these residues were predicted as .5. However, their alternate positions had better occupancy values such as .57, .61, .61, .57, and .64, respectively. Hence, these residues were changed to their alternate positions.

The possible ionization and tautomeric states of the hetero groups were predicted using Epik v2.9 at the pH range of 7 ± 3. Two possible ionization states were predicted for hetero group (inhibitor NITD-107). As the original state had the lowest penalty (0.53 kcal/mol) than the other (1.53 kcal/mol), the original state was retained in the 3D structure.

The hydrogen bonds were optimized to avoid steric clashes. The water molecules which form less than 3 hydrogen bonds were removed. Finally, the entire 3D structure was refined with restrained minimization using Impact Refinement (impref) module and the force field OPLS-2005 with the convergence criteria of .3 Å RMSD.

### 2.2. Selection of fragment library

The Enamine Golden Fragment Library (EGFL) consists of 1300 fragments (as on September 2013) (http://www.enamine.net). These fragments adhere to Astex’s Rule of Three (molecular weight ≤ 300 Da; number of hydrogen bond donor ≤ 3; number of hydrogen bond acceptor ≤3; ClogP ≤ 3). They are highly diverse in nature and the diversity coefficient is .916. Besides, EGFL excludes toxicophores and reactive groups. This makes GFL suitable for FBDD and was downloaded from Enamine database in structure data file format (sdf).

### 2.3. Ligand preparation

The selected fragments were imported into the Maestro workspace and were prepared using LigPrep v3.1 (Foudah, Sallam, Akl, & El Sayed, 2014). The ligands were desalted. The structures were converted from 2D into 3D. The hydrogen atoms were added. All possible ionization states and tautomeric states of each of the ligands were prepared at pH range of 7 ± 2 using Epik v2.9 (Shelley et al., 2007) and were verified using Henderson–Hasselbalch Equation. However, the chirality of the ligands was not altered. The ligands were energy minimized using force field OPLS 2005. This ligand preparation process resulted in one low-energy conformation for each of the ligand. The resultant structures were saved in maestro file format.

### 2.4. Grid generation

The receptor grid generation panel in Glide v6.4 (Kawatkar, Wang, Czerminski, & Joseph-McCarthy, 2009) was used for grid box generation. The grid was generated on the active site of the protein. The inhibitor NITD-107 bound active site was chosen for placing the grid box. The center of the grid box was set in the center of the workspace ligand NITD-107. To soften the potential for non-polar parts of the receptor, the van der Waals radius scaling factor and partial charge cutoff were set to 1 and .25, respectively. The bound inhibitor NITD-107 was excluded from the grid and the grid box was generated.

### 2.5. Docking

The prepared EGFL fragments were screened against DV RdRp using Glide v6.4 (Grid-based Ligand Docking with Energetics) (Kawatkar et al., 2009) in standard precision (SP). The prepared fragments were placed in different locations in the grid box with possible conformations. As the ligands in this study are small molecular weight fragments, to widen the docking funnel, the initial number of poses per ligand were increased to 50,000; the scoring window was widen to 500; the number of minimized poses per ligand were increased to 1000; and expanded sampling was applied.

During Glide SP docking, the receptor was set as rigid molecule and the ligand as flexible molecule. The van der Waals radii of non-polar atoms were calculated to decrease close contact penalties between ligand and active site residues. No constraints were applied. The ligands were allowed to bind in more than one meaningful conformation within the grid box. The resultant hit molecules were selected based on the docking score.

### 2.6. Fragment linking

The fragment hits (FH1-5) which showed the highest binding affinity in Glide SP docking were combined with other fragments which were proximal to them in the active site. Fragments were combined either through direct joining or linking. In direct joining, fragments were linked directly without linkers, whereas in fragment
linking, linkers were used. Schrodinger’s combine
fragments script was used for linking fragments.

2.7. Substructure search
The substructures were searched for the fragment hits
FH1-5 from the PubChem (http://pubchem.ncbi.nlm.nih.
gov) and ChemSpider databases (www.chemspider.com/).
The input was given in simplified molecular-input line-
entry system (SMILES) format to enhance the fast search
against the databases. Compounds that satisfy Lipinski’s
rule of five (molecular weight ≤500 Da; hydrogen bond
acceptor ≤10, hydrogen bond donor ≤5; log P ≤ 5) were
selected and saved in sdf.

2.8. Neighborhood search
The top 5 hits from substructure search (SSH1-5) were
searched for similar structures from the PubChem data-
base. The structural information of the input molecule
was given in SMILES file format. Similarity cutoff was
set to 90%. To obtain only the drug-like compounds, the
Lipinski rule of five filters was applied. The resultant
similar compounds were collected and saved in sdf.

2.9. Fragment growing
The fragment hits FH1-5 obtained in Glide SP docking
were chosen as the initial core scaffold. Hydrogen atoms
were added with these fragment hits and were grown
using AutoGrow v2.0.4 (Durrant, Amaro, & McCammon,
2009). Fragment hits were grown in number of genera-
tions. Each generation includes 3 operators: selection,
mutation, and crossover. The top 10 ligands with the best
binding affinity values were selected from the previous
generation. In each of the ligand, an appropriate hydrogen
atom was selected randomly, and replaced with the frag-
ments from small fragment library. Twenty such substit-
tuted ligands, which were similar but distinct from other
members of the population, were created. In each genera-
tion, except the first, pair of ligands was selected and a
new hybrid ligand was formed by random mixing and
matching of the attached moieties. Twenty such child/hy-
brid ligands were derived in each generation. All these 50
ligands were docked with the target protein DV RdRp
using Glide v6.4 in SP. Maximum number of generations
was set to 5 and maximum number of atoms in the grow-
ing fragment was set to 500. The indices of scaffold
hydrogen atoms to which AutoGrow will not add frag-
ments were set as 30 for FH1, and 38 and 39 for FH5.

2.10. ADME/toxicity prediction
The 3D structure of lead-like molecules obtained in frag-
ment growing and linking processes were neutralized
and energy minimized. The pharmacokinetic properties
such as absorption, distribution, metabolism, excretion,
and toxicity were predicted using QikProp v4.1
(Jorgensen & Duffy, 2002). This predicts both physically
significant descriptors and pharmacologically relevant
properties quickly and reliably.

2.11. Molecular dynamics simulations (MDS)
The molecular dynamics simulations (MDS) of DV
RdRp in apo form and in complex with the lead mole-
cule and the reference ligand was carried out using
Desmond/Maestro version 2014.2 with optimized poten-
tial for liquid simulations-all atoms (OPLS-AA) (2005)
force field (Liu et al., 2013). The system for MD
simulation was built using System Builder module in
Maestro. The protein–ligand complex was placed in a
10 Å orthorhombic box solvated with simple point
charge water molecules. The system was neutralized
with 3 Cl– ions and periodic boundary conditions were
applied in each direction. The Na+ and Cl– ions were
added in a concentration of .15 M. Electrostatic interac-
tions were computed using particle mesh Ewald method
and a cutoff of 10 Å was set for Lennard-Jones interac-
tion. The SHAKE algorithm was used to restrict the
motions of all covalent bonds involving hydrogen
atoms.

Before MD simulations were performed, the entire
system was relaxed using a series of minimization and
short MD simulations. The solvated system was mini-
imized initially with solute non-hydrogen atoms
restrained and then further it was minimized using a
hybrid method of steepest decent and the limited-
memory Broyden–Fletcher–Goldfarb–Shanno algorithm
with maximum iteration of 2000 which includes the
initial 10 iterations of steepest descent. In this minimiza-
tion stage, no restraints were applied on the solute atoms.
Once the system is minimized, the system was simulated
with restraint on the solute non-hydrogen atoms in NVT
ensemble at 10 K for a short period of 12 ps; and in
NPT ensemble at 10 K for 12 ps; and in NPT ensemble
at 300 K for 12 ps. Finally without any restraints, the
system was simulated in NPT ensemble at 300 K for
24 ps. During these short periods of simulations, the
temperature and pressure were controlled using Berend-
sen thermostats and barostats, respectively.

After the system was relaxed, it was simulated in
NPT ensemble at a temperature of 300 K and
1.01325 bar pressure for 10 ns with a time step of 2 fs.
Nose–Hoover thermostat and Martyna–Tobias–Klein
barostat were used to control the temperature and pres-
sure. Atomic coordinate data and system energies were
recorded for every 9.6 ps. Using trajectories, the stability
of the protein–ligand complex and the consistency of the
hydrogen bonding patterns were analyzed.
3. Results and discussion

3.1. Fragment screening

The 1300 fragments in EGFL were prepared using LigPrep v3.1 and were screened against DV RdRp enzyme using Glide SP docking. Based on the docking score, the top 10% of the fragments were selected for detailed analysis. The orientation and the binding poses of these 130 fragments reveal three different regions in the active site that has high propensity for ligand binding. Based on these regions where the fragments bind, the fragments are clustered into three groups: G1, G2, and G3, which comprises 6, 8, and 48 fragments, respectively. However, the remaining 68 fragments occupy the region where both G2 and G3 fragments bind in the active site. These 68 fragments are clustered into group G4. The Enamine ID of the fragments cluster in each group is shown in Table S1. The three different regions in the active site preferred by EGFL fragments are shown in Supplementary Figure S1.

The G1 and G2 fragments are mainly stabilized by hydrogen bonds. In G1, Asp663 plays a crucial role in making hydrogen bonds, whereas in G2, Tyr606 is found to be important. In G3 and G4, the primary stabilization factor for protein–ligand complexes is cation–π interactions. Lys401 is the only residue that made this significant interaction with EGFL.

The G1 fragments interact with drug target mainly through hydrogen bonds. The side chain carboxylic acid in Asp663 forms hydrogen bonds with phenolic or carboxylic group or with secondary amines present in the fragments. Asp663 is a component of GDD motif or motif C and plays a primary role in divalent metal ion coordination of the enzyme. In few of the G1 fragments, π–π interactions are also observed with Tyr606 and Hid798. However, cation–π interactions are not formed by G1 fragments.

In category G2, the protein–ligand complexes are stabilized only through hydrogen bonds. Tyr606 plays an essential role in making these interactions. The amine group in Tyr606 forms hydrogen bonds with either the carboxyl oxygen or heterocyclic nitrogen atoms in the fragments. None of the G2 fragments established π–π and cation–π interactions.

The main stabilization factor for the G3 and G4 fragments is cation–π interactions. The cationic side chain amine present in Lys401 is highly reactive and forms this interaction with the π electrons of the aromatic groups in the fragments. In addition, π–π interactions and hydrogen bonds are also observed with G3 and G4 fragments. Phe398, Phe485, and Phe412 play a significant role in making π–π interactions. Phe412 is an essential residue, which regulates the access of ssRNA substrate. In G3 fragments, the residues involve in forming hydrogen bonds were Asn405 and Val411. But in G4, this residue is Tyr606. Asn405 is a residue in nuclear localization sequence (α/β-NLS) and it is responsible for binding cellular factor α/β-importin. Lys401 and Asn405 play an important role in establishing interactions with the previously reported DV RdRp inhibitor NITD-107. It is notable that the top 25 fragments were from category G3 and G4. This shows the importance of Lys401 and the formation of cation–π interactions with the drug target.

3.2. Fragment hit identification

Among 130 fragments, a G3 fragment 5-(6,7-dihydro-4H-thieno[3,2-c]pyridine-5-carbonyl)thiophene-2-carboxylic (Enamine ID Z271743752) which shows the highest binding affinity with DV RdRp, is found to be the top most fragment hit in the overall screening process. This fragment hit (FH1) has a molecular mass of 293 Da, partition coefficient (logP) of .144 and polar surface area (PSA) of 57.61 Å². It has 3 hydrogen bond acceptors, one hydrogen bond donor, and two rotatable bonds. FH1 has thiophene and thienopyridine rings in its structure. The side chain carboxylic group in the thiophene ring form hydrogen bonds with Thr413 while the π electrons in thienopyridine ring make cation–π interaction with Lys401 as well as π–π interaction with Phe412. FH1 form hydrophobic interactions with the residues Phe398, Val402, Val411, Trp477, Phe485, and Val603. FH1 shows the highest docking score of 7.661 and ligand efficiency of .403. The 2D structure of the top fragment hit (FH1) and its interactions with the drug target DV RdRp are shown in Figures 1 and 2.

3.3. Generation of lead compounds

As fragments are small molecular weight (molecular weight ≤300 Da) compounds, they usually form fewer interactions with the drug targets. Hence, they can be developed into a larger lead molecule (molecular

![FH1](image-url)
weight ≤ 500 Da) which can form much more strong interactions with the drug target. Fragment growing, fragment merging and fragment linking are the three different methods by which lead can be generated. In this work, fragment linking and fragment growing methods were applied. To increase the possibilities of getting the best lead molecule, the top 5 fragment hits were selected for lead generation. The fragment hits chosen for lead generation and the interactions they made with the drug target DV RdRp are provided in Table 1.

3.3.1. Fragment linking

Fragment linking is one of the efficient ways to develop fragments into lead molecule. In fragment linking, two or more fragments are linked together either directly or using linkers. However, the fragments which are going to be linked together should be closer to each other in the active site. As fragment linking process required minimum of two fragments, one of them was restricted as fragment hit. Other fragments which were closer to it in the active site were chosen for linking.

The fragment hit 1 (FH1) was linked with 8 such proximal fragments, whereas the hit 2 (FH2) with 1, hit 3 (FH3) with 5, and hits 4 (FH4) and 5 (FH5) with 6 proximal fragments. Thus, the fragment linking process produced 26 linked fragments (LF).

In 4 of the fragment linking process (linking of FH1 and FH67; FH3 and FH26; FH3 and FH55; and FH4 and FH42), the fragments were directly linked together. No linkers were involved in these fragment linking processes. But the rest of the fragment linking processes either ethylene or methylene linkers were used. The fragments which were possibly linked with each fragment hits (FH1-5) and the linkers used are shown in Table 2.

All the linked fragments were screened against the target protein DV RdRp using Glide v6.4 in SP. The linked fragments which showed the highest binding affinity for each of the fragment hits and the interactions they made with the target protein DV RdRp are shown in Table 3. The docking study clearly shows that each of the linked fragments maintain the interactions that their parent fragments made with the drug target. Similarly, most of the linked fragments also maintain the binding poses and orientations of their parent fragments. However, only in 5 linked fragments, the binding affinity increased than their corresponding initial fragments. These correspond to LF1.67 (FH1 linked with hit 67), LF1.127 (FH1 linked with 127), LF2.26 (FH2 with 26), LF4.55 (FH4 with 55), and LF5.104 (FH5 with 104).

Table 1. Fragment hits chosen for lead generation and their interactions with DV RdRp.

| Fragment hit | Enamine ID   | DS   | LE | Hbond  | π–π          | cation–π | Hydrophobic                  |
|--------------|--------------|------|----|--------|--------------|----------|---------------------------- |
| FH1          | Z271743752   | −7.66| .403| Thr413, Gly601 | Phe412 | Lys401 | Phe398, Lys401, Val402, Val411, Trp477, Phe485, Gly601, Val603 |
| FH2          | Z431763650   | −7.35| .387| Tyr606  | Phe398 |        | Phe398, Lys401, Val402, Asn405, Val411, Phe412, Asn492, Gly602, Val603, Gly604, Thr605 |
| FH3          | Z1139336709  | −7.25| .346| Tyr606  | Phe485 | Lys401 | Phe398, Lys401, Val402, Val411, Phe412, Phe485, Asn492, Val603, Gly604, Thr605 |
| FH4          | Z104341206   | −7.21| .380| Tyr606  | Phe398 |        | Phe398, Lys401, Val411, Phe412, Phe485, Val603, Gly602 |
| FH5          | Z650353014   | −7.20| .360| Gly604  | Phe485, Phe398 | Lys401 | Phe398, Phe412, Leu488, Gly489, Asn492, Val603, Phe485, Lys401 |

Notes: DS – Docking score; LE – Ligand efficiency.
Among the 26 linked fragments, LF2.26 (fragment hit 2 linked with hit 26) shows the highest binding affinity (DS = -7.457) in the fragment linking processes. However, it has ligand efficiency of only 0.226. The LF2.26 retains the binding poses and the orientations of the initial fragments FH2 and FH26. The LF2.26 superimposed very well with the FH26 and the RMSD of 0.0016 is found, and LF2.26 not exactly overlapped with the FH2. However, LF2.26 maintained the same interactions as FH2. An RMSD of 1.5680 is observed between FH2 and LF2.26. The FH2 and FH26 which were superposed with LF2.26 are shown in Figure 3.

The FH2 (Enamine ID Z431763650) has thienopyridine and pyridine rings in its structure. The π electrons in the thienopyridine ring form cation-π interaction with Lys401. Similarly, the oxygen atom of the pyridine-2-one form hydrogen bond with Tyr606. These interactions are also observed in the linked fragment LF2.26.

FH26 (Enamine ID Z57069342) is a quinolone derivative. The hydroxyl and the secondary amine present in FH26 form two hydrogen bonds with the side chain carboxylic acid in Asp663. Similarly, the oxygen atom of quinolone-2-one forms a hydrogen bond with side chain amide in Asn609. These essential interactions are also maintained by the linked fragment LF2.26. Likewise, the hydrophobic interactions of FH2 with the residues Val411, Phe412, Gly604, and Thr605, and FH26 with Ser661 and Asp664 are also identified in the linked fragment LF2.26. The structure of the FH2, FH26, and LF2.26 are shown in Figure 4 and its interactions with the drug target DV RdRp are shown in Figures 5–7.

### Fragment growing

One way to develop a fragment into a lead molecule is fragment growing. This can be done in two different ways: explicitly growing fragments and implicitly growing fragments. In explicit fragment growing method, either atoms or groups of atoms are added with the existing fragment hit in a manner that the newly grown fragment can make additional favorable interactions with the drug target. In implicit fragment growing method, the fragment hit is grown by selecting existing compounds that include the structure of the fragment hit as its substructure and finding the best one from the selection using a more efficient high-throughput screen. This can be done using substructure and/or similarity search.

---

| Fragment hit | Enamine ID | Fragment hit | Enamine ID | Linked fragment | Name of the linker |
|--------------|------------|--------------|------------|-----------------|-------------------|
| FH1          | Z271743752 | FH26         | Z57069342  | LF1.26          | Ethylene          |
| FH1          | Z271743752 | FH42         | Z875025832 | LF1.42          | Ethylene          |
| FH1          | Z271743752 | FH55         | Z56796061  | LF1.55          | Ethylene          |
| FH1          | Z271743752 | FH67         | Z15842101  | LF1.67          | –                 |
| FH1          | Z271743752 | FH104        | Z1171979415| LF1.104         | Ethylene          |
| FH1          | Z271743752 | FH116        | Z446011260 | LF1.116         | Ethylene          |
| FH1          | Z271743752 | FH120        | Z114743980 | LF1.120         | Ethylene          |
| FH1          | Z271743752 | FH127        | Z1190031301| LF1.127         | Methylene         |
| FH2          | Z431763650 | FH26         | Z57069342  | LF2.26          | Ethylene          |
| FH3          | Z1139336709| FH26         | Z57069342  | LF3.26          | –                 |
| FH3          | Z1139336709| FH26         | Z57069342  | LF3.26          | –                 |
| FH3          | Z1139336709| FH26         | Z57069342  | LF3.26          | –                 |
| FH3          | Z1139336709| FH26         | Z57069342  | LF3.26          | –                 |
| FH4          | Z104341206 | FH26         | Z57069342  | LF4.26          | Methylene         |
| FH4          | Z104341206 | FH42         | Z875025832 | LF4.42          | –                 |
| FH4          | Z104341206 | FH55         | Z56796061  | LF4.55          | Ethylene          |
| FH4          | Z104341206 | FH104        | Z1171979415| LF4.104         | Methylene         |
| FH4          | Z104341206 | FH116        | Z446011260 | LF4.116         | Ethylene          |
| FH5          | Z650353014 | FH26         | Z57069342  | LF5.26          | Methylene         |
| FH5          | Z650353014 | FH42         | Z875025832 | LF5.42          | Methylene         |
| FH5          | Z650353014 | FH55         | Z56796061  | LF5.55          | Methylene         |
| FH5          | Z650353014 | FH104        | Z1171979415| LF5.104         | Ethylene          |
| FH5          | Z650353014 | FH116        | Z446011260 | LF5.116         | Methylene         |
| FH5          | Z650353014 | FH120        | Z114743980 | LF5.120         | Ethylene          |

*Fragments linked directly without any linkers.
†Substituted atom was F.
§Substituted atom was Br.
Table 3. Hits from fragment linking processes and their interactions with DV RdRp.

| Hit | FH1 | FH7 | LF1.67 |
|-----|-----|-----|--------|
| Enamine ID | Z217143752 | Z15842101 | |
| DS | −7.661 | −6.324 | −7.107 |
| LE | .403 | .486 | .222 |
| Hbond | Thr413, Gly601 | Try606, | | |
| π–π | Phe412 | Gly604 | | |
| Cation–π | Lys401 | Lys401 | | |
| Hydrophobic | Phe398, Lys401, Val402, Val411, Trp477, Phe485, Gly601, Val603 | Phe398, Lys401, Asp485, Gly489, Asn492, Gly493, Val603 | | |
| Hit | FH2 | FH26 | |
| Enamine ID | Z431763650 | | |
| DS | −7.351 | −6.969 | −7.457 |
| LE | .387 | .558 | .226 |
| Hbond | Tyr606 | Asp663, Asn609 | | |
| π–π | Phe398 | | | |
| Cation–π | Lys401 | | | |
| Hydrophobic | Phe398, Lys401, Val402, Asn405, Val411, Phe412, Asn492, Gly602, Val603, Gly604, Thr605 | Ser661, Asp664, Tyr606, Val411, Phe412, Gly604, Thr605 | | |
| Hit | FH3 | FH116 | |
| Enamine ID | Z1139336709 | | |
| DS | −7.258 | −6.127 | −5.564 |
| LE | .346 | .438 | .155 |
| Hbond | Tyr606 | Asp663 | | |
| π–π | Phe485 | | | |
| Cation–π | Lys401 | | | |
| Hydrophobic | Phe398, Lys401, Val402, Val411, Phe412, Phe485, Asn492, Val603, Gly604, Thr605 | Ser661, Gly662, Ile797, Val411, Phe412, Trp477, Phe485, Glu493, Glu507, Val603 | | |
| Hit | FH4 | FH55 | |
| Enamine ID | Z104341206 | | |
| DS | −7.216 | −6.375 | −6.72 |
| LE | .38 | .49 | .198 |
| Hbond | Tyr606 | Asp663, Asn609 | | |
| π–π | Phe398 | Hid798, Tyr606 | | |
| Cation–π | Lys401 | Lys401 | | |
| Hydrophobic | Phe398, Lys401, Val411, Phe412, Phe485, Gly602, Val603 | Ser661, His798, Phe398, Val411, Phe412, Val603, Thr606, Gly662, Ser796, Trp795, Ile797, LF5.104 | | |
| Hit | FH5 | FH104 | |
| Enamine ID | Z650353014 | | |
| DS | −7.208 | −6.177 | −6.332 |
| LE | .36 | .562 | .192 |
| Hbond | Gly604 | Asp663, Ile797, Gly507 | | |
| π–π | Phe485, Phe398 | Hid798, Phe412 | | |
| Cation–π | Lys401 | Lys401 | | |
| Hydrophobic | Phe398, Lys401, Phe412, Leu488, Gly489, Asn492, Val603, Phe485 | Ser661, Gly662, Phe398, Lys401, Asn405, Phe412, Gly493, Glu507, Glu602, Val603, Thr605, Tyr606, Ile797 | | |

Notes: DS – Docking score; LE – Ligand efficiency; Hbond – Amino acids involved in Hydrogen bond formation; π–π – Amino acids involved in π–π interactions; Cation–π – Amino acids involved in cation–π interactions; Hydrophobic – Amino acids involved in hydrophobic interactions.

‡Similar interactions observed in both linked fragment and corresponding parent fragments.

3.3.2.1. Explicit fragment growth

The top 5 fragment hits (FH1-5) were considered as a core molecule for fragment growing processes. AutoGrow is a genetic algorithm which iterates for a give number of generations. It terminates the program when the molecular weight of the grown fragment exceeds 500 Da. The top fragment hit 1 (FH1) and FH2 were grown for 2 generations. Similarly, FH3, FH4, and FH5 were grown for 4, 1, and 3 generations, respectively. The top hits obtained in each generation for each of the fragment hits (FH1-5) are shown in Supplementary Figure S2.

The fragment hit 4 (FH4) after generation 1 produced a molecule with the highest binding affinity (DS = −7.985) with drug target DV RdRp. The carboxylate side chain
Figure 3. The binding pose and orientation of fragment hits FH2 (orange) and FH26 (blue) superposed on linked fragment hit LF2.26 (green).

Figure 4. Structure of (a) FH26, (b) FH2, and (c) LF2.26. FH2, FH26, and the linker are shown in red, blue, and green colors.
in the benzene ring of this AutoGrow hit 1 (AGH1) form salt bridge with Lys401 and hydrogen bonds with Asn492 and Lys401. The \( \pi \) electrons in the benzenothiazole ring forms cation–\( \pi \) interaction with Lys401 and \( \pi \)–\( \pi \) interaction with Phe398. The structure of AutoGrow hit 1 and its interactions are shown in Figures 8 and 9. AGH1 formed the hydrophobic interactions with the residues Phe398, Val402, Val411, Phe412, Phe485, Val603, Trp795, and Ile797.

### 3.3.2.2. Implicit fragment growth

#### 3.3.2.2.1. Substructure search.

The top fragment hits FH1-5 obtained in the initial screening process were searched against the PubChem and ChemSpider databases for all possible chemical compounds that include in the benzene ring of this AutoGrow hit 1 (AGH1) form salt bridge with Lys401 and hydrogen bonds with Asn492 and Lys401. The \( \pi \) electrons in the benzenothiazole ring forms cation–\( \pi \) interaction with Lys401 and \( \pi \)–\( \pi \) interaction with Phe398. The structure of AutoGrow hit 1 and its interactions are shown in Figures 8 and 9. AGH1 formed the hydrophobic interactions with the residues Phe398, Val402, Val411, Phe412, Phe485, Val603, Trp795, and Ile797.

#### 3.3.2.2. Implicit fragment growth

#### 3.3.2.2.1. Substructure search.

The top fragment hits FH1-5 obtained in the initial screening process were searched against the PubChem and ChemSpider databases for all possible chemical compounds that include

Figure 7. The binding poses obtained for LF2.26 and its interactions with the drug target DV RdRp.

Notes: The protein is shown in cartoon and ligand in ball and stick representations. Hydrogen bonds, \( \pi \)–\( \pi \) interactions, and cation–\( \pi \) interactions are shown in red, green, and magenta dashed lines. Only the interacting residues are shown for clarity in wire representation.

Figure 8. Structure of AutoGrow hit 1 (AGH1).
the structures of the fragment hits as substructures. The substructure search resulted in 4, 2, and 13 compounds from ChemSpider database for FH1, FH2, and FH4, respectively. Similarly, the search predicted 10, 2, and 306 compounds from PubChem database for FH1, FH3, and FH4, respectively. However, FH5 did not find any compounds from both the databases in the substructure search. Lipinski rule of five filters was applied with all the resultant compounds. This eliminated 225 compounds from the PubChem results of FH4 and provided 81 compounds. All the remaining compounds complied with the filter. Altogether 112 compounds were predicted for fragment hits FH1-5 in the substructure search.

All these 112 structures were prepared using LigPrep v3.1 and were screened against the target protein DV RdRp using Glide in SP. Based on the Docking score, top 5 compounds (Table 4) were selected for similarity or neighborhood search. The PubChem compound 24082586 exhibits the highest binding affinity with DV RdRp during Glide SP docking. The docking score of this top substructural hit 1 (SSH1) is −8.307 and the ligand efficiency is .244. The hydrogen bonds are formed between the amide side chain in piperidine ring with Ser796 and Hisd801, and the carbonyl oxygen with Lys401. The π electrons in the benzothiazole ring formed π–π interactions with Phe398 and cation–π interactions with Lys401. The structure and the interactions formed by this top substructural hit (SSH1) are shown in Figures 10 and 11. The hydrophobic interactions are formed between SSH1 and the amino acids Phe398, Val402, Met408, Val411, Phe412, Trp477, Phe485, Val603, Tyr606, Trp795, Ile797, and Ala799.

### 3.3.2.2. Neighborhood search

To get all possible compounds which are similar to the substructure search hits (SSH), similarity or neighborhood search was performed against the PubChem database. The similarity search found 245, 151, 741, 85, and 78 compounds which are similar to SSH1–5 in the PubChem database. Lipinski rule of five filter resulted in 108, 13, 389, 44, and 19 drug-like compounds in neighborhood search. Thus, a total of 573 compounds were obtained in the similarity search. These compounds were prepared using LigPrep v3.1 and were screened against the target protein using Glide in SP. The docking results predicted the same substructure search hit 1 (SSH1) as the top 1 hit in the neighborhood search. The 2D structure of the top 5 hits from this neighborhood search is shown in Supplementary Figure S3. The PubChem ID, docking score, and ligand efficiency values of the top 5 hits are shown in Table 5.

### 3.4. Selection of lead molecule

Fragment linking is successful when the linked fragments retain the binding poses and the orientations of the individual fragments. However, there should be an increase in its binding affinity when it bounds with the drug target. Even though LF2.26 (FH2 linked with FH26) shows the highest binding affinity (DS −7.457; LE −.226) with DV RdRp in the fragment linking process, it is less than that of the top fragment hit (FH1; DS: −7.661; LE: .403). Hence, the linked fragment LF2.26 is not considered as possible lead molecule.

### Table 4. Hits obtained in substructure search.

| Substructure search hit | PubChemID | Docking score | Ligand efficiency |
|-------------------------|-----------|---------------|-------------------|
| 1                       | 24082586  | −8.307        | .244              |
| 2                       | 23831983  | −8.071        | .212              |
| 3                       | 23887899  | −7.955        | .221              |
| 4                       | 24163859  | −7.858        | .212              |
| 5                       | 41576717  | −7.803        | .229              |

3.3.2.2.2. Neighborhood search. To get all possible compounds which are similar to the substructure search hits (SSH), similarity or neighborhood search was performed against the PubChem database. The similarity search found 245, 151, 741, 85, and 78 compounds which are similar to SSH1–5 in the PubChem database. Lipinski rule of five filter resulted in 108, 13, 389, 44, and 19 drug-like compounds in neighborhood search. Thus, a total of 573 compounds were obtained in the similarity search. These compounds were prepared using LigPrep v3.1 and were screened against the target protein using Glide in SP. The docking results predicted the same substructure search hit 1 (SSH1) as the top 1 hit in the neighborhood search. The 2D structure of the top 5 hits from this neighborhood search is shown in Supplementary Figure S3. The PubChem ID, docking score, and ligand efficiency values of the top 5 hits are shown in Table 5.

Figure 9. Interactions of AutoGrow hit (AGH1) with DV RdRp.
Notes: The target is shown in cartoon representation and the ligand in ball and stick representation. Hydrogen bonds, π–π interaction, and cation–π interactions are shown in red, green, and magenta dashed lines. Only the interacting residues are shown for clarity in wire representation.

Table 4. Hits obtained in substructure search.
Similarly, the binding affinity of AGH1 obtained in the explicit fragment growing process using AutoGrow is also less when compared with that developed implicitly. Hence, this AutoGrow hit (AGH1; DS: −7.985; LE: .307) is also not considered as possible lead molecule.

The implicit fragment growing methods such as substructure search and similarity or neighborhood search predicted a PubChem compound (PubChem ID 24082586) with the highest binding affinity (DS: −8.307) with the drug target DV RdRp. This is greater than that of the top fragment hit (DS: −7.661). In case of small molecular fragments, the concept of LE plays an important role. However, for larger lead-like molecules, the overall binding affinity should be considered. Hence, even though the LE of this PubChem compound (SSH1) reduced prominently during fragment growth, it is reported as a possible lead molecule for the development of a non-nucleoside inhibitor for DV RdRp.

3.5. Strain energy and binding free energy of the lead molecule

We have calculated the strain energy (the difference between the energy of the ligand in bound and unbound forms) and the ligand binding free energy for both the proposed lead molecule (PubChem ID 24082586) and the reference ligand NITD-107 using OPLS 2005. Interestingly, the strain energy of the proposed lead molecule is .019 kcal/mol, which is less than the reference ligand, which has 2.644 kcal/mol.

The binding free energy, $\Delta G_{\text{bind}}$, is calculated using Prime MMGBSA method. As the Prime MM-GBSA calculations require pose view files of the docked complexes in extra precision (XP), the target protein and the ligands were docked using Glide in XP. Prime MM-GBSA method combines OPLS-AA force field, molecular mechanics energies (EMM), SGB solvation model for polar solvation (GSGB), and a non-polar solvation term (GNP). The non-polar salvation term comprises non-polar...
solvent accessible surface area and van der Waals interactions. The total binding free energy is calculated with the equation: $\Delta G_{\text{bind}} = G_{\text{complex}} - (G_{\text{protein}} + G_{\text{ligand}})$. The binding free energy ($\Delta G_{\text{bind}}$) of the proposed lead molecule is $-78.661 \text{ kcal/mol}$, which is less than the binding free energy ($\Delta G_{\text{bind}}$) of the reference ligand, $-49.347 \text{ kcal/mol}$. Both the strain energy and the binding free energies of the lead molecule show that the lead molecule would be a better inhibitor than the existing DV RdRp inhibitor NIT-107.

3.6. Molecular dynamics simulation

In order to validate the stability of the docked complex, the binding mode, and the interaction pattern between the DV polymerase and lead molecule, MD simulation was carried for DV-RdRp–lead complex, in a physiological condition using explicit water molecule as solvent medium for 10 ns each. The best binding mode predicted from the docking result was taken for MD simulation.

The stability of the protein–ligand complex was assessed by measuring C$\alpha$ RMSD of the simulated structures with respect to their reference frame (starting structure). The analysis shows that RMSD values increases up to 3 ns and then become stable till 10 ns. This depicts that the initial conformation rearranges slightly in both the complexes and then becomes stable for the entire period of the simulation. The C$\alpha$ atoms and side chain atoms of the DV polymerase deviates within a range of $0.67$–$1.93 \text{ Å}$ and $1.09$–$2.84 \text{ Å}$, respectively, when it is in complex with the lead molecule (Figure 12). Similarly, the ligand RMSD with respect to their position in the active site was measured for the lead molecule and it is almost stable throughout the entire simulation period within a range of $1.2$–$3.2 \text{ Å}$ (Figure 13). Both the protein and ligand RMSD values ensure that the lead molecule is stable in the active site of the protein.

To predict the conformational flexibility of the complex, the root-mean-square fluctuation (RMSF) was calculated for both C$\alpha$ atoms and side chain atoms for each residue in the complex. The C$\alpha$ and side chain RMSF is plotted in Figure 14. The C$\alpha$ and side chain atoms of the polymerase residues fluctuate in a range of $0.36$–$2.31 \text{ Å}$ and $0.47$–$3.27 \text{ Å}$, respectively. The residues which interact with the ligand fluctuate less than non-interacting residues. The C$\alpha$ and side chain atoms of these interacting residues fluctuates in a range of $0.5$–$1.13 \text{ Å}$ and $0.66$–$1.85 \text{ Å}$, respectively, which ensures that the interactions are stable. The lesser fluctuation of C$\alpha$ atoms and side chain atoms ensure that the complex is less flexible or rigid in nature. The ligand RMSF was measured for all the atoms in the lead molecule. The ligand atoms fluctuate in a range of $0.68$–$1.18 \text{ Å}$. This shows that the ligand is more stable in the binding pocket. Both the RMSD and RMSF analysis depicts that the conformation of the polymerase–lead complex is stable and less flexible.

The conformation of the secondary structures is stable throughout the entire period of simulation. Only the residues form $\alpha$-helices and loops take part in interactions, whereas the residues from $\beta$-sheets do not involve in any interactions with the lead molecule.

| Neighborhood search hit | PubChem ID | Docking score | Ligand efficiency |
|-------------------------|------------|---------------|------------------|
| 1                       | 24082586   | $-8.307$      | .244             |
| 2                       | 23887899   | $-7.955$      | .221             |
| 3                       | 4829774    | $-7.944$      | .284             |
| 4                       | 73269831   | $-7.922$      | .264             |
| 5                       | 41576717   | $-7.803$      | .229             |

Table 5. Hits obtained in neighborhood search.

Figure 12. C$\alpha$ and side chain RMSD of DV polymerase in DV-RdRp–lead complex.
Next to the conformational stability studies using RMSD and RMSF analysis, trajectories obtained in MD simulation are analyzed for the interactions intermolecular interactions in DV-RdRp–lead complex and their consistency throughout the simulation period. The trajectories shows the various interactions formed by the polymerase with the lead molecule such as hydrogen bonds, hydrophobic, ionic, cation–π and π–π interactions. Besides, MDS also provide insights into the significance of water molecules which facilitates the polymerase–ligand interactions. The interacting residues, type of interactions formed by them, and their period of interactions are analyzed for the DV-RdRp–lead complex.

This analysis reveals that hydrogen bonds formed by Arg792, Thr794, Asn405, and Ser796 play a major role in making the DV-RdRp–lead complex stable. Next to hydrogen bond interactions, the hydrophobic interactions by Phe412, Phe398, and Ile797 and the water-mediated interactions by Val411, Asn405, Arg792, and Glu493 contribute equally for the stability of the polymerase–ligand interactions (Figure 15). In addition to these residues which form major interactions, hydrogen bonds formed by Lys401 and Ala799; π–π interactions formed by Phe412, Phe398, and Phe485; hydrophobic interactions formed by Val402, Met408, Val411, Phe485, Val603, Tyr606, and Trp795; ionic interactions and cation–π interactions formed by Lys401; and water-mediated interactions formed by Lys401, Thr404, Thr413, Asn492, and His800 are less significant. The binding pose of the lead molecule observed from the MD simulation of DV-RdRp–lead complex is shown in Figure 16.

In order to validate our docking results, the interaction pattern obtained from Glide docking is compared with the trajectories obtained in MD simulation (Table 6). A similar pattern of hydrogen bonds, hydrophobic interactions, cation–π interactions, and π–π interactions are observed in both docking and MD simulation. These includes the hydrogen bonds formed by Arg792, Ser796, Lys401; the π–π interaction between the π electrons in Phe398 and the benzene ring of benzothiazole part of the lead molecule; and the cation–π interaction formed between the cationic ε-amino group (NH$_3^+$) of Lys401 with the π electrons in the thiazole ring of benzothiazole part of the lead molecule.

The analysis of trajectories obtained from MD simulations conclude that the conformation of the DV-RdRp–lead complex and the intermolecular interactions are stable and consistent throughout the simulation period. The hydrogen bonds formed by Arg792, Thr794,
Ser796, and Asn405 and the hydrophobic interactions formed by Phe398, Phe412, and Ile797 are the major contributors in making this complex stable.

3.7. Synthetic accessibility

The proposed lead molecule SSH1 (PubChem CID 24082586) is available from the vendor Angene Chemical.
This overcomes the synthetic accessibility problems associated with lead molecules in drug discovery.

3.8. Conformational changes induced by lead molecules

The 3D structure of the target protein in complex with the proposed lead molecule is compared with its apo form (PDB ID 4HHJ) [27]. The lead molecule induces several conformational changes in the drug target DV RdRp. This includes disorder into order transition and changes in the secondary structures (Supplementary Figure S4).

The disorder to order transition is observed in the loops (Ala407-Asp419 and Glu458-Lys468) when lead molecule binds with the target protein. The hydrogen bonds formed by Arg792, Thr794, and Ser796; and hydrophobic interactions formed by Phe412, Phe398, and Ile797 might induce this conformational changes. As these interactions are stable, this might keep the polymerase in closed conformation.

In addition to this disorder to order transition, changes in the secondary structures are also observed. The amino acids Thr343-Val358 which is in α-helical structure in the apo form has been changed into loop when it binds with the lead molecule. But, the amino acids Trp302-Asn321 which is in the loop conformation in apo structure form β-sheets when the ligand SSH1 binds with the target protein.

During the initiation of polymerization, the polymerase has to shift from closed conformation into open conformation, so that RNA template can bind with the polymerase. The lead molecule SSH1 induces the conformational changes induced by the lead molecule makes the interactions strong and the complex stable. Thus, the polymerase might be kept in the closed
conformation which could not be changed into the open conformation and thereby the activity gets inhibited. This shows that the lead molecule might be potent NNI for DV RdRp.

### 3.9. Competitive inhibition

The target protein DV RdRp in complex with the lead molecule was compared with the 3D structure of hepatitis C virus (HCV) polymerase in complex with RNA template (PDB ID 4E7A) (Mosley et al., 2012). As the lead molecule SSH1 binds in the RNA-binding groove of the polymerase, there is steric hindrance between the lead molecules and the RNA template strand (Supplementary Figure S5). This suggests that the lead molecule may compete with RNA for its substrate binding site and thus the inhibition might be competitive.

### 3.10. ADME/toxicity prediction

QikProp predicted the most important properties that are useful in determining the physicochemical and pharmacokinetic properties of lead molecule. These properties are within the specified range to become a successful drug. The most important descriptor predicted by QikProp is Stars. It is the number of properties or descriptors that fall outside the 95% range of similar values for known drugs. The range of this descriptor is 0–5. The lower value of this descriptor reflects more drug likeliness, whereas the higher value denotes less drug likeliness. A value of 0 is obtained for this descriptor. This shows that the lead molecule has drug-like properties.

QikProp predicted 51 significant descriptors. The predicted values of the descriptors such as SASA (Total solvent accessible surface area), FOSA (Hydrophobic component of the SASA), FISA (Hydrophilic component of the SASA), PISA (π (carbon and attached hydrogen) component of the SASA), and WPSA (Weakly polar component of the SASA) fall within the specified ranges in QikProp physicochemical properties. Likewise, the values for the van der Waals surface area of polar nitrogen and oxygen atoms and carbonyl carbon atoms are fulfilling the limit in physicochemical calculation. The polar surface area (PSA) and the total solvent accessible...
volume (volume) are also within the limits, which show that these favor the molecules for binding with the target.

The physicochemical descriptors like molecular weight, H-bond donors, H-bond acceptors, and QPlogPo/w (octanol/water partition coefficient) are less than 5 which show that these lead molecules obey Lipinski rule of five. The molecular weight indicates that the compounds can diffuse very well in human body. Similarly, the number of hydrogen bond donors and acceptors present in the lead molecules shows that they can form sufficient number of interactions with the drug target. Likewise, the values of QPlogPo/w, QPlogPocht, QPlogPw, and QPlogKp show that the lead molecules can distribute very well within the body.

Aqueous solubility (QPlogS) of the lead molecule is in the allowed solubility range. The value for the qualitative human oral absorption is predicted as 3 in a scale of 1–3. This shows that the lead molecule has high oral absorption. The percentage of human oral absorption is greater than 70. The skin permeability (QPlogKp) value is also in the allowed permeability range. The Caco-2 cell permeability value of 71.714 shows that they would have an average absorption across the human intestinal

| QikProp descriptor | Property | Range or recommended values | SSH1 |
|--------------------|----------|-----------------------------|------|
| #stars             | Number of property or descriptor values that fall outside the 95% range of similar values for known drugs | 0–5 | 0 |
| mol MW             | Molecular weight of the molecule | 130.0–725.0 | 473.545 |
| Donorhb            | Estimated number of hydrogen bonds that would be donated by the solute to water molecules in an aqueous solution | 0.0–6.0 | 2 |
| Accepthb           | Estimated number of hydrogen bonds that would be accepted by the solute from water | 2.0–20.0 | 9 |
| Dipole             | Computed dipole moment of the molecule | 1.0–12.5 | 9.763 |
| Volume             | Total solvent accessible volume in cubic angstroms using a probe with a 1.4 Å radius | 500.0–2000.0 | 1389.902 |
| PSA                | van der Waals surface area of polar nitrogen and oxygen atoms and carbonyl carbon atoms | 7.0–200.0 | 118.152 |
| QPPolrz            | Predicted polarizability in cubic angstroms | 50.669 | 751.909 |
| SASA               | Total solvent accessible surface area (SASA) in square angstroms using a probe with a 1.4 Å radius | 300.0–1000.0 | 715.228 |
| FOSA               | Hydrophobic component of the SASA | 0.750.0 | 153.53 |
| FISA               | Hydrophilic component of the SASA | 7.0–330.0 | 171.484 |
| PISA               | π (carbon and attached hydrogen) component of the SASA | 0.450.0 | 394.499 |
| WPSA               | Weakly polar component of the SASA | 0.175.0 | 32.395 |
| IP(eV)             | PM3 calculated ionization potential (negative of HOMO energy) | 7.9–10.5 | 8.932 |
| EA(eV)             | PM3 calculated electron affinity (negative of LUMO energy) | −9.9–1.7 | 1.291 |
| QPlogS             | Predicted aqueous solubility, log S. S in moldm−3 is the concentration of the solute in a saturated solution that is in equilibrium with the crystalline solid | −6.5–5 | −3.995 |
| CIQProplogS        | Conformation-independent aqueous solubility | −6.5–5 | −5.095 |
| QPlogPo/w          | Octanol/water partition coefficient | −2.0–6.5 | 2.278 |
| QPlogPocht         | Octanol/gas partition coefficient | 8.0–35.0 | 26.219 |
| QPlogPw            | Water/gas partition coefficient | 4.0–45.0 | 21.081 |
| QPlogPC16          | Hexadecane/gas partition coefficient | 4.0–18.0 | 15.842 |
| QPlogKhsa          | Prediction of binding to human serum albumin | −1.5–1.9 | −2.03 |
| QPlogBB            | Predicted brain/blood partition coefficient | −3.0–1.2 | −1.38 |
| QPPCaco            | Caco-2 cell permeability in nm/s. Caco-2 cells is a model for the gut–blood barrier | <25 poor; >500 great | 71.719 |
| QPPMDCK            | Predicted apparent MDCK cell permeability in nm/s. MDCK cells are considered to be a good mimic for the blood–brain barrier | <25 poor; >500 great | 155.076 |
| QPlogKp            | Predicted skin permeability, log Kp | −8.0–1.0 | −2.906 |
| Human oral absorption | Human oral absorption | 1, 2, or 3 for low, medium, or high | 3 |
| Percent human oral absorption | Percentage human oral absorption | >80% is high; <25% is poor | 73.498 |
| CNS                | Predicted central nervous system activity | −2 to +2 | −2 |
| #metab            | Number of likely metabolic reactions | 1–8 | 4 |
gut wall. The values of the descriptor QPlogKhsa indicates that the lead molecule can bind well with human serum albumin. Overall, the physicochemical and pharmacokinetic properties predicted by QikProp program shows that the lead molecule is suitable for human use. The important descriptors and the values predicted for lead molecule SSH1 are shown in Table 7. The acceptable range of values defined by the QikProp program for each descriptor is also provided.

3.11. Similar features with existing RdRp inhibitors

The structure of the proposed lead molecule was compared with existing RdRp inhibitors of the dengue virus (DV), HCV, and human norovirus (HNV) (Table S2). We observed several similar features in the structure of the inhibitor as well as in the type of stabilizing interactions.

The HCV RdRp inhibitors BIW, HI4, 08E, SX3, SX5, SX6, and 23E have a bicyclic structure, which contain a six-membered benzene ring fused to a five-membered ring with one nitrogen atom. Our proposed lead molecule consists of the same bicyclic structure with an additional sulfur atom. The π electrons in these bicyclic structures are essential for the formation of both π−π and cation−π interactions with the receptor.

Similarly, the 4-piperidinecarboxamide group in the lead resembles the 4-piperidinecarboxylic acid group in the HCV RdRp inhibitor BIW. This carboxamide group forms several hydrogen bonds with DV polymerase. Likewise, the 1-(carbonyl)piperidine-4-carboxylic acid group in the lead molecule is also observed in the HCV RdRp inhibitors BIW, SX3, SX5, and SX6. This 1-carbonyl group, which form hydrogen bond with Arg501 in HCV, is essential in forming the stable cation−π interaction with Lys401 in DV RdRp.

The Phc398 forms a π−π interaction with the proposed lead molecule. Similar π−π interactions are also observed in the HCV RdRp inhibitors PH7 and PH9 with Tyr448; 1 MB, CCT, RNA, VR1, VRX and VXR with Trp528; and NN2 with Phc193. The significant cation−π interaction between the Lys401 and the lead molecule is also observed between Arg501 and the HCV RdRp inhibitor SX4.

The hydrogen bond between Ser796 and the proposed DV RdRp inhibitor is also observed with the HCV RdRp inhibitors 1 MB, PH9, 1M9, 08F, 221, CCT, NN3, RNA, SX5, SX6, VR1, and VRX. Similarly, the hydrogen bond interaction with Arg792 is also observed with HNV RdRp inhibitor 21D and HCV RdRp inhibitors SX3, SX4, SX5, SX6, BIW, RNA, 23E, and VGC.

The hydrophobic interaction between the proposed DV RdRp inhibitor with Ile797 is observed in the HCV RdRp inhibitors 1M9 and 23E. Similarly, the hydrophobic interaction with Val411 is observed in the HCV RdRp inhibitors BIW, VGC, and 23E. Likely the hydrophobic interaction with Phc398 and Phc412 is found in the DV RdRp inhibitor VWS and the HCV RdRp inhibitor NN2. These observations show that the proposed lead molecule would be a potent DV RdRp inhibitor.

4. Conclusion
Dengue is a painful, debilitating mosquito-borne viral infection which does not have any medication to treat. We applied FBDD technique and identified a potent inhibitor for DV RdRp enzyme. Fragment screening predicted a top fragment hit with docking score of −7.661 and ligand efficiency of .403 which was extended further by fragment growing and linking methods. This predicted the compound, [2-(4-carboxypiperidin-1-yl)-2-oxoethyl] 8-(1,3-benzothiazol-2-yl)naphthalene-1-carboxylate as a potent inhibitor for DV RdRp. MD simulation studies reveals that the conformation of DV-RdRp–lead complex is stable and the interaction pattern is consistent throughout the simulation. The favorable physicochemical and pharmacokinetic properties of this lead molecule ensure that this compound might be a potent inhibitor for DV polymerase. Further experimental studies and optimization of this lead would certainly result in a potent antiviral medication for dengue.

List of abbreviations
ADME Absorption, distribution, metabolism, and excretion
AGH AutoGrow hit
CID Compound identifier
DF Dengue fever
DHF Dengue hemorrhagic fever
DHODH Dihydroorotate dehydrogenase
DS Docking score
DSS Dengue shock syndrome
DV Dengue virus
EGFL Enamine golden fragment library
FBDD Fragment-based drug designing
FDA Food and drug administration
FH Fragment hit
HCV Hepatitis C virus
HNV Human norovirus
HTS High-throughput screening
LE Ligand efficiency
LF Linked fragment
MDS Molecular dynamics simulations
MTase Methyltransferase
NI Nucleoside/nucleotide inhibitors
NLS Nuclear localization sequence
NMR Nuclear magnetic resonance
NNI Non-nucleoside inhibitors
Supplementary material

The supplementary material for this paper is available online at http://dx.doi.org/10.1080/07391102.2015.1081620.

Acknowledgments

We thank High Performance Computing and Indian Institute of Technology Madras for computational facilities and Department of Biotechnology (DBT), Government of India, New Delhi, for Bioinformatics Infrastructure Facility (Madras University and Indian Institute of Technology Madras). We thank Professor KK. Balasubramanian for fruitful discussions.

Disclosure statement

The authors declare that they have no conflict of interest.

Funding

This work was supported by the Department of Science and Technology (DST), Government of India, New Delhi [grant number SR/WOS/LS62/2013].

References

Back, A. T., & Lundkvist, A. (2013). Dengue viruses – An overview. Infection Ecology & Epidemiology, 3, 1983–19860. doi:10.3402/iee.v3i0.19839

Bollag, G., Tsai, J., Zhang, J., Zhang, C., Ibrahim, P., Nolop, K., & Hirth, P. (2012). Vemurafenib: The first drug approved for BRAF-mutant cancer. Nature Reviews Drug Discovery, 11, 873–886. doi:10.1038/nrd3847

Bollati, M., Alvarez, K., Assenberg, R., Baroní, C., Canard, B., Cook, S., ... Bolognesi, M. (2010). Structure and functionality in flavivirus-NS-proteins: Perspectives for drug design. Antiviral Research, 87, 125–148. doi:10.1016/j.antiviral.2009

Dayan, G. H., Galan-Herrera, J. F., Forrat, R., Zambrano, B., Bouckenooghe, A., Harenberg, A., ... Lang, J. (2014). Assessment of bivalent and tetavalent dengue vaccine formulations in flavivirus-naïve adults in Mexico. Human Vaccines & Immunotherapeutics, 10, 2853–2863. doi:10.4161/hv.2014.972131

Durrant, J. D., Amaro, R. E., & McCammon, J. A. (2009). AutoGrow: A novel algorithm for protein inhibitor design.

Chemical Biology & Drug Design, 73, 168–178. doi:10.1111/j.1747-0285.2008.00761.x

Foudah, A. I., Sallam, A. A., Akl, M. R., & El Sayed, K. A. (2014). Optimization, pharmacophore modeling and 3D-QSAR studies of siphonolas as breast cancer migration and proliferation inhibitors. European Journal of Medicinal Chemistry, 73, 310–324. doi:10.1016/j.ejmech.2013.11.039

Gubler, D. J. (2006) Dengue/dengue haemorrhagic fever: history and current status. Novartis Foundation Symposium, 277, 3–16. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/?term=17319151

Hung, N. T. (2012). Fluid management for dengue in children. Paediatrics and International Child Health, 32, 39–42. doi:10.1179/0261268712Z.00000000051

Jorgensen W. L., & Duffy E. M. (2002, March 31). Prediction of drug solubility from structure. Advanced Drug Delivery Reviews, 54, 355–366. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/?term=11922952

Kawatkar, S., Wang, H., Czerminski, R., & Joseph-McCarthy, D. (2009). Virtual fragment screening: An exploration of various docking and scoring protocols for fragments using glide. Journal of Computer-Aided Molecular Design, 23, 527–539. doi:10.1007/s10822-009-9281-4

Kumar, A., Voet, A., & Zhang, K. Y. (2012) Fragment based drug design: From experimental to computational approaches. Current Medicinal Chemistry, 19, 5128–5147. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/?term=22934764

Lim, S. P., Somntag, L. S., Noble, C., Nilar, S. H., Ng, R. H., Zou, G., ... Shi, P. Y. (2011). Small molecule inhibitors that selectively block dengue virus methyltransferase. Journal of Biological Chemistry, 286, 6233–6240. doi:10.1074/jbc.M111.179184

Lim, S. P., Wang, Q. Y., Noble, C. G., Chen, Y. L., Dong, H., Zou, B., ... Shi, P. Y. (2013). Ten years of dengue drug discovery: Progress and prospects. Antiviral Research, 100, 500–519. doi:10.1016/j.antiviral.2013.09.013

Liu, M., Bender, S. A., Cuny, G. D., Sherman, W., Glicksman, M., & Ray, S. S. (2013). Type II kinase inhibitors show an unexpected inhibition mode against Parkinson's disease-linked LRRK2 mutant G2019S. Biochemistry, 52, 1725–1736. doi:10.1021/bi3012077

Low, J. G., Sung, C., Wijaya, L., Wei, Y., Rathore, A. P., Mosley, R. T., Edwards, T. E., Murakami, E., Lam, A. M., Watanabe, S., ... Shi, P. Y. (2013). Type II kinase inhibitors show an unexpected inhibition mode against Parkinson’s disease-linked LRRK2 mutant G2019S. Biochemistry, 52, 1725–1736. doi:10.1021/bi3012077

Moses-Isizi, T., Edwards, T. E., Murakami, E., Lam, A. M., Grice, R. L., Du, J., ... Otto, M. J. (2012). Structure of hepatitis C virus polymerase in complex with primer-template RNA. Journal of Virology, 86, 6503–6511. doi:10.1128/JVI.00386-12

Nguyen, N. M., Tran, C. N., Phung, L. K., Duong, K. T., Huynh, H. L., Farrar, J., ... Simmons, C. P. (2013). A randomized, double-blind placebo controlled trial of balapavir, a polymerase inhibitor, in adult dengue patients. Journal of Infectious Diseases, 207, 1442–1450. doi:10.1093/infdis/jit470

Niyomrattanakit, P., Chen, Y. L., Dong, H., Yin, Z., Qing, M., Glickman, J. F., ... Shi, P. Y. (2010). Inhibition of dengue virus polymerase by blocking of the RNA tunnel. Journal of Virology, 84, 5678–5686. doi:10.1128/JVI.02451-09
Velmurugan, D., Mythily, U., & Rao, K. (2014). Design and docking studies of peptide inhibitors as potential antiviral drugs for dengue virus ns2b/ns3 protease. *Protein and Peptide Letters, 21*, 815–827. Retrieved from http://www.ncbi.nlm.nih.gov/published?term=23855663

Vitek, C. J., Gutierrez, J. A., & Dirrigl, F. J., Jr. (2014). Dengue vectors, human activity, and dengue virus transmission potential in the lower rio Grande Valley, Texas, United States. *Journal of Medical Entomology, 51*, 1019–1028. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/?term=25276932

Wang, Q. Y., Bushell, S., Qing, M., Xu, H. Y., Bonavia, A., Nunes, S., … Shi, P. Y. (2011). Inhibition of dengue virus through suppression of host pyrimidine biosynthesis. *Journal of Virology, 85*, 6548–6556. doi:10.1128/JVI.02510-10

Wang, Q. Y., Kondreddi, R. R., Xie, X., Rao, R., Nilar, S., Xu, H. Y., … Shi, P. Y. (2011). A translation inhibitor that suppresses dengue virus *in vitro* and *in vivo*. *Antimicrobial Agents and Chemotherapy, 55*, 4072–4080. doi:10.1128/AAC.00620-11

Wang, Q. Y., Patel, S. J., Vangrevelinge, E., Xu, H. Y., Rao, R., Jaber, D., … Vasudevan, S. G. (2009). A small-molecule dengue virus entry inhibitor. *Antimicrobial Agents and Chemotherapy, 53*, 1823–1831. doi:10.1128/AAC.01148-08

Watanaveeradej, V., Gibbons, R. V., Simasathien, S., Nisalak, A., Jarman, R. G., Kerdpunich, A., … Innis, B. L. (2014). Safety and immunogenicity of a rederived, live-attenuated dengue virus vaccine in healthy adults living in Thailand: A randomized trial. *American Journal of Tropical Medicine and Hygiene, 91*, 119–128. doi:10.4269/ajtmh.13-0452

Weaver, S. C., & Vasilakis, N. (2009). Molecular evolution of dengue viruses: Contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral disease. *Infection, Genetics and Evolution, 9*, 523–540. doi:10.1016/j.meegde.2009.02.003

Wei, J., Chen, H., & An, J. (2014). Recent progress in dengue vaccine development. *Virologica Sinica, 29*, 353–363. doi:10.1007/s12250-014-3542-7

Whitehorn, J., & Farrar, J. (2010). Dengue. *British Medical Bulletin, 95*, 161–173. doi:10.1093/bmb/ldq019

Xie, X., Wang, Q. Y., Xu, H. Y., Qing, M., Kramer, L., Yuan, Z., & Shi, P. Y. (2011). Inhibition of dengue virus by targeting viral NS4B protein. *Journal of Virology, 85*, 11183–11195. doi:10.1128/JVI.05468-11

Yap, T. L., Xu, T., Chen, Y. L., Malet, H., Egloff, M. P., Canard, B., … Lescar, J. (2007). Crystal structure of the dengue virus RNA-dependent RNA polymerase catalytic domain at 1.85-angstrom resolution. *Journal of Virology, 81*, 4753–4765. Retrieved from http://www.ncbi.nlm.nih.gov/Published?term=17301146

Yin, Z., Patel, S. J., Wang, W. L., Chan, W. L., Ranga Rao, K. R., Wang, G., … Keller, T. H. (2006). Peptide inhibitors of dengue virus NS3 protease. Part 2: SAR study of tetrapeptide aldehyde inhibitors. *Bioorganic & Medicinal Chemistry Letters, 16*, 40–43. Retrieved from http://www.ncbi.nlm.nih.gov/Published?term=16246563

Yin, Z., Chen, Y. L., Kondreddi, R. R., Chan, W. L., Wang, G., Ng, R. H., … Keller, T. H. (2009). N-sulfonylanilinic acid derivatives as allosteric inhibitors of dengue viral RNA-dependent RNA polymerase. *Journal of Medicinal Chemistry, 52*, 7934–7937. doi:10.1021/jm901044z

Yin, Z., Chen, Y. L., Schul, W., Wang, Q. Y., Gu, F., Duraswamy, J., … Shi, P. Y. (2009). An adenosine nucleoside inhibitor of dengue virus. *Proceedings of the National Academy of Sciences, 106*, 20435–20439. doi:10.1073/pnas.0907010106