Research Article
CODEN (USA): IJPLCP

[Jasrai et al., 3(6): June, 2012]
ISSN: 0976-7126

INTERNATIONAL JOURNAL OF PHARMACY & LIFE SCIENCES

Exploring the polymerase activity of chikungunya viral non structural protein 4 (nsP4) using molecular modeling, e-pharmacophore and docking studies

S. Prasanth Kumar[1], Ravi G. Kapopara[1], Mehul I. Patni[1], Himanshu A. Pandya[1], Yogesh T. Jasrai[1]* and Saumya K. Patel[2]

1, Department of Bioinformatics, Applied Botany Center (ABC), Gujarat University, Ahmedabad, (Gujrat) - India

2, Department of Zoology, University School of Sciences, Gujarat University, Ahmedabad, (Gujrat) - India

Abstract

Chikungunya viral RNA-dependent RNA polymerase (RdRp) activity is conferred by non structural protein 4 (nsp4), an important protein target towards the development of antiviral compounds. The present study deals about the development of homology model of nsP4 followed by molecular docking with known RdRp inhibitors experimented in Hepatitis C virus (HCV), HIV-1, Paramyxovirus, etc. The predicted catalytic site and two allosteric binding sites were docked with nucleosidic and non-nucleosidic inhibitors. The best top five scoring ligands were selected based upon the interaction profiles and a common pharmacophore was developed. Further, RNA template-primer complex was docked within the template tunnel of modeled nsP4 to study the mode of polymerase activity.

Key-Words: Chikungunya, nsP4, RdRp, Docking, Pharmacophore, Polymerase activity

Introduction

Chikungunya virus (CHIKV), a member of the Alphavirus genus belongs to the family *Togoviridae* and it is primarily transmitted to humans by two main vectors, *A.aegypti* and *A. albopictus* [1]. With unprecedented order of magnitude going-on in the Indian Ocean territories since January 2005 [2] and some confirmed reports of infection in Southern India by Indian Council of Medical Research (ICMR), India [3], it is alive in Indian subcontinent after an interval of 32 years. Recurrent episodes of infections were also reported in some parts of Southeast Asia subcontinent. The scarcity of scientific knowledge on various epidemiological aspects intimidated the outburst of epidemic and the unavailability of suitable vaccine and/or specific antiviral agent added fuel to the fire. Hence, there is an immediate need to initiate research on this newly re-emerging evolutionary potent CHIKV infection.

*** Corresponding Author**
E.mail:yjasrai@yahoo.com,
prasanthbioinformatics@gmail.com

More difficulties are currently augmented to distinguish CHIKV infections from a spectrum of other viral infections as its symptoms are very much similar to other viral symptoms including nausea, vomiting, myalgia, rash and arthralgia and in some instance, the observation of painful puffy feet and ankles experiencing the chronic polyarthralgia, a discernible symptom of rheumatoid arthritis [4]. The CHIKV is enveloped and its genome consists of a linear, positive-sense, single-stranded RNA molecule of about 12 kb. Approximately 5' two-third of the genomic RNA encodes non-structural proteins viz. nsP1 (cytoplasmic capping enzyme), nsP2 (N-terminal: RNA trisphosphatase and RNA helicase activity; C-terminal: protease activity), nsP3 (subgenomic 26S mRNA synthesis) and nsP4 (RNA dependent RNA polymerase, RdRp) while 3' one-third corresponds to 26S subgenomic RNA encodes structural proteins C, E3, E2, 6K and E1 essential for other cellular activities including infection [5]. CHIKV nsP4 (EC 2.7.7.48) localizes to 2014-2462 sequence position in the polyprotein nsP1234 constitute 449 amino acids with polymerase activity [6]. In this report, we explained the structural properties of CHIKV nsP4 by developing a homology model based on the crystallographic data of Sapporo viral RdRp. With the comparison of structure-

based modeling approaches, we identified potential catalytic site bearing a triad and two allosteric binding sites with similar observation in Hepatitis C virus (HCV) ns5B polymerase (RdRp enzymatic activity). Literature published nucleosidic and non-nucleosidic inhibitors used in other viral infections towards the inhibition of RdRp were collected and studied their interactions using molecular docking approach. Further, the best ranked conformers were selected to generate a minimal pharmacophore which would suggest a minimum scaffold prerequisite for a molecule to have site-binding affinity. Finally, we developed a viral RNA duplex mimicking the template-primer complex was docked into the template tunnel to explore the mechanism of template-dependent replication and its electrostatic point of view.

Material and Methods

Sequence retrieval and sequence analysis

The protein sequence of CHIKV nsP4 and HCV ns5B was recovered from NCBI Reference Section (RefSeq) with the accession ids: NP_690588.1 and AAP33889.1 in FASTA format (best PDB template in FASTA format as described later) [7] and it was aligned with EMBOSS Stretcher [8] with the specification of BLOSUM62 scoring function and gap opening and extension penalty as 12 and 2. Motif analysis was carried out using default search settings over PROSITE database [9] to distinguish conserved domains of RdRps.

Homology model development

The model was developed based on the known templates with significant similarities and identities. The templates were identified using Swiss-Model template identification [10] and DaliLite version 3 [11] programs, respectively. The best ten templates were selected which scored significant e value in case of Swiss-Model and largest Z score in the analysis of DaliLite provided both of the hits secured plausible similarities and identities. Templates with Z score in the range between 41.3 and 45.9 were found with lowest e values as recognized by Swiss-Model. Hence, these crystallographic coordinates were selected to develop model.

Comparative modeling was executed using Modeller 9v7 which employs genetic algorithm protocol by selecting a set of initial alignments made with the target and progresses to iterative re-alignment with genetic operators such as mutations and cross-over to generate new alignments satisfying spatial restraints. The generated models can be judged by evaluating its scoring function parameters such as GA341 score (a composite of Z-score estimated using atomic statistical potential, target-template identity and structural compactness value) and DOPE (Discrete Optimized Potential Energy) score [12]. The model secured a DOPE score of -45298 with a GA341 score less than 0.6 was selected for further structural analysis. The loops were modeled by molecular dynamics (MD) based loop program engineered in Modeller 9v7.

Energy minimization and model assessment

The generated model was energy minimized using YASARA force field with the help of YASARA energy minimization server and returned a scene file operable in YASARA View (academic license) [13]. Now, the model was checked for stereochemistry and structure correctness using Ramachandran plot [14] implemented in SAVES (Structure Analysis and Verification Server) [15]. The secondary structural information was obtained using ExPASy PDBsum program [16]. The model was also graphically validated by superimposing the model with the template using Superimpose server. YASARA View returned root mean squared deviation (RMSD) value by providing the superimposed coordinates generated using Superimpose program to understand the structural arrangements.

Binding site prediction

It is well known that viral RdRps possesses catalytic and allosteric sites for its enzymatic activity. Q-SiteFinder was employed to identify ligand binding sites which utilize an interaction energy scheme to locate energetically favorable binding sites between the protein and a simple van der Waals probe. Spatial proximal scheme is used to cluster energetically favorable sites and then ranked by accounting its estimated interaction energies [17]. Cavities predicted over palm and thumb domains (described in next section) encompassed with a plausible volume were considered and cross-checked with the crystallographic data of HCV ns5B polymerase.

RNA duplex generation, docking and estimation of intermolecular energy

A RNA duplex (double-stranded RNA) with stem region constituted with 7 nucleotides (nts) was developed using the literature reported CHIKV secondary structure RNA model with the exception of hairpin. CONTRAfold program [18] was used to develop the duplex by specifying the customized input. The program takes advantage of stochastic context-free grammars (SCFGs) with enhanced fully-automated statistical learning algorithms to model RNA structure. The phosphodiester bonds were manually removed in one strand using a standard structure visualizer for a reason described in Results and Discussion section. Hex version 5 [19] was used to dock the RNA template inside the core cleft of CHIKV nsP4. Hex uses

spherical polar Fourier (SPF) representation to evade grid sampling in order to enhance 6D correlation master equation which has been calculated by each sampled pairwise interaction which is then used to derive generating functions (GFs) for 5D, 3D and 1D fast Fourier transformation (FFT) rotational correlations over angular terms constrained search space. Shape plus electrostatic samples were defined to spawn plausible docked samples. Ascalaph Graphics from Agile Molecules Inc. [20] was used to calculate the intermolecular energy of RNA duplex-CHIKV nsP4 complex using OPLS (Optimized Potentials for Liquid Simulations) force field.

Ligand dataset

Nucleosidic and non-nucleosidic inhibitors reported in bibliographic literature pertaining to viral RdRps were collected and drawn using Marvin Sketch utility of Marvin applications [21]. The drawn structures were then 'cleaned' using Clean in 3D utility. The dataset was further distinguished into nucleosidic and non-nucleosidic inhibitors (Supplementary Figure 1, 2A and 2B). Care was taken to consider one unique scaffold in case of derivatives to better understand the diversity of molecular interaction with predicted binding sites. To select a random derivative, the reference molecule with chemical space was subjected to undergo combinatorial expansion using ORVIL (ORganic Virtual Library), a perl based program to generate virtual organic library within an interval of 2-3 minutes [22].

Small molecular docking and pharmacophore development

Docking simulations was performed on a single machine equipped with Intel Dual-Core processor with 4 GB RAM and 150 GB hard disk operating on an Ubuntu environment using iGEMDOCK version 2 program [23]. iGEMDOCK computes a ligand conformation and relative orientations owing to search space relative to the protein target binding site by applying genetic evolutionary methods (GA). Screening was carried out with the following specifications: population size: 500, number of generations: 100 and number of solutions: 10. Followed by generation of interaction profiles, the pharmacological interactions were mined and assigned a Z score. We sorted the interaction profiles in the following order: binding energy, H bonding term and van der Waals term with electrostatic term remained undisturbed. The top most clusters comprised of five best docked conformers were considered to develop minimal 3D pharmacophore using features-based alignment in Ligand Scout (trial version) [24]. Ligand Scout in its first step discerns and corrects molecules by ring perception, hybridization steps assignment, bond order corrections progressed with abstraction level chemical features schema to predict hydrogen bonding, charge and hydrophobic interactions.

Results and Discussion

Homology model of CHIKV nsP4

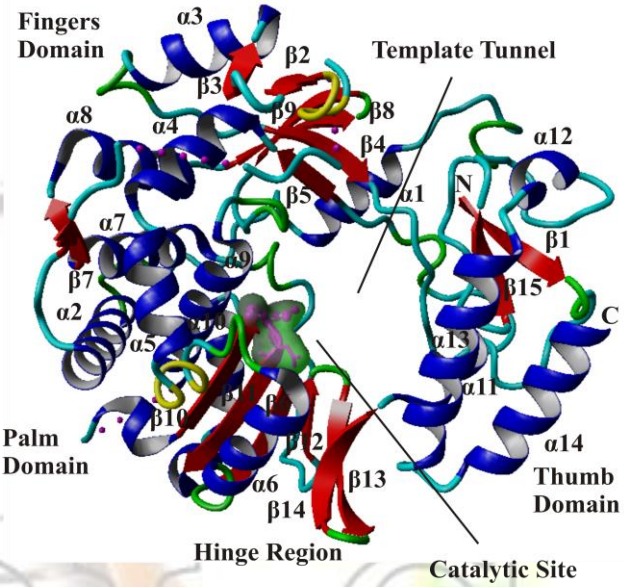


Fig. 1: Homology model of CHIKV nsP4

Homology modeling was performed to determine the structure of CHIKV nsP4 using Modeller 9v7. Suitable templates were selected based upon the consensus generated by Swiss-Model template identification and DaliLite version 3 programs. The templates were selected based upon the scoring functions employed in the respective programs.

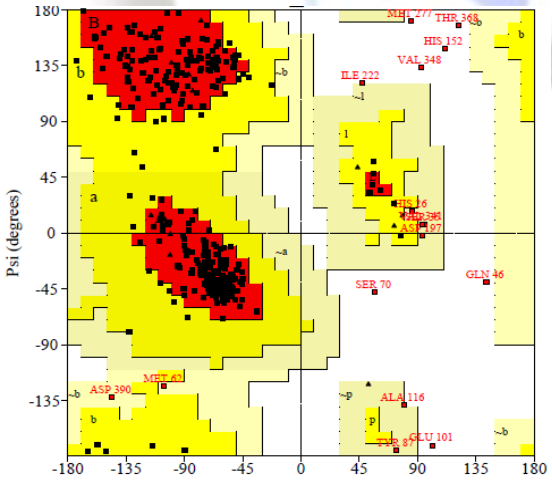


Fig. 2: Ramachandran plot of CHIKV nsP4 favoring 95.8 % in core region

Templates identified with a lowest e value (Swiss-Model) and largest Z score was considered for modeling. This analysis yielded best five templates viz. Sapporo virus RdRp (PDB code: 2UUT) D346G mutant protein, Sapporo virus RdRp (2UUW) D347G and D348G mutant protein, Sapporo virus RdRp (2CKW), Sapporo virus dimeric RdRp (2WK4) D347G and D348G mutant protein and Rabbit hemorrhagic disease virus RdRp complexed with Lutetium (III) ion (1KHV). The RMSD of 1.0 and largest Z score (2UUT: 45.9 in comparison with 2CKW: 43.2) was the choosing factor to consider a mutant protein instead of a native one. The model with a lowest DOPE score and a GA341 score less than 0.6 was considered and loop refinement was carried out using in-built MD based loop program (Figure 1).

The homology model was in good correspondence with the templates assigned. The N-terminal region (12 amino acids) was not modeled due to the template-model poor alignment. Ramachandran plot exhibited 95.8% in ψ-φ core areas with an exception of seven amino acids excluding Gly and Pro occupied in disallowed regions (Figure 2). The CHIKV nsP4 constituted fourteen helices and fifteen beta sheets in comparison with twenty helices and eighteen beta sheets observed in the template. Upon referring the secondary structural information of 2UUT [25], the N-terminal region comprised of two helices connected by a turn and in some regions, the beta sheets were modeled as loop regions in the model due to the low similarity and identity values. Global alignment of template with CHIKV nsP4 using EMBOSS Stretcher using BLOSUM62 protein scoring matrix revealed an identity of 16.2 %, similarity of 32.9 % with a gap of 14.9 % overall making a score of -144. Hence, it is evident that *Alphaviridae* and *Calciviridae* family are sufficiently diverged.

Structural features of CHIKV nsP4

The CHIKV nsP4 model was assigned their secondary structural information (Figure 3) based upon read-through from the N-terminal and no nomenclature was considered due to the unavailability of information related to bibliographic citations in the templates (unpublished bibliographic data as of 23.04.2012 pertaining to PDB template coordinates: 2UUT, 2UUW). Thus, the Sapporo viral (native RdRp) crystallographic studies were utilized in the present study to conclude [25]. The CHIKV nsP4 and the HCV ns5B are related to each other due to RdRp activity despite the low sequence conservation (identity: 4.9 %, similarity: 12.2 %, gaps: 75.5 % and score: -604; calculated using EMBOSS Stretcher). The model resembled with a right hand configuration usually observed in RdRps and distinguished into three domains viz. fingers, thumb and palm domain with a wide hollow tunnel for RNA template and/or primer interaction and rNTPs binding similar to HCV ns5B polymerase [26].

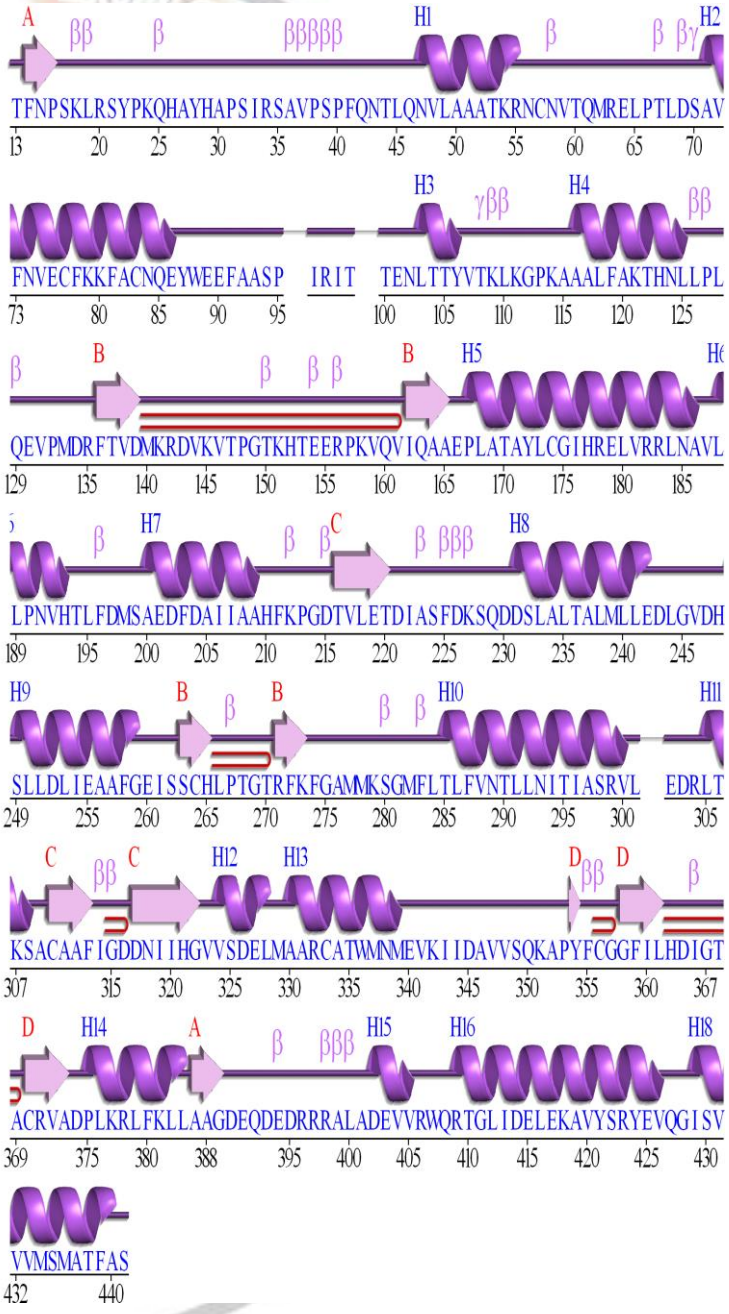


Fig. 3: Secondary structure information of CHIKV nsP4

Fig. 4: Superimposed view of CHIKV nsP4 with Sapporo viral RdRp

The N-terminal domain comprises two antiparallel beta strands forming a twisted configuration followed by extended interwoven loops connecting alpha helices and beta strands projected between the finger and thumb domains. This appearance has been previously documented in RdRp structures [26]. It is also known that the N-terminal region (residues in the range of 12 to 54) undergoes significant conformational variation which in turn, related to the open (inactive) or closed (active) conformation required for polymerase activity [27]. In fact, the unmodeled 12 amino acids in the N-terminal corresponding with the two beta strands of the Sapporo viral RdRp perhaps required for polymerase activity. Further, the modeled protein was superimposed with the chosen target 2UUT and returned an RMSD of 0.502 by chemically flipping equivalent chemical groups in YASARA View. Superimpose server yielded precise information on the geometrical topology between the template and the generated model (Figure 4). The developed model was submitted to Protein Model Database [28] and assigned the PMID: PM0078108.

Catalytic and allosteric binding sites

CHIKV nsP4 encompassed with a highly curated catalytic triad, Gly315, Asp316 and Asp317 (GDD motif) responsible for polymerase activity with a volume of 2769.41 Å^3 whereas allosteric binding sites with a volume of 73.216 Å^3 and 55.904 Å^3, were identified on both of the palm and thumb domains in similar topology to HCV ns5B predicted by Q-SiteFinder. Hence, we anticipated that these cavities will act as allosteric binding sites (Table 1, Figure 5). Further, a volume of 2769.41 Å^3 essentially paves way for interaction with primers and template. This cavity displays certain loop segments similar to Sapporo viral RdRp known to be prominent for RNA-primer duplex. Motif analysis based on primary sequence over ExPASy PROSITE database yielded no information over conserved motifs as studied in HCV, HIV-1 RdRps but only distinguished the catalytic triad.

Table 1: Catalytic and allosteric sites predicted by Q-SiteFinder

Site*	Localization	Interfacing amino acids in the cavities	Predicted volume
Catalytic	Adjacent to template tunnel	Ile314, Gly315**, Asp316**, Asp317**, Asn318, Phe225, Asn293, Cys356	2769.41 Å^3 (This cavity shares its surface area with template tunnel)
Allosteric	Palm	Pro95, Ser94, Pro113, Lys114, Phe119, Ala171, Gly175	73.216 Å^3
Allosteric	Thumb	Asn15, Pro16, Ser17, Lys18, Ala388, Gly389, Asp390, Val405, Arg409, Leu412, Ile413, Phe439, Ala440	55.904 Å^3

* Exposed to surface within the proximity of 5 Å
** Catalytic triad usually observed in RdRps.

Mechanism of RNA replication

It is known that plant alpha virus –like virus takes the advantage of *de novo* (primer independent) mechanism for initiation of replication including the members of the *Flaviviridae* family viz. HCV, Dengue 2, etc [29]. However, the presence of internal polyadenylation site (I-poly(A)) suggested the mechanism of template-dependent and observed experimentally [6]. A wide central cleft having a volume of 2769.41 Å^3 may accommodate a template-primer duplex and will be

easily accessible for active site (Figure 6). Contrasting to HCV ns5B polymerase, the thumb domains does not appear to protrude and occlude the active site, which makes it a more compact molecule. Electrostatic view demonstrates the need of positively charged clusters around the tunnel in a view to allow RNA template and NTP substrate to access the active site. The occupancy of Lys, His and Arg residues contributed to the positive region of the tunnel and found to be properly distributed over the surface (Figure 7A).

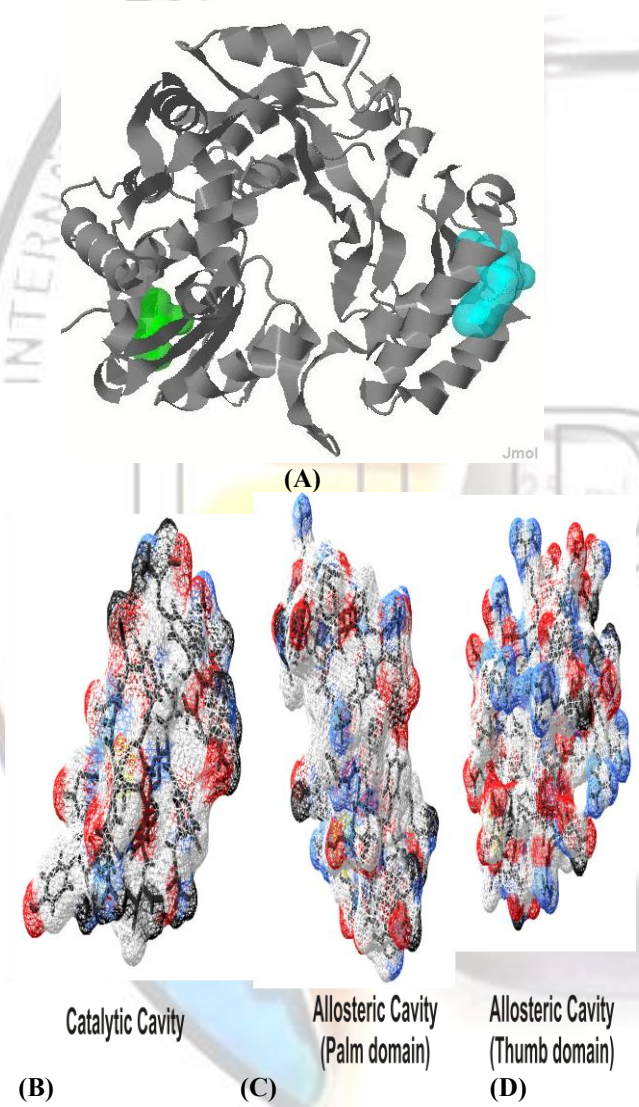


Fig. 5: (A) The location of predicted allosteric binding sites on the palm and thumb domain. Electrostatic view of catalytic (B), allosteric cavity on palm domain (C) and on thumb domain (D).

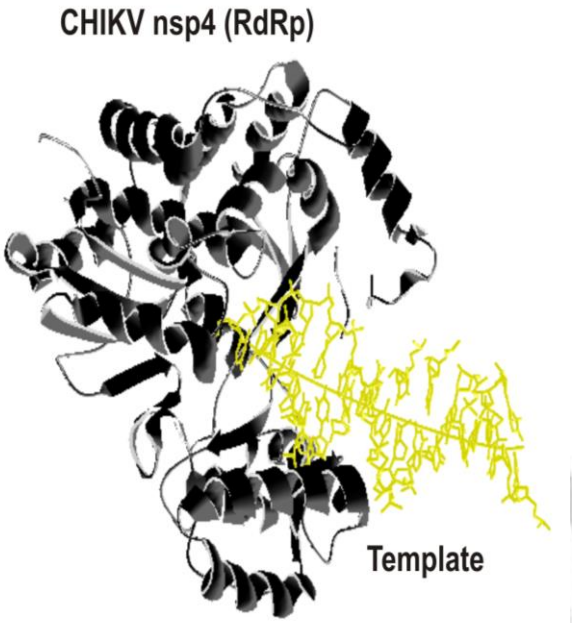


Fig. 6: Docked pose of RNA template-primer inside the CHIKV nsP4 core cleft

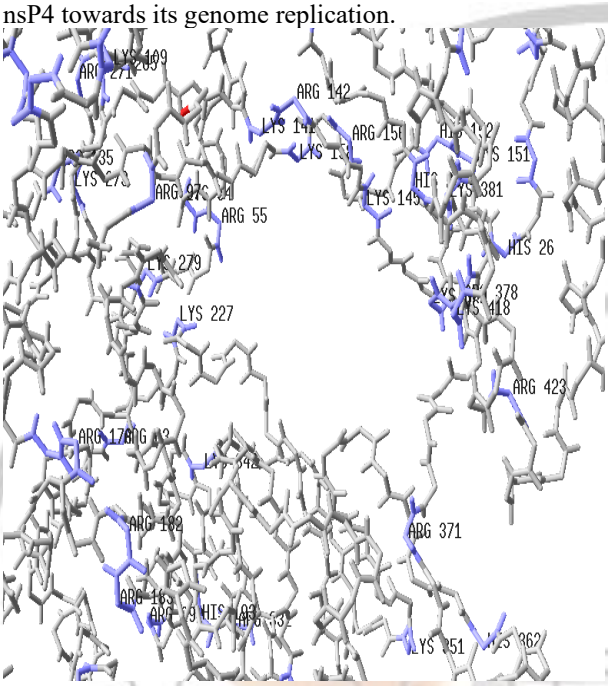
A RNA duplex consisting of viral template with rNTP was developed using the secondary structure published by Khan *et al.*, 2002 [6]. The stem region (7 nts) was considered for generation of duplex and hairpin was excluded from our analysis. To mimic a condition of placing the rNTP in the catalytic triad, we manually removed the phosphodiester linkage of adjacent nucleotides in one of the strand whereas the template strand remains undisturbed. This effort ensured the plausible placement of rNTPs and template in the template tunnel and manually checked by translating the RNA duplex inside the cavity using standard structure visualizer.

Interaction of RNA template-primer duplex inside CHIKV nsP4 cavity

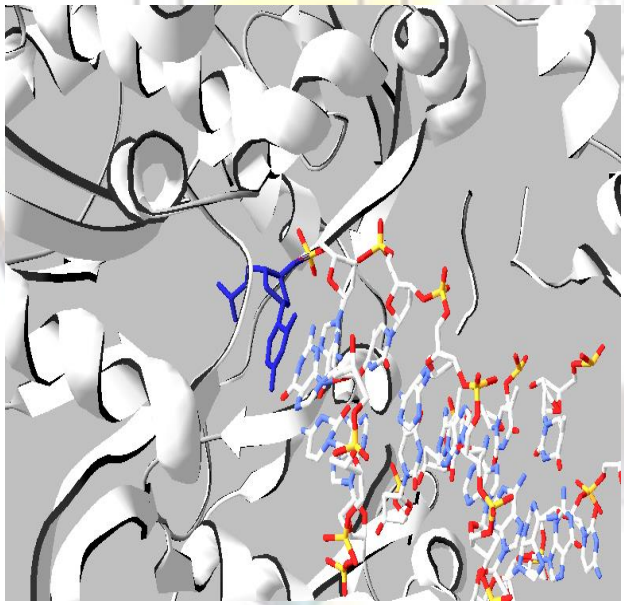
The conformational space of tunnel cleft is very large in magnitude compared to HCV ns5B. Hex version 5 protein-nucleic acid docking procedure was adopted which is known for its multi-dimensional rotational FFT algorithm with shape plus electrostatic correlations. The docked conformation reveals that rNTP (highlighted in dark blue color) occupies the catalytic cavity whereas RNA template (colored by atom-type definitions) protrudes towards the hollow cleft (Figure 7B). The intermolecular energy of CHIKV nsP4 and RNA template together with rNTPs was calculated using Ascalaph Graphics. The energy was estimated as 4.10777e+ 012 Kcal/mol with van der

[Jasrai et al., 3(6): June, 2012]
ISSN: 0976-7126

Waals component energy as the most promoting interaction component. Thus, we believe that a template-based mechanism was espoused by CHIKV nsP4 towards its genome replication.



(A)

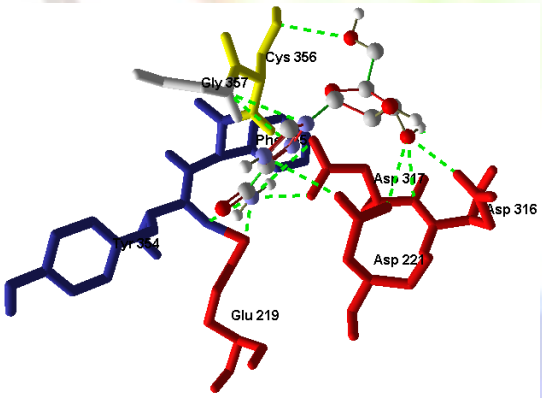


(B)

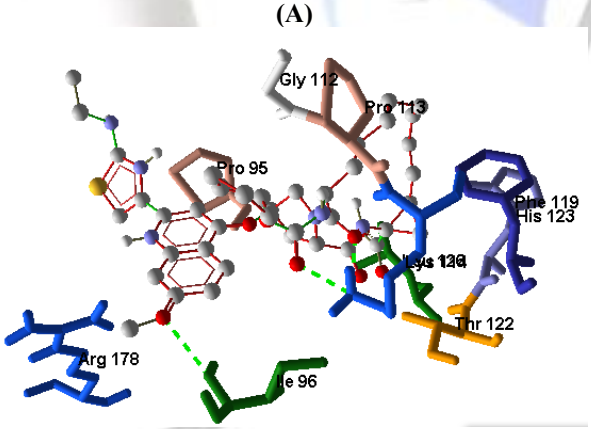
Fig. 7: (A) The positive charged amino acids at the core cleft of CHIKV nsP4 and (B) the environment of RNA template-primer (primer highlighted in dark blue color) resembling the state of template-dependent replication mechanism

Small molecular docking of nucleosidic and non-nucleosidic inhibitors

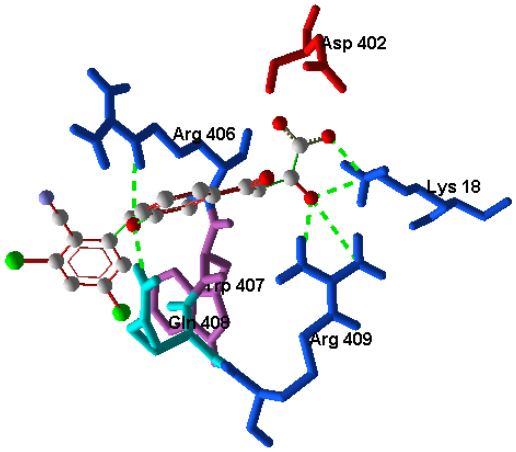
The catalytic triad (Gly315, Asp316 and Asp317) along with its neighbours within the proximity of 5 Å was implemented as active site (binding) region overall making a volume of 2769.41 $Å^3$ and nucleosidic inhibitors (21 molecules) reported in bibliographic literatures were considered. Two allosteric binding sites predicted over the palm and thumb domains by Q-SiteFinder were taken. The interfacing amino acids in the allosteric sites were opted by initially selecting the centroid amino acid in the respective cavities and extended to a span of 5 Å and exported as potential binding site for non-nucleosidic inhibitors (25 molecules). After the definition of appropriate binding site and ligand dataset in iGEMDOCK version 2 molecular docking program with GEMDOCK scoring function, the docking simulations were executed by multiple-iterative experiments and the best cluster (higher occurrence rate of similar clusters) is presented in this study (Figure 8).



(A)



(B)



(C)

Fig. 8: Docked pose of (A) Ribavirin, (B) BILN2061 and (C) diketo acid derivative interacting with the amino acids present in the catalytic and allosteric sites on palm and thumb domains

Table 2: Top most cluster constituting best docked poses in various docking experiments

Nucleosidic Inhibitors	Total Energy	VDW	H Bond
Ribavirin	-85.79	-46.1575	-39.6325
2'-C methyl guanosine	-85.4173	-56.4992	-28.9181
5- substituted cytidine analog 3	-81.7609	-55.6309	-26.13
2'-C methyladenosine	-78.7082	-54.5621	-24.1461
5- substituted cytidine analog 4	-78.3301	-39.0003	-39.3298
5- substituted cytidine analog 1	-78.049	-54.5715	-23.4775
GS-327073	-76.8737	-70.1595	-6.7142
5-bromo-2'-deoxycytidine	-76.7051	-50.1566	-26.5485
5-hydroxy-2'-deoxycytidine	-75.966	-49.1361	-26.8298
5- substituted cytidine analog 2	-74.7989	-52.5615	-22.2374
Non-nucleosidic Inhibitors (Palm domain)			
BILN2061	-94.294	-83.8392	-10.4548
JTK109	-91.998	-82.3032	-9.05339
6,7-dihydro-5H-benzo[5,6][1,4]diazepino[7,1-a]indoles	-91.1732	-76.0949	-14.0254
Benzimidazole derivative	-88.5878	-81.581	-7.00675
MSC204	-88.0289	-67.4467	-20.5822
4-hydroxy-2(1*H*) quinolinone derivative	-87.7853	-69.0233	-18.7619
Benzothiadiezine derivative	-86.1588	-72.1905	-13.9683
Acyl pyrrolidine derivative	-85.3859	-72.012	-11.6829
Diketo acid derivative	-80.9219	-58.6972	-19.6739
Phenyl derivative	-80.5368	-74.9324	-5.60442
Non-nucleosidic Inhibitors (Thumb domain)			
Diketo acid derivative	-95.6317	-71.635	-24.1654
Acyl pyrrolidine derivative	-93.6228	-75.4513	-12.9505
Piperidine derivative	-93.5658	-76.6599	-16.9059
JTK109	-91.5251	-75.2158	-16.3093
Benzimidazole derivative	-91.3149	-77.2781	-13.7247
BILN2061	-91.1136	-85.7133	-5.40032
Thiophene derivative	-91.0267	-72.7634	-18.2633

4-hydroxy-2(1*H*) quinolinone derivative	-89.063	-83.2217	-5.8413
739W94	-88.6384	-71.5732	-17.0652
Benzothiadiezine derivative	-86.9676	-79.9676	-7

The interaction profile generated over the docking experiments was sorted in the following order: binding energy, H bonding term and van der Waals term (Table 2). Ribavirin, the widely prescribed drug for treating HCV infections was found to be the top ranked docked conformer with an energy of -85.79 Kcal/mol in accordance with our previous observation on HCV ns5B polymerase [30] whereas the nucleosidic analogues resembling the core scaffold of guanosine, cytidine, adenosine was observed to be the best docked conformations down the clustered hierarchy with an energy distributed over the range of -85.4173 and -78.049 Kcal/mol.

Non-nucleosidic inhibitors were docked on both of the allosteric sites localized in palm and thumb domains, respectively. This blind docking helped us to identify molecules with greater binding affinity towards the site and discriminated the specificity to develop site-specific pharmacophores. Palm allosteric site analysis exhibited that BILN 2061, a potent HCV serine protease inhibitor, was the best docked molecules in a view of its energy term, -94.294 Kcal/mol. JTK-109, an effective inhibitor of HCV ns5B with a low nanomolar biochemical potency and a submicromolar cellular potency and MSC204, a non-nucleosidic reverse transcriptase inhibitor (NNRTI) of HIV-1 achieved the top cluster with an energy of -91.998 and -88.0289 Kcal/mol. Within the best five docked poses, an indole (6,7-dihydro-5H-benzo[5,6][1,4]diazepino[7,1-a]indoles) and benzimidazole derivatives was found to be better interacting with CHIKV nsP4 palm allosteric site. The top most clusters showed that van der Waals energy component was the major interaction component rather than H bonding term as it scored with an energy in the range of -83.8392 and -67.4467 Kcal/mol. Thumb allosteric site analysis revealed that diketo acid (energy: -95.6317 Kcal/mol), acyl pyrrolidine (-93.6228 Kcal/mol), piperidine derivatives (-93.5658 Kcal/mol) clustered in the top rank whereas JTK-109 (-91.5251 Kcal/mol) and benzimidazole derivative (-91.3149 Kcal/mol) followed its interaction profile in the same cluster.

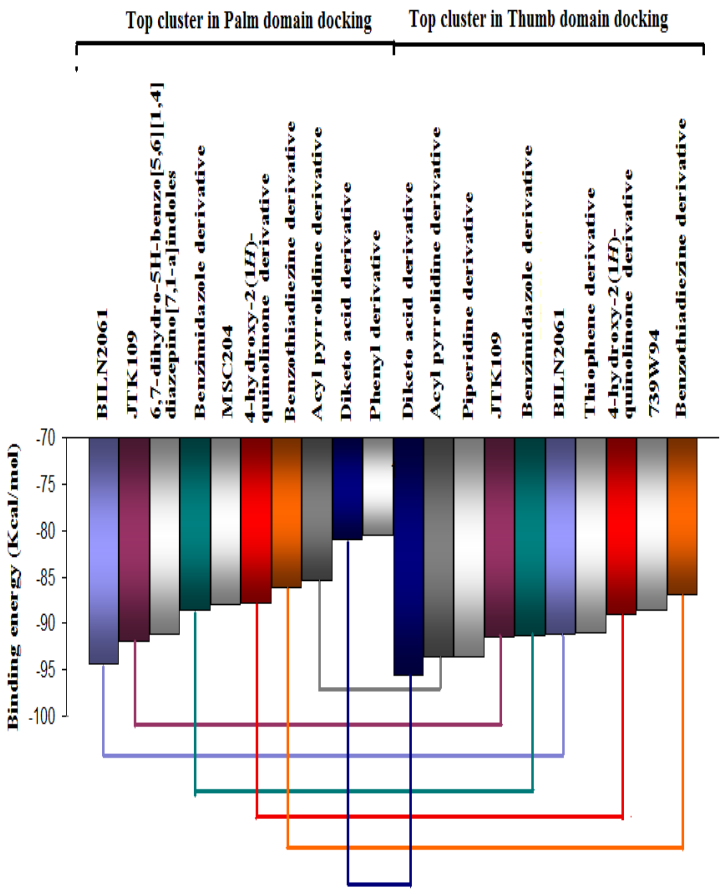
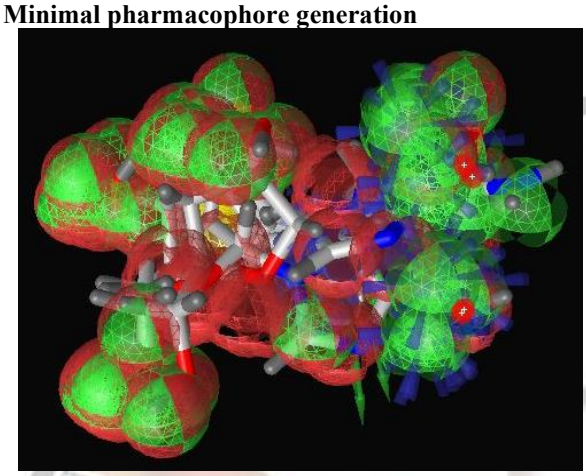


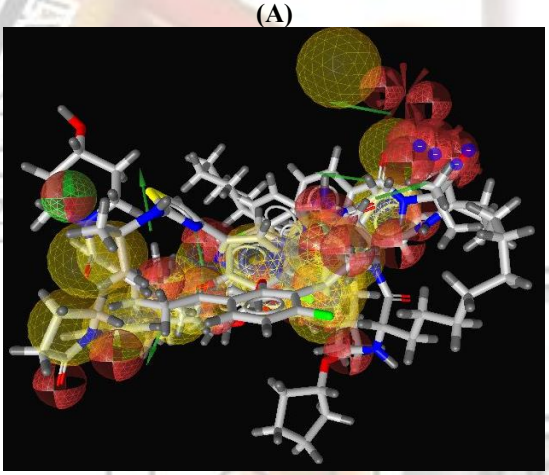
Fig. 9: Molecules showing plausible bi-specific interactions with both of the allosteric sites (interconnected by nodes). Unique molecules are highlighted in white blocks

Upon referring the top clusters of docked conformations pertaining to palm and thumb domains specific docking experiments, the following molecules were able to bind with both of the allosteric sites with desirable binding energies: BILN 2061, JTK 109, benzimidazole, 4-hydroxy-2(1*H*)-quinolinone, benzothiadiezine, acyl pyrrolidine and diketo acid derivatives (Figure 9). It is evident that the specificity of molecules towards the binding sites were well discriminated. For example, BILN 2061 was the best docked conformer of palm domain whereas its interaction with thumb domain was significantly less as it occupied 6th position in the energy based sorting. Following this approach, some unique molecules including 6,7-dihydro-5H-benzo[5,6][1,4]diazepino[7,1-a]indoles, MSC 204 and phenyl derivative in palm allosteric site whereas piperidine, thiophene derivatives and 739W94 in thumb allosteric sites shown certain level of specificity.

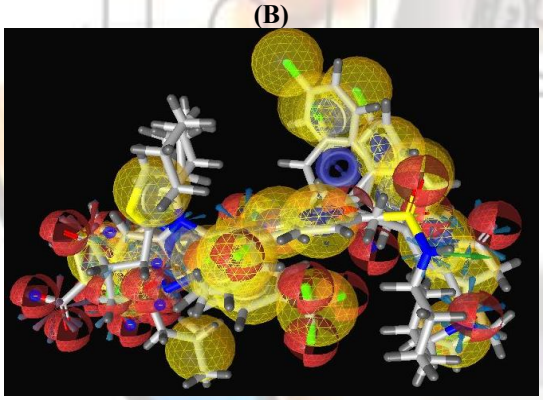
Minimal pharmacophore generation



(A)



(B)



(C)

Fig. 10: Shared pharmacophore of (A) nucleosidic (catalytic) and non-nucleosidic inhibitors on palm (B) and thumb (C) allosteric sites selected from top most cluster in various docking experiments.

The common pharmacophores were generated by selecting the top most five docked conformations in accordance with individual site-specific docking experiments (Figure 10). The minimum core for catalytic specific (nucleosidic) inhibitors was found to be 4 H bond donors (HBDs), 7 H bond acceptors (HBAs) and 1 aromatic ring (R). The minimum core for non-nucleosidic inhibitors related to palm allosteric site was observed to be 1 HBD, 4 HBAs, 2 Rs and 3 hydrophobic regions (Hs; 2 from aromatic ring and 1 from aliphatic group) while 6 HBAs, 2 Rs and 3 Hs (2 halogens and 1 from aromatic ring) pertaining to thumb allosteric site. Thus, it was demonstrated that nucleosidic inhibitors require a minimum scaffold of NTPs whereas non-nucleosidic inhibitors necessitates the scaffold with aromatic rings and hydrophobic regions in concordance with the electrostatic surface requirements with certain orders of H bonding chemical functionalities to form interaction with side and/or main chains.

A recent Chikungunya outbreak reported in Indian Ocean territories with unprecedented magnitude and some reports in Southern India. CHIKV nsP4 promises a potential target for development of antiviral agents similar to HCV infections. In the present study, we attempted to develop a homology model of CHIKV nsP4 and distinguished their interactions with known inhibitors widely used in treating viral infections. The model looks like a typical Alphavirus RdRps and bear a resemblance to Sapporo virus RdRp. Catalytic site along with two allosteric binding sites were predicted in a similar fashion related to HCV ns5B crystallographic data. Ribavirin, a nucleosidic analogue (inhibitor) interacted with the catalytic site with significant binding energy. Non-nucleosidic inhibitors such as BILN 2061 and diketo acid derivative better interacted with palm and thumb allosteric sites, respectively. The top most clustered conformations were utilized to develop minimal pharmacophore. Nucleic acid based docking experiment showed the preference of template-dependent mechanism and studied its interaction with positive regions encircled the template tunnel. We anticipate that this study will help to contribute a significant level in the Chikungunya drug discovery pipeline.

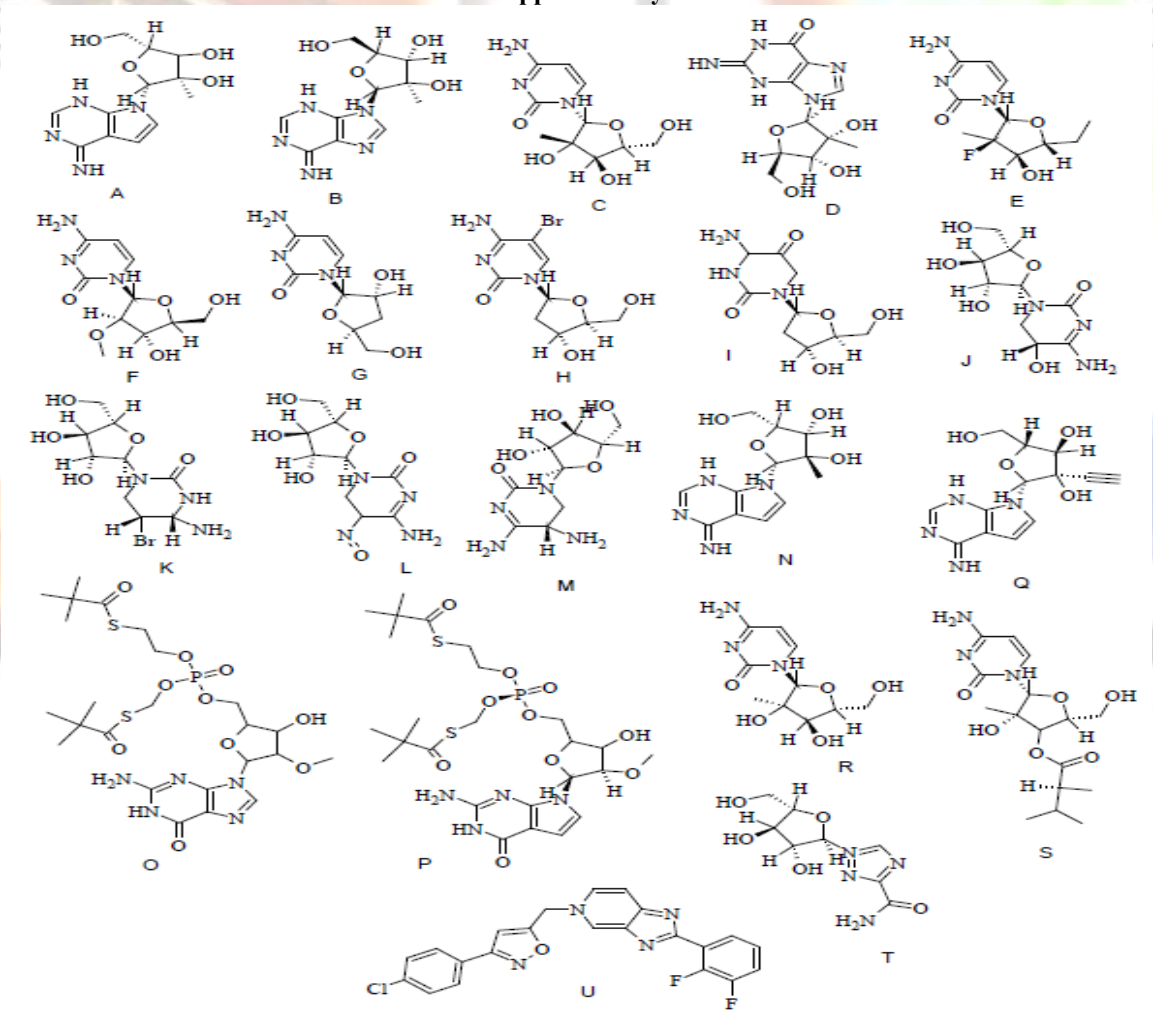
References

1. Sourisseau M, Schilte C, Casartelli N, Trouillet C, Guivel-Benhassine F, Rudnicka D, Sol-Foulon N, Roux KL, Prevost M-C, Fsihi H, Frenkiel M-P, Blanchet F, Afonso PV, Ceccaldi PE, Ozden S, Gessain A, Schuffenecker I, Verhasselt B, Zamborlini A, Saib A, Rey FA, Arenzana-Seisdedos F,Despres P, Michault A, Albert ML, Schwartz O. Characterization of reemerging

Chikungunya virus. PLoS Pathog. 2007; 3:804–817.
2. Schuffenecker I, Iteman I, Michault A, Murri S, Frangeul L, Vaney MC, Lavenir R, Pardigon N, Reynes JM, Pettinelli F, Biscornet L, Diancourt L, Michel S, Duquerroy S, Guigon G, Frenkiel MP, Brehin AC, Cubito N, Despre P, Kunst F, Rey FA, Zeller H, Brisse S. Genome Microevolution of Chikungunya Viruses Causing the Indian Ocean Outbreak. PLoS Med. 2006; 3(7): e263.
3. Samuel PP, Krishnamoorthi R, Hamzakoya K, Aggarwal CS. Entomo-epidemiological investigations on chikungunya outbreak in the Lakshadweep islands, Indian Ocean. Indian J. Med. Res. 2009; 129:442-445.
4. Vanlandingham DL, Hong C, Klingler K, Tsetsarkin K, McElroy KL, Powers AM, Lehane MJ, Higgs S. Differential infectivities of O'nyong-nyong and Chikungunya virus isolates in anopheles gambiae and Aedes aegypti mosquitoes. Am. J. Trop. Med. Hyg. 2005; 72(5):616–621.
5. UniprotKB reviewed annotation on CHIKV Non-structural proteins: Q8JUX6 (accessed on 25.04.2012). http://www.uniprot.org/uniprot/Q8JUX6.
6. Khan AH, Morita K, Parquet MC, Hasebe F, Mathenge EGM, Igarashi A. Complete nucleotide sequence of chikungunya virus and evidence for an internal polyadenylation site. J. Gen. Virol. 2002; 83(12): 3075–3084.
7. Pruitt KD, Tatusova T, Klimke W, Maglott DR. NCBI Reference Sequences: current status, policy, and new initiatives. Nucl. Acids Res. 2009; 37(Database issue):D32-36.
8. Rice P, Longden I, Bleasby A. EMBOSS: The European Molecular Biology Open Software Suite. TIG. 2000; 16(6): 276-277.
9. Sigrist CJA, Cerutti L, de Castro E, Langendijk-Genevaux PS, Bulliard V, Bairoch A, Hulo N. PROSITE, a protein domain database for functional characterization and annotation. Nucl. Acids Res. 2010; 38(Database issue):161-166.
10. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modeling. Bioinformatics 2006; 22(2): 195-201.
11. Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. Nucl. Acids Res. 2010; 38:W545-549.
12. John B, Sali A. Comparative protein structure modeling by iterative alignment, model building and model assessment. Nucl. Acids Res. 2003; 31(14): 3982-3992.
13. Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, Tyka M, Baker D, Karplus K. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. Proteins 2009;77 Suppl 9:114-22.
14. Ramachandran GN, Ramakrishnan C, Sasisekharan V. Stereochemistry of polypeptide chain configurations. J. Mol. Biol. 1963; 7(1):95-99.
15. Structural Analysis and Verification Server (SAVES). NIH MBI Laboratory for Structural Genomics and Proteomics. http://nihserver.mbi.ucla.edu/SAVES/
16. Laskowski RA. PDBsum new things. Nucl. Acids Res. 2009; 37(Database issue):D355-359.
17. Laurie AT, Jackson RM. Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. Bioinformatics 2005; 21(9):1908-1916.
18. Do CB, Woods DA, Batzoglou S. CONTRAfold: RNA secondary structure prediction without physics-based models. Bioinformatics 2006; 22(14):e90-98.
19. Ritchie DW, Kozakov D, Vajda S. Accelerating and focusing protein–protein docking correlations using multi-dimensional rotational FFT generating functions. Bioinformatics 2008; 24(17): 1865-1873.
20. Ascalaph Graphics is a program for molecular graphics and dynamics, as well as a graphical shell to package of molecular dynamics MDynaMix. http://www.biomolecular-modeling.com/Ascalaph/Ascalaph_Graphics.html
21. Marvin Applications for chemical drawing and manipulation program, ChemAxon Ltd.
22. Kumar SP. Organic Virtual Library (ORVIL) - A Combinatorial Library Construction based on Organic Constituents and without Scaffold Hopping. Int. J. Appl. Res. Info. Tech. Comp. 2011; 2(2): 57-62.
23. Hsu KC, Chen YF, Lin SR, Yang JM. iGEMDOCK: a graphical environment of enhancing GEMDOCK using pharmacological interactions and post-screening analysis. BMC Bioinformatics 2011; 12 Suppl 1: S33.
24. Wolber G, Langer T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. J. Chem. Inf. Model. 2005; 45(1):160-169.

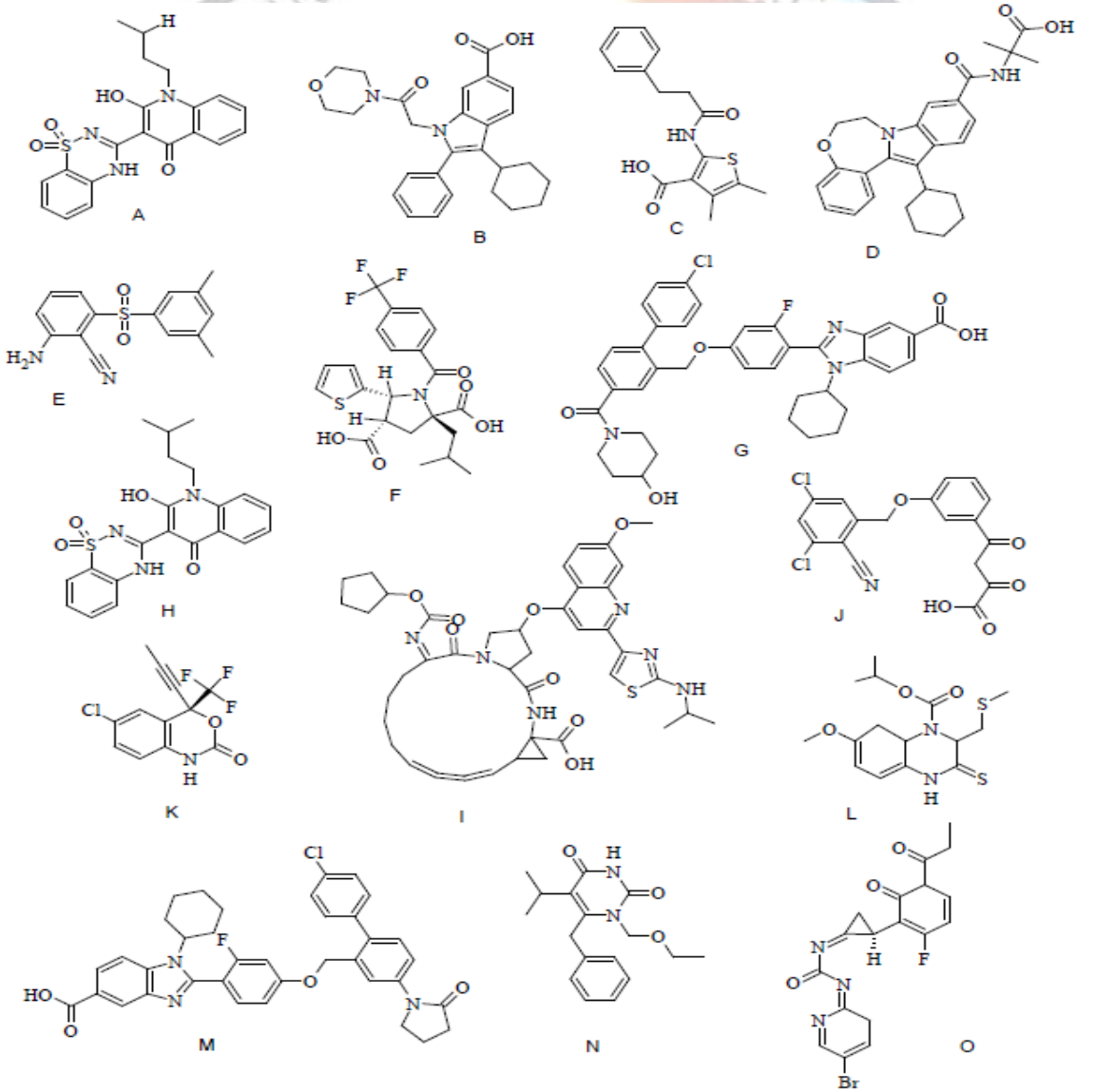
25. Fullerton SWB, Blaschke M, Coutard B, Gebhardt J, Gorbalenya A, Canard B, Tucker PA, Rohayem J. Structural and Functional Characterization of Sapovirus RNA-Dependent RNA Polymerase. J. Virol. 2007; 81(4): 1858-1871.
26. Ferrer-Orta C, Arias A, Escarmis C, Verdaguer N. A comparison of viral RNA-dependent RNA polymerases. Curr. Opin. Struct. Biol. 2006; 16(1):27–34.
27. Biswal BK, Cherney MM, Wang M, Chan L, Yannopoulos CG, Bilimoria D, Nicolas O, Bedard J, James MNG. Crystal structures of the RNA dependent RNA polymerase genotype 2a of hepatitis C virus reveal two conformations and suggest mechanisms of inhibition by non nucleoside inhibitors. J. Biol. Chem. 2005; 280(18):18202-18210.
28. Castrignano T, Meo PDD, Cozzetto D, Talamo IG, Tramontano A. The PMDB Protein Model Database. Nucl. Acids Res. 2005; 34 (Suppl 1): D306-D309.
29. Strauss JH, Strauss EG. The alphaviruses: gene expression, replication, and evolution. Microbiol. Rev. 1994; 58(3): 491-562.
30. Srinivasan P, Sudha A, Hameed AS, Kumar SP, Karthikeyan M. Screening of medicinal plant compounds against NS5B polymerase of hepatitis C virus (HCV) using molecular docking studies. J. Pharm. Res. 2011; 4(1): 136-140.

Supplementary

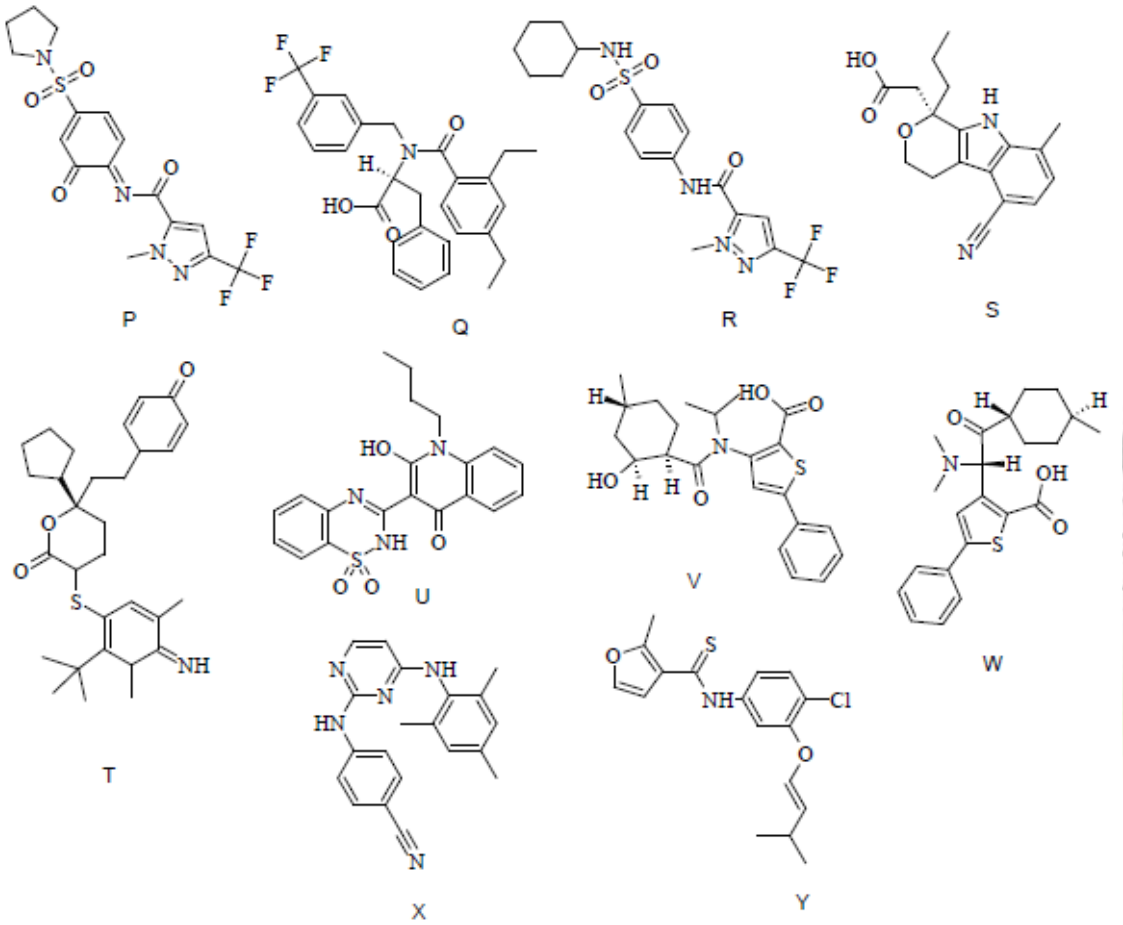


Supplementary Fig. 1: Nucleosidic inhibitors (A) 2'-C –methyl-7-deaza adenosine (MK0608), (B) 2'-C methyladenosine, (C) 2'-C methyl cytidine, (D) 2'-C methyl guanosine, (E) 2'-deoxy-2'-fluoro-2'-C methylcytidine,

(F) 2'-O-methylcytidine, (G) 3'-deoxycytidine, (H) 5-bromo-2'-deoxycytidine, (I) 5-hydroxy-2'-deoxycytidine, (I) 5- substituted cytidine analog 1, (J) 5- substituted cytidine analog 2, (K) 5-substituted cytidine analog 3, (L) 5-substituted cytidine analog 4, (M) (2R,3S,4S,5R)-5-(hydroxymethyl)-2-(4-imino-1,4-dihydro-7H-pyrrolo[2,3-d]pyrimidin-7-yl)-3-methyltetrahydrofuran-3,4-diol, (N) Bis(S-pivaloyl-2-thioethyl]phosphotriester derivative 1, (O) Bis(S-pivaloyl-2-thioethyl]phosphotriester derivative 2, (P) NITD008, (Q) NM107, (S) NM283, (T) Ribavirin and (U) GS-327073.



Supplementary Fig. 2A: Non-nucleosidic inhibitors (A) 4-hydroxy-2(1H)-quinolinone derivative, (B) 1, 3-cyclohexyl-1-(2-morpholin-4-yl-2-oxoethyl)-2-phenyl-1H-indole-6-carboxylic acid, (C) 4,5- Dialkyl analogs, (D) 6,7-dihydro-5H-benzo[5,6][1,4]diazepino[7,1-a]indoles, (E) 739W94, (F) Acyl pyrrolidine derivative, (G) Benzimidazole derivative, (H) Benzothiadiezine derivative, (I) BILN2061, (J) Diketo acid derivative, (K) Efavirenz, (L) HBY097, (M) JTK109, (N) MKC442, (O) MSC204.



Supplementary Fig. 2B: Non-nucleosidic inhibitors (continued) (P) Phenyl derivative, (Q) Phenylalanine derivative, (R) Piperidine derivative, (S) Pyranoindole derivative, (T) Pyranone derivative, (U) Thiadiazine derivative, (V) Thiophene carboxylic acid, (W) Thiophene derivative, (X) TMC120, (Y) UC781.