

RESEARCH ARTICLE

Predicting and designing therapeutics against the Nipah virus

Neeladri Sen[☉], Tejashree Rajaram Kanitkar[☉], Ankit Animesh Roy[☉], Neelesh Soni^{ORCID}, Kaustubh Amritkar, Shreyas Supekar, Sanjana Nair, Gulzar Singh, M. S. Madhusudhan^{ORCID}*

Indian Institute of Science Education and Research, Pune, India

☉ These authors contributed equally to this work.

* madhusudhan@iiserpune.ac.in



Abstract

Despite Nipah virus outbreaks having high mortality rates (>70% in Southeast Asia), there are no licensed drugs against it. In this study, we have considered all 9 Nipah proteins as potential therapeutic targets and computationally identified 4 putative peptide inhibitors (against G, F and M proteins) and 146 small molecule inhibitors (against F, G, M, N, and P proteins). The computations include extensive homology/*ab initio* modeling, peptide design and small molecule docking. An important contribution of this study is the increased structural characterization of Nipah proteins by approximately 90% of what is deposited in the PDB. In addition, we have carried out molecular dynamics simulations on all the designed protein-peptide complexes and on 13 of the top shortlisted small molecule ligands to check for stability and to estimate binding strengths. Details, including atomic coordinates of all the proteins and their ligand bound complexes, can be accessed at <http://cospi.iiserpune.ac.in/Nipah>. Our strategy was to tackle the development of therapeutics on a proteome wide scale and the lead compounds identified could be attractive starting points for drug development. To counter the threat of drug resistance, we have analysed the sequences of the viral strains from different outbreaks, to check whether they would be sensitive to the binding of the proposed inhibitors.

OPEN ACCESS

Citation: Sen N, Kanitkar TR, Roy AA, Soni N, Amritkar K, Supekar S, et al. (2019) Predicting and designing therapeutics against the Nipah virus. *PLoS Negl Trop Dis* 13(12): e0007419. <https://doi.org/10.1371/journal.pntd.0007419>

Editor: Jeanne Salje, University of Oxford, UNITED KINGDOM

Received: April 26, 2019

Accepted: November 4, 2019

Published: December 12, 2019

Copyright: © 2019 Sen et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files. The coordinates of the models of proteins and complexes with the inhibitors is publicly available at <http://cospi.iiserpune.ac.in/Nipah/>

Funding: MS Madhusudhan would like to acknowledge the Wellcome Trust-DBT India alliance for a senior fellowship. Neeladri Sen and Sanjana Nair would like to acknowledge CSIR-SPMF for funding. Kaustubh Amritkar would like to acknowledge INSIPRE-SHE fellowship. The funders

Author summary

Nipah virus infections have killed 72–86% of the infected individuals in Bangladesh and India. The infections are spread via bodily secretions of bats, pigs and other infected individuals. Even though, the disease was first detected in the human population in 1998, there are no approved drugs/vaccines against it. In this study, we have tried to model the 3D structures of the Nipah virus proteins. We have then used these models to design/predict inhibitory molecules that would bind them and prevent their function. We have also analysed the different strains of the virus to identify conservation patterns of amino acids in the proteins, which in turn informs us about the potential target sites for the drugs. The designed/docked compounds, as well as the protein models, are freely accessible for experimental validation and hypothesis testing.

had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

The May 2018 outbreak of the Nipah Virus (NiV) in Kerala, India, claimed the lives of 21 of the 23 infected people [1,2]. This zoonotic pathogen was first detected to infect humans in an outbreak in Malaysia in 1998 [3]. Since then, the mortality rate, especially in the Indian sub-continent has been high with Bangladesh and India reporting 72% and 86% fatalities respectively [4–6]. Though the overall number of fatalities linked with each outbreak has never been more than 105, NiV poses a deadly threat and could potentially become pandemic [7–9]. Considering its high mortality and transmission rates, NiV features in the WHO R&D Blueprint list of epidemic threats that need immediate R&D action [4]. In the light of this, the Coalition for Epidemic Preparedness Innovations (CEPI) has extended US\$ 25 million support to Profectus BioSciences Inc. and Emergent BioSolutions Inc. for the development of vaccines against NiV in 2018 [10]. NiV is currently classified as a Biosafety Level 4 (BSL-4) pathogen [11] with no licensed drugs or vaccines. Ribavirin and 4-Azidocytidine have been investigated as putative therapeutics against Paramyxoviruses [12,13]. However, the efficacy of ribavirin against NiV is unclear [14]. During the 1998–1999 Malaysian outbreak, it showed a 36% reduction in mortality compared to the control group [12]. The control group, however, consisted of patients who were admitted prior to the availability of ribavirin and hence did not necessarily follow the same treatment regimen which could have contributed to higher mortality. It was also administered to patients during the Kerala outbreak and as post-exposure prophylaxis to medical professionals. None of the medical personnel who were administered prophylactic ribavirin acquired the disease. The only two survivors were given ribavirin, although it is not clear how many others also received it as 6 fatalities had been reported before confirmation of disease etiology [1,14]. While ribavirin efficacy *in vivo* is uncertain, 4-azidocytidine trials against Hepatitis C Virus and Dengue Virus were halted due to low efficacy and extreme toxicity [15–17]. The drug favipiravir [18] protects against lethal doses of NiV in hamster models and is in Phase II of clinical trials (for influenza, which like NiV is a member of the Paramyxoviridae family). However, *in vitro* studies have shown the emergence of resistance to this drug among members of the influenza family [19]. A monoclonal antibody, m102.4 [20] acts against the G protein of the virus has been shown to be effective on animal models but human trials are yet to be conducted, though preliminary indications appear promising [21]. In principle, structure based rational design of therapeutics and drugs could help combat the disease and also address the concerns of drug resistance.

The NiV genome encodes six structural proteins viz. Glycoprotein (G), Fusion protein (F), Matrix protein (M), Nucleoprotein (N), RNA-directed RNA polymerase (L), Phosphoprotein (P) and three non-structural proteins named W, C and V [22]. The G protein helps in viral attachment to host cell ephrin receptors and the F protein mediates viral fusion [23–25]. The P protein binds to the N protein and maintains it in a soluble form and increases its specificity towards viral RNA instead of non-specific cellular RNA. The N-P protein complex binds the viral RNA forming the nucleocapsid [26]. This nucleocapsid coated viral RNA acts as a template for viral polymerase L to replicate itself and the host machinery is then utilized to translate its proteins [27]. After replication, the M protein homodimerizes and the dimers form arrays at the plasma membrane. These dimer-dimer interactions induce a curvature in the membrane that enables budding/release of new viral particles [28,29]. The non-structural proteins W, V, and C act against interferon signalling to escape the host immune response [30]. All these proteins are potential targets for rational drug design. Some studies in the recent past have targeted epitopes of these viral proteins [31,32]. However, to the best of our knowledge, the whole proteome modeling of NiV for drug discovery has not been attempted.

In this study, we have used the experimentally determined structures of the NiV proteins and built models for the remaining proteins in trying to find putative lead compounds against the virus. Four proteins (F, G, N and P proteins) have structural data available in the Protein Data Bank (PDB) [33] with varying degrees of structural coverage (Table 1). Using homology based methods, we have extended the structural coverage of these proteins and built models for four of the remaining proteins using either homology modeling or threading/*ab initio* methods. We designed peptide inhibitors targeting interacting sites on G protein-human ephrin-B2 receptor, F protein trimer and M protein dimer. Binding stability of inhibitory peptides was assessed with molecular dynamics (MD) simulations. In addition, to quantifying the binding affinities, binding free energies of the designed peptide inhibitors to their respective targets were also evaluated, based on conformations from MD simulations. We have predicted putative drug like molecules using molecular docking that could bind to NiV proteins. The stability of a few of our top docked protein-inhibitor complexes was evaluated based on MD simulations and binding free energy calculations. Our proposed inhibitors should potentially bind to viral proteins and hinder their function thereby preventing viral life-cycle progression. Finally, we have compared the proteomes of Malaysian, Bangladesh and Indian NiV isolates for sequence variations and mapped them onto their protein structures. This enables us to delineate the consequences (if any) of sequential variation among strains on the efficacy of proposed drugs.

Methods

Protein structure modeling

At the time of modeling, the sequence of the Indian strain was not available and so all the modeling was carried out using the Malaysian strain (AY029768.1) [34]. From our experience, using one strain over another would only minimally affect the computed models (Refer to the result section on sequence variation in NiV isolates for details). Monomeric structures of the proteins were built using the homology modeling pipeline ModPipe-2.2.0 [35,36] and their multimeric complexes were built using MODELLER v9.17 [37,38]. The templates for homology modeling were identified using both sequence-sequence and profile-sequence search methods. Profile-sequence search methods improve the identification of distant homologs that have sequence identity lower than 30%. The sequence profiles of the target proteins were generated using PSI-BLAST [36] against the UniRef90 database [39] with three iterations and an e-value threshold of 0.001. Models were built with dynamic Coulomb (electrostatic) restraints and were subjected to the ‘very slow’ mode of refinement with two rounds of optimization. The quality of the generated models was assessed using the Modpipe quality score, GA341, Discrete Optimized Protein Energy (DOPE) and Normalized DOPE scores [40]. Protein structure models were considered for further analyses only if they had a Normalized DOPE score less than or equal to zero.

Protein domains/regions that could not be reliably modeled by MODELLER (either greater than zero Normalized DOPE score or with less than 50% structural coverage) were rebuilt using meta-threading and *ab initio* methods on the I-TASSER web server [41]. Models built using I-TASSER were assessed with Normalized DOPE scores along with their C-scores, predicted TM scores and RMSD scores provided by the webserver [41].

Prediction of putative small molecules that can bind to NiV proteins

Docking was used to identify putative small molecules that can potentially bind and inhibit the activities of the NiV proteins. In this exercise, NiV proteins (G, N, F, P and M proteins) that had crystal structures or models built from templates with high identity (>90%) and high

Table 1. List of NiV proteins with their lengths, PDB codes of crystal structures along with their resolution in parenthesis, coverage of crystal structures, coverage of models, additional coverage obtained by the models and the overall sequence coverage. In cases where models have increased the coverage over existing crystal structures, the original coverage is in parentheses.

Sr. no.	Protein	Length	X-ray structures (Resolution)	X-ray coverage	Model coverage	Additional coverage	% Overall coverage
1	Pre-fusion F protein	546	5EVM (3.4Å), 1WP7 (2.2Å), 3N27 (1.8 Å)	27–482	27–482	0	84
	Post-fusion F protein		-	-	72–418	347	64
2	G protein	602	2VSM (1.8 Å), 2VWD (2.25 Å), 3D11 (2.3 Å), 3D12 (3.0 Å)	176–602	98–597	79	84 (71)
3	N protein	532	4CO6 (1.7 Å)	32–371	39–414	44	72 (64)
4	P protein	709	4CO6 (1.7 Å), 4GJW (3.0 Å), 4N5B (2.2 Å), 6EB8 (2.5 Å), 6EB9 (1.9 Å)	1–38	655–709	55	37 (29)
				471–578			
5	M protein	352	-	-	45–352	308	88
6	L protein	2244	-	-	1814–2024	210	9
7	V protein	456	-	-	1–38	297	65
					87–243		
					313–414		
8	W protein	450	-	-	1–38	266	59
					87–243		
					321–391		
9	C protein	166	-	-	-	-	-

<https://doi.org/10.1371/journal.pntd.0007419.t001>

coverage (> 80%) were used as targets for ligand screening. The screening library consisted of a 70% non-redundant set of 22,685 ligands constructed from ~13 million clean drug like molecules of the ZINC database [42,43]. The 70% library was chosen as a practical measure to ensure wide coverage. Further, we envisage that during experimental trials all structurally similar small molecules to our predicted hits would be tested. The binding pockets for docking on the targets were predicted using the DEPTH server [44,45]. The parameters of DEPTH included a minimum number of neighbourhood waters set to 4 and the probability threshold for binding site of 0.8. Evolutionary information was also included by the server in binding site predictions [45]. The druggability of the binding pocket was predicted using PockDrug [46] and CavityPlus [47], but no consensus prediction could be obtained (S8 Table). Hence the druggability of the pocket was not taken into consideration during docking. Docking was performed using Autodock4 [48], and DOCK6.8 [49,50]. The target proteins were prepared for docking by Autodock4, by adding missing polar hydrogen atoms and Gasteiger charges. The ligand docking site, marked by affinity grids were generated using the Autogrid module of Autodock. The centre of the grid, number of grid points in X, Y, and Z directions and separation of grid points were chosen based on the predicted binding pockets using the ADT viewer from MGL tools [48]. The number of Genetic Algorithm runs was set to 20. The final energies reported by Autodock4 were used for evaluation and selection of the putative leads. The target proteins were prepared for docking by DOCK6.8 using Dock Prep tool [49] from Chimera [51]. Missing hydrogen atoms were added to the target proteins using Chimera. Charges on atoms of the protein were determined using AMBER. Molecular surface of the target was generated using the DMS tool from Chimera. The sphgen program from DOCK6.8 was used to generate spheres from the molecular surface. The cluster of spheres were selected according to the binding sites predicted by DEPTH. The grid box and grid were created by showbox and

grid programs respectively. Flexible ligand docking was performed using DOCK6.8. The final energies reported by DOCK6.8 were used for evaluation and selection of the putative leads.

Assessing the stability of inhibitory peptides and small molecules against the NiV proteins

One peptide inhibitor was computationally designed against each of the F and M proteins while 2 inhibitors were designed against the G protein. Additionally, 13 small molecules were predicted with high confidence to bind different NiV proteins. Details of the procedures for modeling/predicting peptide/small molecule inhibitors are stated in the results section. MD simulations were carried out in triplicates for all four predicted protein-peptide inhibitor complexes. The simulations were carried out using GROMACS [52,53] with the Amber99SB-ILDN force field [54]. Parameters for the small molecules were generated using Antechamber [55,56]. The Amber99SB-ILDN force field has been used for the MD simulations of protein-peptide and protein-ligand complexes extensively [57–61]. In an earlier study, we used the same force field to study various protein-ligand interactions and validated one such purported complex experimentally [62]. In the cases where the small molecule ligand dissociated from the binding site, we re-simulated the system using the CHARMM27 force field [63], another popularly used molecular mechanics package. We did the second simulation to ascertain that binding was indeed weak. Parameters for the small molecules in the CHARMM27 simulations were generated using SwissParam [64].

A water box whose sides were at a minimum distance of 1.2 nm from any protein atom was used for solvating each of the systems (S15 Table). Sodium or chloride counter ions were added to achieve charge neutrality (S15 Table). Electrostatic interactions were treated using the particle mesh Ewald sum method [65] and LINCS [66] was used to constrain hydrogen bond lengths. A time step of 2 fs was used for the integration. The whole system was minimized for 5000 steps or till the maximum force was less than 1000 kJ/mol/nm. The system was then heated to 300K in an NVT ensemble simulation for 100 ps using a Berendsen thermostat [67]. The pressure was stabilized in an NPT ensemble simulation for 100 ps using a Berendsen barostat. The systems were simulated (NPT) for a maximum of 100 ns (for protein-peptide inhibitor complexes) or for 50 ns (for protein-small molecule inhibitor complexes) where pressure was regulated using the Parrinello-Rahman barostat [68]. Structures were stored after every 10ps. The temperature, potential energy and kinetic energy were monitored during the simulation to check for anomalies.

Free energy of binding of the putative peptide inhibitors/small molecules provides an important quantitative description of its efficacy. In this study, the extensive MD simulations of protein-inhibitor complexes were post-processed to obtain binding free energy estimates using the molecular mechanics Poisson-Boltzmann surface area (MM/PBSA) approach [69,70]. The MM/PBSA method employs an implicit solvation model to estimate the free energy of binding by evaluating ensemble averaged classical interaction energies (MM) and continuum solvation free energies (PBSA) of the protein-ligand complex conformations from the MD trajectories. Snapshots of protein-peptide complexes were obtained at every 100 ps from the last 50 ns of the MD trajectories, thus totalling 500 snapshots. The last 50 ns of protein-peptide inhibitors were selected for MM/PBSA treatment to ensure sampling of equilibrium conformations for appropriate MM/PBSA energy evaluations (S2–S5 Figs for the RMSD and the distance between the centre of peptide and protein). The MM/PBSA calculations of the protein-small molecule inhibitors were calculated based on the last 40 ns trajectory with snapshots obtained after every 1000 ps, totalling to 40 snapshots. The MD snapshots were energy minimized for 2000 steps before evaluation of interaction and solvation free energies.

The protein and solvent were modeled with dielectric constants of $\epsilon = 2$ and $\epsilon = 80$, respectively. APBS suite [71] and GMXPBSA [72] were used for implicit solvent calculations. In this study, we attempted to calculate the entropic estimate of binding using the interaction entropy formalism [73]. However, converged entropic values with reasonable error estimates for protein-peptide trajectories could not be obtained, which is often the case when evaluating entropic contributions from molecular simulations. We, therefore, neglected entropic contributions to the binding free energies, as estimated entropy change upon binding is often negligible and can be ignored for relative binding free energies calculations [74,75]. The enthalpies of binding obtained from MM/PBSA calculations are reported as binding energies for the protein-peptide complexes.

Mapping strain variants onto structure

Protein sequences of 15 different NiV isolates, 7 from Malaysia (AY029768.1, A J564621.1, AJ627196.1, AY029767.1, AJ564622.1, AJ564623.1, AF212302.2) [34], 3 from Bangladesh (AY988601.1, JN808857.1, AY988601.1) [76] and 5 from India (MH523641.1, MH523642.1, MH396625.1, MH523640.1, FJ513078.1) [77] were retrieved from their translated genomes deposited in the NCBI nucleotide database [78] and were used to identify sequence variations in proteins. We also verified that the translated protein sequences of the Malaysian strain matched with those of the protein sequences deposited in SwissProt [79]. Multiple sequence alignments of the sequences obtained from the 15 isolates were performed with MUSCLE [80]. Positions with amino acid variations were mapped onto the structures. Amino acid variations within 5 Å at inhibitor binding sites were identified.

Results

Structural coverage of the NiV proteome

Homology modeling the Nipah proteome. In this study, we first focused on characterizing the structures of the NiV proteins. Partial structures for 4 (F, G, N and P protein) of the 9 NiV proteins are available in the PDB (Table 1). Computationally, we attempted to extend the structural coverage of these 4 proteins and to build models for the remaining 5 proteins using homology modeling (with MODELLER), *ab initio* modeling and threading (with I-TASSER). Model accuracies were carefully scrutinized before attempting to design/predict inhibitors against all possible proteins in the proteome. In this section, we only present the results of homology modeling as all models built using I-TASSER resulted in structures that were not favourably assessed (Normalized DOPE > 0) (S1 Table)

Multiple models were constructed for each of the proteins using all available templates. All proteins, except C, had at least one model with a normalized DOPE score of less than or equal to zero. All models built for proteins with existing X-ray structure conferred additional sequence coverage except for the F protein (Table 1). The structural coverage of the N, P and G proteins increased by 8–13% after modeling (Table 1). Overall, we increased the structural coverage of the NiV proteome by 90%, from ~23% (1364 residues) to ~43% (2623 residues).

Modeling the post-fusion F protein. For NiV to enter a host's cell, its G protein binds the host ephrin receptor and the F protein is instrumental in fusing the viral envelope with the host cell membrane [24]. The F protein undergoes a conformational change from the pre-fusion to the post-fusion state triggered by the binding of the G protein to the ephrin receptor. These conformational changes are characteristic of class I viral fusion proteins [24,81–85]. The structure of only the pre-fusion state of the NiV F protein has been determined experimentally (PDB id: 5EVM). We modeled the post-fusion state using the structure of the human Parainfluenza Virus 3 (PDB id: 1ZTM) as a template since it is also a class I fusion protein. Though

the NiV and human parainfluenza virus fusion proteins are only 26.4% identical in sequence, their pre-fusion conformations take on similar folds with a structure overlap of 67% and an RMSD of 0.2 nm (as calculated using CLICK [86]). The rationale for modeling the post-fusion state of NiV using the Parainfluenza virus template is further corroborated by reports in literature of the common mode of conformational change in post-fusion states of class I viral fusion proteins [24,87–91] despite their low sequence identity (S1 Fig) leading to the formation of a 6-helix bundle. The target-template alignment was done using CLUSTALW-1.7, and the model was constructed using MODELLER v9.17. It has previously been shown that Hendra virus (HeV) and NiV infection can be inhibited by peptides derived from the heptad repeat regions of the human Parainfluenza Virus 3 [92]. This occurs as a result of the inhibition of 6 helix bundle formation, due to interactions between the native heptad repeat regions of NiV/HeV and peptide heptad repeats derived from Parainfluenza virus 3. The interaction of the heptad repeats of the human Parainfluenza Virus 3 with those of NiV/HeV along with their sequence conservation (S1 Fig) could be suggestive of similarities in the post-fusion structure of these viruses, supporting our choice of template for modeling the post-fusion conformation of the F protein.

Modeling the M protein dimer. The M protein in NiV is crucial in initiating the budding of the virus. This protein homodimerizes before homo-oligomerizing and forming the viral matrix [29]. A monomer of M protein was modeled using the HeV M protein (PDB id: 6BK6) which had a sequence identity of 94% (refer to Methods Section on protein structure modeling for details). Utilizing this monomeric structure and the crystal structure (PDB id: 4G1G) of a dimer of another Paramyxovirus, the Newcastle Disease Virus as templates, a homology model of M protein dimer was built. The target-template (Newcastle virus as template) sequence identity was 19%, going up to 27% at the interface (29 identical residues out of 70). The model was energy minimized with GROMACS using the Amber99SB-ILDN force field [54] and evaluated using our empirical knowledge based scoring scheme, PIZSA [93]. PIZSA has been benchmarked previously for its efficacy in identifying true binding interfaces [93,94]. The dimer had a PIZSA Z score of 1.69, well above the binding threshold of 1.50 (for the distance threshold of 4 Å). We also attempted to build several host-pathogen protein complexes but none of the models were evaluated favourably by FoldX [95]/PIZSA (S1 Text).

Design and stability of protein peptide inhibitor complexes

Peptide inhibitor of the post fusion F protein. Protein F contains two helical domains identified as HRA and HRB. The HRA domain forms coiled-coil trimer that associates with three helices of the HRB domain to form a 6-helix coiled-coil bundle (sometimes referred to as 6HB) [96] (Fig 1), which is essential for its fusion with the host membrane. One strategy to inhibit the formation of this 6-helix bundle hexamer (which in turn would prevent the fusion of the host and viral membranes), is to design a peptide that would competitively bind to HRA domains, preventing its binding to the HRB helices. The 6-helix bundle forming regions of HRA and HRB, have heptad repeat sequence pattern [97] of a-b-c-d-e-f-g, such that hydrophobic amino acids occupy positions a/d and charged amino acids occupy e/g positions (Fig 1A and 1B). An amino acid sequence of the inhibitor (IKKSKSYISKAQELL) was designed to mimic the HRB domain (LQQSKDYIKEAQRLL) such that all hydrophobic amino acids occupy a/d heptad positions (Fig 1C and 1D). Further, the inhibitor sequence was designed to ensure that the atomic density in the core was optimized, similar to that observed in other coiled-coil proteins. Effectively, this meant changing the N terminal Leu in HRB to Ile in the inhibitor. Other amino acid replacements were done to ensure salt bridging between the inhibitor and the HRA domain (Fig 1D). Amino acids at non a/d heptad positions of the inhibitor

repulsion causing instability of the unmodified construct. Hence Lys was modified to Asn to reduce the size and repulsive forces. The penultimate residue, Glu was modified to Ile to improve hydrophobic contacts with its neighbours on the M protein. The modified peptide RRTAGSTIN was used as a putative M protein dimerization inhibitor for further analysis. Prevention of M protein dimerization could potentially prevent the virus from budding out of cells.

Peptide inhibitors of the G protein-ephrin interaction. The NiV infection is initiated by the binding of the G protein to the ephrin receptors on the host cell [98] (PDB id: 2VSM). Inhibiting this protein-protein interaction could prevent viral entry. In this study, we have tested the feasibility of using 2 peptides to inhibit the G protein-ephrin interaction. One peptide (FSPNLW) is the part of the ephrin-B2 receptor that interacts with the G protein [99]. The other peptide (LAPHPSQ) is a part of a monoclonal antibody, m102.3, that binds [21] to both NiV and HeV. A crystal structure of the antibody bound to HeVs G protein (PDB id: 6CMG) was used as a template (79% target-template sequence identity) to construct the antibody-NiV G protein complex. 3D structural models of the speculated G protein-peptide interactions were also constructed using MODELLER v9.17.

Computational prediction of the stability of the protein-inhibitor complexes. Three independent MD simulations of 100 ns each were performed to assess the stability of each of the four protein-peptide complexes. The peptide inhibitors designed against F and M proteins bind a hydrophilic pocket while the binding interactions of the G protein to its inhibitor are predominantly hydrophobic. For each of the trajectories, the total potential energy, the distance between the centre of the protein and the peptide, RMSD and RMSF of the peptide after superimposition of protein were analysed and found to be consistent across independent runs (S2–S5 Figs and S2–S7 Tables). The F and M-peptide complexes are stabilized by hydrogen bonds. A few of them (3 and 2 hydrogen bonds in F and M complexes respectively) (S3 and S5 Tables) are retained on average in over 50% of the trajectories. Hydrogen bond analysis was not done for the G protein-peptide inhibitor complexes since their binding is mediated mainly by hydrophobic interactions and there were no stable hydrogen bonds. The protein-peptide complex was stable during the simulations as can be inferred by the peptide RMSDs, peptide RMSFs and the distances between the protein and peptide. The distance of the centre of the protein to that of the peptide fluctuated with a standard deviation of 0.03–0.09 nm (S2, S4, S6 and S7 Tables and S2–S5 Figs) around the average distance. While these measures are all indicative of tight binding, we used the trajectories to determine the binding energy of association using the MM/PBSA protocol. The inhibitors of the F and M proteins bind tightly (~110 kJ/mol) to their targets (S2, S4 and S6 and S7 Tables). However, in case of G protein inhibitors, the inhibitors FSPNLW and LAPHPSQ bind the G protein with ~-100 and ~-60 kJ/mol, respectively, suggesting that ephrin-B2 receptor based design binds 40 kJ/mol stronger. This trend is also reflected in the RMSD/RMSF values (S4 and S5 Figs).

Prediction of putative small molecules that can bind to NiV proteins

The crystal structures of the G, N, P and F proteins were used in docking studies to find plausible small molecule inhibitors. A homology model of the M protein was also included in the docking exercise as it was based on a template with high (94%) sequence identity and coverage (88%). We were conservative with the docking approach and did not use our models of the structures of the W and V proteins in this exercise. Even though V and W proteins share a large portion of their sequence with P (N-terminal 407 amino acids), there was no crystal structure of P corresponding to the identical regions of V and W proteins (except residue no 1–38, which is too small a stretch for binding site prediction). The V and W protein models

cover ~60% of the whole protein length (297 and 266 residues of a total length of 456 and 450 for V and W respectively) in discontinuous fragments, sometimes with target-template sequence identities of ~30%. (<http://cospi.iiserpune.ac.in/Nipah/>).

First, we predicted the plausible binding pockets on each of the proteins using the DEPTH server, that we had earlier benchmarked for binding site prediction accuracy [45]. A total of 12 binding pockets were predicted in G (2), N (4), P (2), F (1) and M (3) proteins (S8 Table). Two of the predicted binding pockets, one on the M protein and another on the G protein, are on the dimer interface and host protein (ephrin receptor) binding interface respectively. As mentioned in Methods section previously, these sites are important drug targets. All 12 binding sites were used to screen 22685 drug like molecules from the 70% non-redundant ZINC database of clean drug like molecules using two different docking tools, DOCK6.8 and Autodock4. The docking tools provide a docking energy score that was used to select possible high affinity binders. In the absence of an objective measure or threshold to determine strong binders, we chose the top 150 best scoring ligands for each of the pockets from both the docking tools. We then compared the two lists for common molecules. 146 molecules were identified by both Dock6.8 and Autodock4 for G (9), N (56), P (45), F (10) and M [46] proteins (S9 Table). The grid scores for the predicted complexes range between -71 to -32 units for DOCK6.8. The corresponding Autodock4 binding free energies range between -14 kcal/mol to -6 kcal/mol (S9 Table).

To corroborate our predictions, we measured the RMSD between the same ligand [in the common list] as docked by the two different tools (top 5 poses predicted by Autodock4 were compared to the top pose predicted by DOCK6.8), after superimposing the proteins. This measure is referred to as RMSD_lig. 15 unique drug like molecules had an RMSD_lig less than 0.15 nm between their docked poses.

In addition to conformational similarity, we also assessed the similarities in ligand-protein interactions, primarily hydrogen bonding (S10 Table). Further, the hydrogen bonding interactions were ~50% conserved in 9 of these complexes (with RMSD_lig < 0.15 nm). In a few instances, though the hydrogen bonding was not precisely the same, visual inspection of the complexes suggests that these bonds could be formed with small conformational changes.

Interestingly, a known drug (ZINC04829362), an antiasthmatic and antipsoriatic among other uses [100], binds to a pocket of the N protein with RMSD_lig of 0.085 nm. Another drug (ZINC12362922) used in the treatment of depression and Parkinson's disease [101] also binds the N protein with RMSD_lig < 0.15 nm.

10 drug-like molecules docked to N (5), P (4) and M (1) had an RMSD_lig of less than 0.15 nm between their docked poses and were in the top 100 scoring models as predicted by both the docking tools (S9 Fig). The molecule with the best RMSD_lig (0.074 nm) from our screening, ZINC94258558 (Fig 2A), binds the N protein (S9 Table). There are however 3 molecules (S9 Table) that are of interest despite their relatively large RMSD_lig values. The molecule ZINC91252717 is predicted as the best binder to the P protein by Autodock4 (binding energy of -14 kcal/mol) and the second best binder by DOCK6.8 (grid score of -71) (Fig 2B). These scores were among the best achieved during this docking exercise. We selected ZINC00814199 that was docked onto the M protein and was similar to ZINC01725633, which in turn formed 14 and 8 hydrogen bonds with Autodock4 and Dock6.8 respectively. ZINC00814199 was within the top 14 ranked compounds by both methods. Lastly, the hydrophobic molecule ZINC63411510 is predicted to bind the G protein on its ephrin-B2 binding interface. Though both docking methods identified this site, the docking poses were different (RMSD_lig of 0.8 nm). We hypothesize that the hydrophobic nature of the binding pocket and its size could contribute to the difference in docked poses. Note that in our list there are 3 ligands (ZINC12362922, ZINC00814199 and ZINC73641145) that (S11 Table) bind different pockets

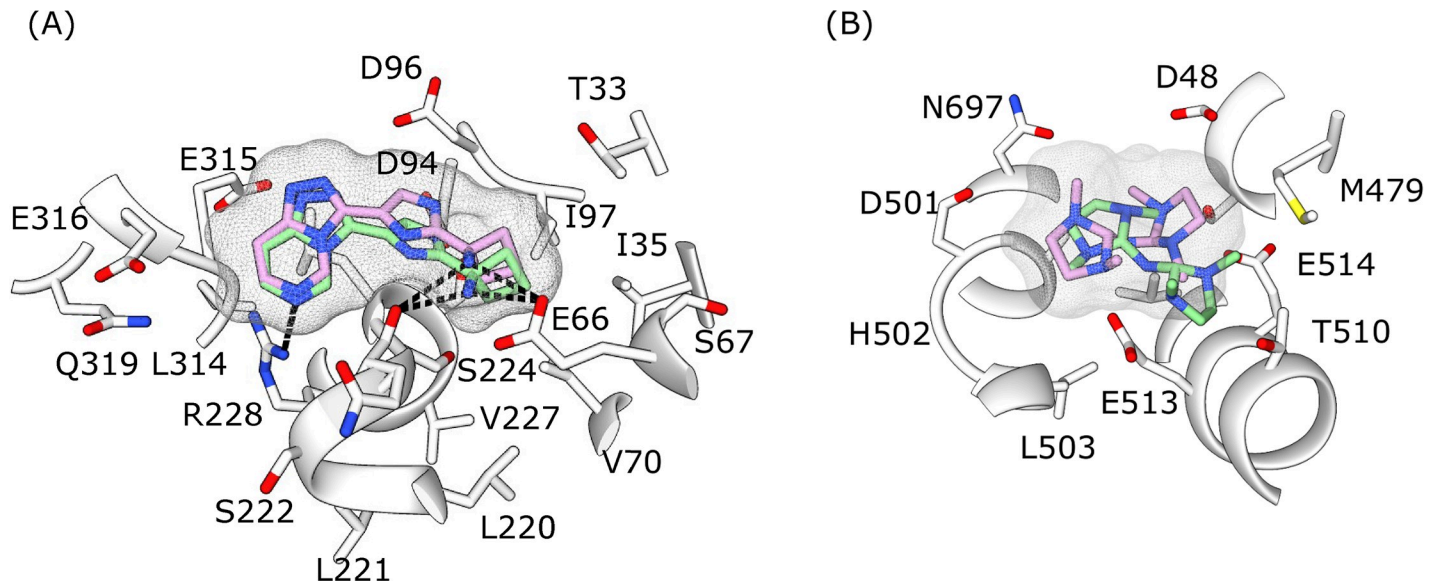


Fig 2. The docked poses of ZINC94258558 bound to N protein (A) and ZINC91252717 bound to P protein (B) as predicted by Autodock4 (green sticks with surface mesh) and Dock6.8 (lilac sticks with surface mesh). The protein is represented in white ribbons with the residues interacting with ligand shown in stick representation. Hydrogen bonds (only displayed in A) are shown as dashed lines.

<https://doi.org/10.1371/journal.pntd.0007419.g002>

on the same protein or pockets on different proteins. The ligand binding pockets (PN4 and PM2) that bind ZINC12362922 and ZINC00814199 have a similar amino acid composition containing Lys/Arg residues, Tyr residue and Leu/Val residues. The two ligands have terminal oxygens that interact with positively charged residues of the binding pocket. Another ligand ZINC73641145 binds to two different pockets on N protein (PN5 and PN4), these pockets are spatially close to one another and the ligand occupies the region between the two pockets in a similar orientation.

Computational prediction of the stability of the protein-inhibitor complexes. To assess the stability of the 13 protein-small molecule ligand complexes, we carried out three independent MD simulations of 50 ns each, using the AMBER99SB-ILDN force field (S12 and S13 Tables). 10 of the 13 ligands have RMSD_{lig} values of less than 0.15 nm (S9 Table) and were in the top 100 scoring models as predicted by both the docking tools. For these ligands the simulations were carried out starting with the DOCK6.8 predicted pose. For each of the trajectories, the distance of the centre of the small molecule ligand to the centre of the binding pocket (based on the starting structure after NPT equilibration) was monitored (S6 and S7 Figs). The triplicate MD simulations were terminated if this distance in 2 of the 3 trajectories exceeded 1 nm from its starting value and these complexes were then re-simulated using the CHARMM27 force field (this was restricted to cases where RMSD_{lig} < 0.15 nm). 5 out of 10 cases were re-simulated as a consequence. For the 3 ligands with RMSD_{lig} > 0.15 nm simulations were carried out starting with both the DOCK6.8 and Autodock4 predicted poses.

We computed binding energies for the protein-ligand complexes using MM/PBSA (as mentioned in Methods section on accessing the stability of inhibitory peptides and small molecules against the NiV proteins). 9 of the binding energies were computed to be negative in at least one of the replicates (3 for N protein, 4 for P protein, 1 for G protein and 1 for M protein). In one case (P protein-ZINC7262705 ligand), the binding energy with the CHARMM force field (after the AMBER simulation was terminated) was computed to have positive binding free

energy. In 3 cases (1 for N, P and M protein each) the ligand did not remain bound to the protein in either CHARMM or AMBER simulations (S12 and S13 Tables).

The two known drugs, ZINC04829362 and ZINC12362922 remained bound to the N protein in all 3 replicates with negative binding energies in at least 2 of the trajectories. For the important druggable site on the G protein (that recognizes the ephrin receptor on the host), the ligand remained bound in all 3 replicates when starting with the Autodock4 bound pose with negative binding energies.

Sequence variations in NiV isolates

At the time of modeling the NiV proteins, the sequence data from the 2018 outbreak was not available [77]. Hence, all the modeling was done by considering that sequence of the Malaysian strain. We rationalized that as the Malaysian and Bangladeshi/Indian strains shared a high degree (79–99%) of sequence similarity, structural models using sequences of one strain would be applicable to the other, which is the basis of comparative modeling. However, we wanted to assess whether the efficacy of the designed/proposed therapeutic molecules would be affected by observed sequence variations between the different strains (7 Malaysian, 3 Bangladeshi and 5 Indian) of NiV.

The amino acid variations (S14 Table) were mapped onto their respective structures. All protein sequences are of equal length except the V protein whose length varies between the different strains. The V and W protein have the least sequence conservation (~79%) while the M protein is the most conserved (98.6%). A general observation is that the Bangladeshi and Indian strains are more similar to one another than they are to the Malaysian sequences (S8 Fig).

We mapped the sequence variations onto all the protein structures/models that were used for peptide inhibitor design and drug docking. No variations in the sequence were found close to the peptide inhibitor binding sites on the F, M and G proteins. We found 1 (Lys236Arg), 2 (Asp188Glu, Gln211Arg), 1 (Asp252Gly) and 1 (Ile331Val) variations close to the docking sites on G, N, F and M protein respectively. All the mutations (except for Asp252Gly on F protein) on the binding site were conservative (similar physico-chemical properties and BLOSUM62 score ≥ 0) and hence we conjectured would not affect the interactions between the protein and the inhibitor. Though there is a non-conservative change (Asp252Gly) in one of the drug/inhibitor binding sites of the F protein, this position is not involved in H-bonding with the ligand. Hence the binding of the inhibitor to the protein is unlikely to be affected. Among the top 13 shortlisted ligands, ZINC04829362 and ZINC12362922 bound to N protein and ZINC63411510 bound to G protein were within 0.5 nm of the amino acids that showed variations. No single sequence variant we have studied appears to show that the drug binding would be directly affected.

Web service and database

We have archived all structures/models of NiV proteins and their inhibitor bound complexes in a consolidated database at <http://cospi.iiserpune.ac.in/Nipah>. The data at this site lists details of modeling, docking features and multiple sequence alignments (between the various NiV strains) such as template PDB code, target-template sequence identity, model quality assessment score, docking energies, docking rank and the RMSD_lig between the docking poses.

Discussions

NiV is a deadly zoonotic virus with a mortality rate of 72% and 86% in Bangladesh and India respectively. There are no approved drugs/therapeutics against NiV. The overarching aim of

this study is to computationally design inhibitors and predict small molecule drugs against NiV proteins. To design/predict therapeutic molecules to act against NiV, we characterized all of its proteins. As a part of this effort, we constructed partial models of 5 NiV proteins viz., M, L, V, W proteins along with the post fusion conformation of the F protein. The structure of the post-fusion conformation of the F protein is modeled for the first time in this study. Our model is based on the post-fusion structures of another class I fusion protein from Human Parainfluenza virus 3.

Our efforts have increased the coverage of existing structures of the G, N and P proteins (by 13%, 8% and 8% respectively) by modeling a fraction of their unresolved residues. No reliable models could be generated for the C protein. Effectively, we doubled the number of amino acids in the NiV proteome that were structurally characterized. While our aim is to use these models to predict/design inhibitors, we believe that many of our models are by themselves quite insightful. They could serve as templates for future structure-guided drug designing efforts against members of the Paramyxoviridae family. We attempted to build complexes of the viral and host protein (host cathepsin-L with NiV F protein and host AP3-B1 with NiV M protein) to target the interactions for inhibitor design. However, we were unsuccessful in making reliable models of host-pathogen protein-protein interaction complexes. With improvements to protein-protein docking methods, the quality of such models of complexes could be improved, which in turn would help in better targeting host-viral interactions.

We next used these models to design 4 peptide inhibitors against the F, M and G proteins. The inhibitor against F protein would putatively prevent the pre to post fusion transition of the F protein, a crucial step for viral entry. Our model of the post fusion conformation of the F protein was crucial in designing this inhibitor. Another inhibitor against the M protein was designed such that it would prevent the dimerization of the protein, hence preventing the budding process. The two inhibitors against the G proteins were selected such that they bind to the ephrin receptor binding pocket, preventing viral attachment to the host cell. The peptides here mimic the ephrin-B2 protein and an antibody (m102.3) that are bound at the same site. We conjectured that these peptides would competitively inhibit the G protein from binding the host ephrin receptors. All of these protein-peptide systems were subjected to triplicate runs of 100 ns MD simulations to assess interaction strengths. The distance of the centre of the inhibitor and the peptide fluctuates with a standard deviation of 0.03–0.09 nm from the mean distance, indicative of the inhibitor remaining bound in the binding pocket. The inhibitors against the F and M proteins also had stable hydrogen bond associations in the MD trajectories. Binding affinity calculations suggest that three of the designed putative inhibitors bind tightly (~100 kJ/mol) to their targets, making them promising leads against NiV proteins.

We screened a set of drug like molecules in a docking exercise to identify potential small molecule inhibitors of NiV. The screen consisted of 22685 compounds of the 70% non-redundant set of clean drug like molecules of the ZINC library. The docking onto the NiV proteins was done using two different docking programs, Autodock4 and Dock6.8. Empirically, we chose the top 150 ligands from each of the two methods and selected those that were common between them. This resulted in 146 compounds that bound the G, N, P, F and M proteins of NiV. As a more stringent test, we whittled down this list to only include those molecules that were docked in similar poses (empirically chosen RMSD of 0.15 nm or smaller) on the same binding site and were in the top 100 scored models by both docking schemes. Hence, we predicted 10 compounds that would inhibit the N (5), P (4) and M (1) proteins of NiV. In addition we also included 3 drugs to the list that did not clear the criteria explained above. These drugs include one that binds the G protein on its ephrin binding interface and two others which bind to P and M proteins. The 13 ligand bound protein complexes were subjected to triplicate MD simulations (50 ns each) to gauge the stability of the association. In 9 of the complexes, at

least one of the trajectories was evaluated to have favourable (negative) binding energy. While the simulations and the energy calculations that follow are not to be construed as indicators of binding strength, they do provide the same general trends and give pointers and/or boost our confidence in the binding efficacy of the ligand-protein complex. Only 3 of the 13 ligands consistently moved away from the original predicted binding pocket even when the simulations were repeated using a different force field. In one other case, though the protein-ligand complex remained conformationally stable throughout the course of the triplicate trajectories, our energy estimates of this interaction were unfavourable (positive energy). In the absence of experimental validation, which we seek to do next, these MD simulations serve as indicators of the viability of the ligands to bind the viral proteins.

Of the 13 ligands, two bind in interface regions, one in the M protein dimer interface and another on the ephrin receptor recognition site of protein G. When not bound to these two sites, the ability of the ligands to functionally impair the virus would only be known with experimental testing. The most important aspect of the docking study is that the molecular screen consists of known drugs or drug like compounds. The implication is that a few of our proposed inhibitors could be readily tested and repurposed. For instance, we have identified Cyclopent-1-ene-1,2-dicarboxylic acid (ZINC04829362) as an inhibitor of the NiV N protein. This compound is a known drug prescribed for antiasthmatic and antipsoriatic among other disorders. Another example is Bicyclo[2.2.1]hepta-2,5-diene-2,3-dicarboxylic acid (ZINC12362922) that we propose also inhibits the N protein, is a drug prescribed against depression and Parkinson's disease. Both these ligands have a negative binding free energy in at least 2 of the 3 replicates.

In all our computational predictions, an independent scoring scheme(s) was used to evaluate results. MD simulations were always carried out in triplicate and sometimes using different force fields. In short, we have taken care to ensure cross validation of our computations to whatever extent practically possible. We cannot overemphasize the importance of these computational predictions, especially for swift acting potent viruses as NiV where mortality rates are high.

Finally, we assessed how effective our proposed inhibitors would be against different strains of the virus and assess the risk of the virus getting drug resistant. For this, we studied 3 Bangladeshi, 7 Malaysian and 5 Indian strains and inferred the variations between the various strains from their multiple sequences alignment. Further, we investigated whether such changes would affect inhibitor binding. Here, we narrowed the changes only to those residues that were in direct contact (< 0.5 nm) from the inhibitors. We precluded the possibility of allosteric interactions. None of the residues contacting the peptide inhibitors showed any variations in their sequence. Only 5 residue positions that were involved in binding the drug like inhibitors were changed between the different strains. 4 of these changes are conservative substitutions where the nature of the mutated residue is not deemed to change the binding property of the protein to its inhibitor. Only 1 amino acid change of Asp252Gly of the F protein is a non-conservative change, however the Asp is not involved in hydrogen bonding with the ligand. We conclude that it is likely that the proposed inhibitors would be potent against all strains of NiV and other related zoonotic viruses that pose a serious epidemic threat. Computational approaches can help identify/design inhibitors that could be rapidly tested or even deployed as they may be drugs previously licensed for other uses. Our study also has connotations for related viruses such as HeV and other Paramyxoviruses. Importantly, our models and the web pages we have created could be modified to serve as a portal to study the epidemiology of the virus should there be further outbreaks.

Supporting information

S1 Table. Model quality evaluation of the protein structures built using I-TASSER web server. The best model predicted by I-TASSER (based on their C-Score) have their Normalized DOPE scores and C-scores in bold. TM-scores and RMSDs are only calculated for the best models. L protein was divided into three domains, indicated by their residue numbers in parentheses, and modeled separately.

(DOCX)

S2 Table. Mean and standard deviation of the energy, distance of the center of the inhibitor with the center of the F protein, number of hydrogen bonds between the inhibitor and the protein, RMSD of the inhibitor and the protein-peptide binding energies obtained from the three 100ns MD simulations of F protein-inhibitor complex.

(DOCX)

S3 Table. Percentage of the snapshots with hydrogen bonds between the chain D of inhibitor with chain C and E of the F protein.

(DOCX)

S4 Table. Mean and standard deviation of the energy, distance of the center of the inhibitor with the center of the M protein, number of hydrogen bonds between the inhibitor and the M protein, RMSD of the inhibitor and the protein-peptide binding energies obtained from the three 100ns MD simulations of the M protein-inhibitor complex.

(DOCX)

S5 Table. Percentage of snapshots with hydrogen bonds between chain B of the inhibitor and chain A of the M protein.

(DOCX)

S6 Table. Mean and standard deviation of the energy, distance of the center of the FSPNLW inhibitor with the center of the G protein, RMSD of the inhibitor and the protein-peptide binding energies obtained from the three 100 ns MD simulations of G protein-FSPNLW inhibitor complex.

(DOCX)

S7 Table. Mean and standard deviation of the energy, distance of the center of the LAPHPSQ inhibitor with the center of the G protein, RMSD of the inhibitor and the protein-peptide binding energies obtained from the three 100 ns MD simulations of G protein-LAPHPSQ inhibitor complex.

(DOCX)

S8 Table. List of pocket lining residues for each pocket of NiV Proteins. The residue name is followed by the residue number. The chain id has been depicted after the dot.

(DOCX)

S9 Table. List of the ranks and energy values of the small drug like molecules that were predicted in the top 150 scoring models by both DOCK6.8 and Autodock4. RMSD_1 – RMSD_5 are the RMSDs of the 5 best Autodock4 poses with the best scoring Dock6.8 pose. The least RMSD is depicted in bold. Cells highlighted in yellow have RMSDs less than 0.15 nm. Pocket number indicates pockets from Autodock4. Some of the Autodock4 pockets have been subdivided by DOCK6.8, which indicates the subsections in each pocket.

(DOCX)

S10 Table. Number of hydrogen bonds that are formed between the selected pose for DOCK6.8 and Autodock4 with the protein. Number of common hydrogen bonds indicates the number of hydrogen bonds that are common between the predicted poses of the ligand from Autodock4 and DOCK6.8.

(DOCX)

S11 Table. Same drug like molecule predicted to bind different pockets of the same or different protein. The binding pocket has been mentioned in parenthesis.

(DOCX)

S12 Table. Binding free energy as predicted using MM/PBSA calculations from molecular dynamics simulations carried out using AMBER and CHARMM force fields for 10 ligands predicted against N, P and M proteins. The binding free energies were not calculated (depicted by -) when the ligand left the binding site in at least 2 out of 3 replicates. CHARMM was only used to run molecular dynamics simulations when the ligand left the binding pocket in AMBER simulations.

(DOCX)

S13 Table. Binding free energy as predicted using MM/PBSA calculations from molecular dynamics simulations carried out using AMBER force fields for 3 ligands predicted against G, M and P proteins for both the predicted DOCK6.8 and Autodock4 poses. The binding free energies were not calculated (depicted by -) when the ligand left the binding site in at least 2 out of 3 replicates.

(DOCX)

S14 Table. The sequence variations between the 15 NiV strains. The mutations are mentioned by the residue number followed by the amino acids present in different strains.

(DOCX)

S15 Table. List containing the number of water molecules and counter ions used during MD simulations.

(DOCX)

S1 Fig. Conformational change of the human Parainfluenza virus 3 (HPIV3) fusion protein and its sequence conservation with Nipah virus (NiV) and Hendra virus (HeV). The fusion protein undergoes a large conformational change from the pre-fusion state (A, PDB id: 6MJZ) to post-fusion state (B, PDB id: 1ZTM) to form the 6 helix bundle by interactions between the HRA domain (Salmon ribbon) and HRB domain (Cyan ribbon) heptad repeat regions. (C) Alignment of the heptad repeat regions between fusion protein sequences of the three viruses (Uniprot ids—HPIV3: P06828, NiV: Q9IH63, HeV: O89342). The alignment is color coded based on ClustalX.

(TIF)

S2 Fig. A) Energy of the F protein-inhibitor complex during 100 ns of MD simulation B) Distance of the center of the inhibitor from the center of the F protein during the simulation C) Number of hydrogen bonds between the F protein-inhibitor complex during the simulations D) Plot showing the formation of hydrogen bonds between inhibitor and F protein over 100 ns trajectories. Y axis shows the 11 different hydrogen bonds identified as numbered index (S3 Table). X axis labels time instant during simulation. Each rectangular color box represents the presence of hydrogen bond for a particular run. E) Root mean square deviation (RMSD) [#] of the designed inhibitor during the simulations F) Root mean square fluctuation (RMSF) [#] of the inhibitory peptide during the simulations. Each of the simulations were run in triplicate,

with each run being color coded as red, green and blue. (# RMSD and RMSF were calculated for the inhibitor by superimposing the protein molecule)
(TIF)

S3 Fig. A) Energy of the M protein-inhibitor complex during 100 ns of MD simulation B) Distance of the center of the inhibitor from the center of the M protein during the simulation C) Number of hydrogen bonds between the M protein-inhibitor complex during the simulations D) Plot showing the formation of hydrogen bonds between inhibitor and M protein over 100 ns trajectories. Y axis shows the 8 different hydrogen bonds identified as numbered index (S5 Table). X axis labels time instant during simulation. Each rectangular color box represents the presence of hydrogen bond for a particular run. E) RMSD # of the designed inhibitor during the simulations F) RMSF # of the inhibitory peptide during the simulations. Each of the simulations were run in triplicate, with each run being color coded as red, green and blue. (# RMSD and RMSF were calculated for the inhibitor by superimposing the protein molecule)
(TIF)

S4 Fig. A) Energy of the G protein-FSPNLW inhibitor complex during 100 ns of MD simulation B) Distance of the center of the inhibitor from the center of the G protein during the simulation C) RMSD # of the designed inhibitor during the simulation D) RMSF # of the inhibitory peptide during the simulation. Each of the simulation were run in triplicate, each run being color coded as red, green and blue. (# RMSD and RMSF were calculated for the inhibitor by superimposing the protein molecule)
(TIF)

S5 Fig. A) Energy of the G protein-LAPHPQS inhibitor complex during 100 ns of MD simulation B) Distance of the center of the inhibitor from the center of the G protein during the simulation C) RMSD # of the designed inhibitor during the simulation D) RMSF # of the inhibitory peptide during the simulation. Each of the simulation were run in triplicate, each run being color coded as red, green and blue. (# RMSD and RMSF were calculated for the inhibitor by superimposing the protein molecule)
(TIF)

S6 Fig. Distance of the centre of the ligand from the centre of the binding site (calculated based on the residues within 5Å of the first snapshot after NPT equilibration) during the simulation. The identity of the ligand, force field and docking strategy used and the target protein has been indicated above each plot.
(TIF)

S7 Fig. Distance of the centre of the ligand from the centre of the binding site (calculated based on the residues within 5Å of the first snapshot after NPT equilibration) during the simulation. The identity of the ligand, force field and docking strategy used and the target protein has been indicated above each plot.
(TIF)

S8 Fig. Heatmap showing the sequence conservation between the different strains of NiV for (A) C protein (B) F protein (C) G protein (D) L protein (E) M protein (F) N protein (G) P protein (H) V protein (I) W protein. The color gradient represents sequence conservation where white indicates 100% conservation and redder shades indicate lesser sequence conservation. The labelling convention is Protein_Country_Genome-accession code.
(TIF)

S9 Fig. Structure of 12 molecules that were predicted as top hits in the study.
(TIF)

S1 Text. Modeling of host-pathogen interactions.
(DOCX)

Acknowledgments

The authors would like to acknowledge Swastik Mishra, Yogendra Ramtirtha, Kundan Kumar, Prof. G. Narahari Sastry and Prof. Raghavan Varadarajan for useful discussions.

Author Contributions

Conceptualization: M. S. Madhusudhan.

Formal analysis: Neeladri Sen, Tejashree Rajaram Kanitkar, Ankit Animesh Roy, Neelesh Soni, Kaustubh Amritkar, Shreyas Supekar, Sanjana Nair, M. S. Madhusudhan.

Funding acquisition: M. S. Madhusudhan.

Investigation: Neeladri Sen, Tejashree Rajaram Kanitkar, Ankit Animesh Roy, Neelesh Soni, Kaustubh Amritkar, Shreyas Supekar, Sanjana Nair, M. S. Madhusudhan.

Methodology: Neeladri Sen, Tejashree Rajaram Kanitkar, Ankit Animesh Roy, Neelesh Soni, Kaustubh Amritkar, Shreyas Supekar, Sanjana Nair, M. S. Madhusudhan.

Project administration: Neeladri Sen, M. S. Madhusudhan.

Software: Gulzar Singh.

Supervision: Neeladri Sen, M. S. Madhusudhan.

Visualization: Gulzar Singh.

Writing – original draft: Neeladri Sen, Tejashree Rajaram Kanitkar, Ankit Animesh Roy, Neelesh Soni, Kaustubh Amritkar, Shreyas Supekar, Sanjana Nair.

Writing – review & editing: Neeladri Sen, M. S. Madhusudhan.

References

1. Arunkumar G, Chandni R, Mourya DT, Singh SK, Sadanandan R, Sudan P, Bhargava B. Outbreak investigation of Nipah virus disease in Kerala, India, 2018. *The Journal of infectious diseases*. 2018 Oct 26; 219(12):1867–78.
2. Spiropoulou CF. Nipah Virus Outbreaks: Still Small but Extremely Lethal. *The Journal of infectious diseases*. 2019 May 24; 219(12):1855–7. <https://doi.org/10.1093/infdis/jiy611> PMID: 30365002
3. Looi LM, Chua KB. Lessons from the Nipah virus outbreak in Malaysia. *Malaysian Journal of Pathology*. 2007; 29(2):63–7. PMID: 19108397
4. WHO | Nipah R&D. WHO [Internet]. World Health Organization; 2018 [cited 2018 Nov 20]; Available from: <http://www.who.int/blueprint/priority-diseases/key-action/nipah/en/>
5. Morbidity and mortality due to Nipah or Nipah-like virus encephalitis in WHO South-East Asia Region Country: India. 2018;(May):2018. Available from: http://www.searo.who.int/entity/emerging_diseases/links/morbidity-and-mortality-nipah-sear-2001-2018.pdf
6. WHO | Nipah virus infection. WHO [Internet]. World Health Organization; 2018 [cited 2018 Nov 21]; Available from: <http://www.who.int/csr/disease/nipah/en/>
7. Thibault PA, Watkinson RE, Moreira-Soto A, Drexler JF, Lee B. Zoonotic potential of emerging Paramyxoviruses: Knowns and unknowns. In *Advances in virus research 2017 Jan 1* (Vol. 98, pp. 1–55). Academic Press. <https://doi.org/10.1016/bs.aivir.2016.12.001> PMID: 28433050

8. Simons R, Gale P, Horigan V, Snary E, Breed A. Potential for introduction of bat-borne zoonotic viruses into the EU: A review. *Viruses*. 2014; 6(5):2084–121. <https://doi.org/10.3390/v6052084> PMID: 24841385
9. Luby SP. The pandemic potential of Nipah virus. *Antiviral research*. 2013 Oct 1; 100(1):38–43. <https://doi.org/10.1016/j.antiviral.2013.07.011> PMID: 23911335
10. CEPI Awards \$25 Million Contract to Profectus BioSciences and Emergent BioSolutions to Develop Nipah Virus Vaccine | CEPI [Internet]. [cited 2018 Nov 20]. Available from: <http://cepi.net/news/cepi-awards-25-million-contract-profectus-biosciences-and-emergent-biosolutions-develop-nipah>
11. Who, Searo. Fact Sheet—Nipah Virus Infection [Internet]. [cited 2018 Nov 21]. Available from: http://www.searo.who.int/entity/emerging_diseases/links/CDS_Nipah_Virus.pdf?ua=1
12. Chong HT, Kamarulzaman A, Tan CT, Goh KJ, Thayaparan T, Kunjapan SR, Chew NK, Chua KB, Lam SK. Treatment of acute Nipah encephalitis with ribavirin. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*. 2001 Jun; 49(6):810–3.
13. Hotard AL, He B, Nichol ST, Spiropoulou CF, Lo MK. 4'-Azidocytidine (R1479) inhibits henipaviruses and other paramyxoviruses with high potency. *Antiviral research*. 2017 Aug 1; 144:147–52. <https://doi.org/10.1016/j.antiviral.2017.06.011> PMID: 28629988
14. Banerjee S, Niyas VK, Soneja M, Shibeesh AP, Basheer M, Sadanandan R, Wig N, Biswas A. First experience of ribavirin postexposure prophylaxis for Nipah virus, tried during the 2018 outbreak in Kerala, India. *Journal of Infection*. 2019 Jun 1; 78(6):491–503.
15. Nguyen NM, Tran CN, Phung LK, Duong KT, Huynh HL, Farrar J, Nguyen QT, Tran HT, Nguyen CV, Merson L, Hoang LT. A randomized, double-blind placebo controlled trial of balapiravir, a polymerase inhibitor, in adult dengue patients. *The Journal of infectious diseases*. 2012 Jul 17; 207(9):1442–50. <https://doi.org/10.1093/infdis/jis470> PMID: 22807519
16. Nelson DR, Zeuzem S, Andreone P, Ferenci P, Herring R, Jensen DM, Asselah T, Dieterich D, Foster GR, Marcellin P, Pockros PJ. Balapiravir plus peginterferon alfa-2a (40KD)/ribavirin in a randomized trial of hepatitis C genotype 1 patients. *Annals of hepatology*. 2012 Jan 15; 11(1):15–31. PMID: 22166557
17. Roberts SK, Cooksley G, Dore GJ, Robson R, Shaw D, Berns H, Hill G, Klumpp K, Najera I, Washington C. Robust antiviral activity of R1626, a novel nucleoside analog: a randomized, placebo-controlled study in patients with chronic hepatitis C. *Hepatology*. 2008 Aug; 48(2):398–406. <https://doi.org/10.1002/hep.22321> PMID: 18553458
18. Dawes BE, Kalveram B, Ikegami T, Juelich T, Smith JK, Zhang L, Park A, Lee B, Komeno T, Furuta Y, Freiberg AN. Favipiravir (T-705) protects against Nipah virus infection in the hamster model. *Scientific reports*. 2018 May 15; 8(1):7604. <https://doi.org/10.1038/s41598-018-25780-3> PMID: 29765101
19. Goldhill DH, te Velthuis AJ, Fletcher RA, Langat P, Zambon M, Lackenby A, Barclay WS. The mechanism of resistance to favipiravir in influenza. *Proceedings of the National Academy of Sciences*. 2018 Nov 6; 115(45):11613–8.
20. Bossart KN, Zhu Z, Middleton D, Klippel J, Crameri G, Bingham J, McEachern JA, Green D, Hancock TJ, Chan YP, Hickey AC. A neutralizing human monoclonal antibody protects against lethal disease in a new ferret model of acute nipah virus infection. *PLoS pathogens*. 2009 Oct 30; 5(10):e1000642. <https://doi.org/10.1371/journal.ppat.1000642> PMID: 19888339
21. Xu K, Rockx B, Xie Y, DeBuysscher BL, Fusco DL, Zhu Z, Chan YP, Xu Y, Luu T, Cer RZ, Feldmann H. Crystal structure of the Hendra virus attachment G glycoprotein bound to a potent cross-reactive neutralizing human monoclonal antibody. *PLoS pathogens*. 2013 Oct 10; 9(10):e1003684. <https://doi.org/10.1371/journal.ppat.1003684> PMID: 24130486
22. Rockx B, Winegar R, Freiberg AN. Recent progress in henipavirus research: molecular biology, genetic diversity, animal models. *Antiviral research*. 2012 Aug 1; 95(2):135–49. <https://doi.org/10.1016/j.antiviral.2012.05.008> PMID: 22643730
23. Bonaparte MI, Dimitrov AS, Bossart KN, Crameri G, Mungall BA, Bishop KA, Choudhry V, Dimitrov DS, Wang LF, Eaton BT, Broder CC. Ephrin-B2 ligand is a functional receptor for Hendra virus and Nipah virus. *Proceedings of the National Academy of Sciences*. 2005 Jul 26; 102(30):10652–7.
24. Lee B, Ataman ZA. Modes of paramyxovirus fusion: a Henipavirus perspective. *Trends in microbiology*. 2011 Aug 1; 19(8):389–99. <https://doi.org/10.1016/j.tim.2011.03.005> PMID: 21511478
25. Weis M, Maisner A. Nipah virus fusion protein: importance of the cytoplasmic tail for endosomal trafficking and bioactivity. *European journal of cell biology*. 2015 Jul 1; 94(7–9):316–22. <https://doi.org/10.1016/j.ejcb.2015.05.005> PMID: 26059400
26. Omi-Furutani M, Yoneda M, Fujita K, Ikeda F, Kai C. Novel phosphoprotein-interacting region in Nipah virus nucleocapsid protein and its involvement in viral replication. *Journal of virology*. 2010 Oct 1; 84(19):9793–9. <https://doi.org/10.1128/JVI.00339-10> PMID: 20668089

27. Jordan PC, Liu C, Raynaud P, Lo MK, Spiropoulou CF, Symons JA, Beigelman L, Deval J. Initiation, extension, and termination of RNA synthesis by a paramyxovirus polymerase. *PLoS pathogens*. 2018 Feb 9; 14(2):e1006889. <https://doi.org/10.1371/journal.ppat.1006889> PMID: 29425244
28. Battisti AJ, Meng G, Winkler DC, McGinnes LW, Plevka P, Steven AC, Morrison TG, Rossmann MG. Structure and assembly of a paramyxovirus matrix protein. *Proceedings of the National Academy of Sciences*. 2012 Aug 28; 109(35):13996–4000.
29. Watkinson RE, Lee B. Nipah virus matrix protein: expert hacker of cellular machines. *FEBS letters*. 2016 Aug 1; 590(15):2494–511. <https://doi.org/10.1002/1873-3468.12272> PMID: 27350027
30. Yoneda M, Guillaume V, Sato H, Fujita K, Georges-Courbot MC, Ikeda F, Omi M, Muto-Terao Y, Wild TF, Kai C. The nonstructural proteins of Nipah virus play a key role in pathogenicity in experimentally infected animals. *PLoS One*. 2010 Sep 15; 5(9):e12709. <https://doi.org/10.1371/journal.pone.0012709> PMID: 20856799
31. Mohammed AA, Shantier SW, Mustafa MI, Osman HK, Elmansi HE, Osman IA, Mohammed RA, Bashier FA, Elnnewery ME, Yousif EM, Mustafa MM. Epitope-based peptide vaccine against glycoprotein G of Nipah henipavirus using immunoinformatics approaches. *bioRxiv*. 2019 Jan 1:678664.
32. Kamthania M, Sharma DK. Epitope-based peptides prediction from proteome of nipah virus. *International Journal of Peptide Research and Therapeutics*. 2016 Dec 1; 22(4):465–70.
33. Berman Helen M., Westbrook John, Feng Zukang, Gilliland Gary, Bhat T. N., Weissig Helge, Shindyalov Ilya N., Bourne Philip E.. The Protein Data Bank. *Nucleic Acids Research*. 2000 Jan 1; 28(1):235–242 <https://doi.org/10.1093/nar/28.1.235> PMID: 10592235
34. Chan YP, Chua KB, Koh CL, Lim ME, Lam SK. Complete nucleotide sequences of Nipah virus isolates from Malaysia. *Journal of General Virology*. 2001 Sep 1; 82(9):2151–5.
35. Sanchez RO, Sali A. Large-scale protein structure modeling of the *Saccharomyces cerevisiae* genome. *Proceedings of the National Academy of Sciences*. 1998 Nov 10; 95(23):13597–602.
36. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*. 1997 Sep 1; 25(17):3389–402. <https://doi.org/10.1093/nar/25.17.3389> PMID: 9254694
37. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *Journal of molecular biology*. 1993 Dec 5; 234(3):779–815. <https://doi.org/10.1006/jmbi.1993.1626> PMID: 8254673
38. Webb B, Sali A. Comparative protein structure modeling using MODELLER. *Curr Protoc Bioinforma*. John Wiley & Sons, Inc.; 2006 Sept;5.6.1–5.6.37.
39. UniProt Consortium. The universal protein resource (UniProt) in 2010. *Nucleic acids research*. 2009 Oct 20; 38(suppl_1):D142–8.
40. Shen MY, Sali A. Statistical potential for assessment and prediction of protein structures. *Protein science*. 2006 Nov; 15(11):2507–24. <https://doi.org/10.1110/ps.062416606> PMID: 17075131
41. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nature methods*. 2015 Jan; 12(1):7. <https://doi.org/10.1038/nmeth.3213> PMID: 25549265
42. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: a free tool to discover chemistry for biology. *Journal of chemical information and modeling*. 2012 Jun 15; 52(7):1757–68. <https://doi.org/10.1021/ci3001277> PMID: 22587354
43. Irwin JJ, Shoichet BK. ZINC— a free database of commercially available compounds for virtual screening. *Journal of chemical information and modeling*. 2005 Jan 24; 45(1):177–82. <https://doi.org/10.1021/ci049714> PMID: 15667143
44. Tan KP, Varadarajan R, Madhusudhan MS. DEPTH: a web server to compute depth and predict small-molecule binding cavities in proteins. *Nucleic acids research*. 2011 May 14; 39(suppl_2):W242–8.
45. Tan KP, Nguyen TB, Patel S, Varadarajan R, Madhusudhan MS. Depth: a web server to compute depth, cavity sizes, detect potential small-molecule ligand-binding cavities and predict the pKa of ionizable residues in proteins. *Nucleic acids research*. 2013 Jun 12; 41(W1):W314–21.
46. Hussein HA, Borrel A, Geneix C, Petitjean M, Regad L, Camproux AC. PockDrug-Server: a new web server for predicting pocket druggability on holo and apo proteins. *Nucleic acids research*. 2015 May 8; 43(W1):W436–42. <https://doi.org/10.1093/nar/gkv462> PMID: 25956651
47. Xu Y, Wang S, Hu Q, Gao S, Ma X, Zhang W, Shen Y, Chen F, Lai L, Pei J. CavityPlus: a web server for protein cavity detection with pharmacophore modelling, allosteric site identification and covalent ligand binding ability prediction. *Nucleic acids research*. 2018 May 10; 46(W1):W374–9. <https://doi.org/10.1093/nar/gky380> PMID: 29750256

48. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry*. 2009 Dec; 30(16):2785–91. <https://doi.org/10.1002/jcc.21256> PMID: 19399780
49. Lang PT, Brozell SR, Mukherjee S, Pettersen EF, Meng EC, Thomas V, Rizzo RC, Case DA, James TL, Kuntz ID. DOCK 6: Combining techniques to model RNA–small molecule complexes. *Rna*. 2009 Jun 1; 15(6):1219–30. <https://doi.org/10.1261/rna.1563609> PMID: 19369428
50. Mukherjee S, Balias TE, Rizzo RC. Docking validation resources: protein family and ligand flexibility experiments. *Journal of chemical information and modeling*. 2010 Oct 29; 50(11):1986–2000. <https://doi.org/10.1021/ci1001982> PMID: 21033739
51. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*. 2004 Oct; 25(13):1605–12. <https://doi.org/10.1002/jcc.20084> PMID: 15264254
52. Berendsen HJ, van der Spoel D, van Drunen R. GROMACS: a message-passing parallel molecular dynamics implementation. *Computer physics communications*. 1995 Sep 2; 91(1–3):43–56.
53. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, Van Der Spoel D, Hess B. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*. 2013 Feb 13; 29(7):845–54. <https://doi.org/10.1093/bioinformatics/btt055> PMID: 23407358
54. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function, and Bioinformatics*. 2010 Jun; 78(8):1950–8.
55. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. *Journal of computational chemistry*. 2004 Jul 15; 25(9):1157–74. <https://doi.org/10.1002/jcc.20035> PMID: 15116359
56. Wang J, Wang W, Kollman PA, Case DA. Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of molecular graphics and modelling*. 2006 Oct 1; 25(2):247–60. <https://doi.org/10.1016/j.jmgm.2005.12.005> PMID: 16458552
57. Dapiaggi F, Pieraccini S, Potenza D, Vasile F, Macut H, Pellegrino S, Aliverti A, Sironi M. Computer aided design and NMR characterization of an oligopeptide targeting the Ebola virus VP24 protein. *New Journal of Chemistry*. 2017; 41(11):4308–15.
58. Maharana J, Patra MC, De BC, Sahoo BR, Behera BK, De S, Pradhan SK. Structural insights into the MDP binding and CARD–CARD interaction in zebrafish (*Danio rerio*) NOD2: a molecular dynamics approach. *Journal of Molecular Recognition*. 2014 May; 27(5):260–75. <https://doi.org/10.1002/jmr.2357> PMID: 24700593
59. Schindler CE, Chauvot de Beauchêne I, de Vries SJ, Zacharias M. Protein-protein and peptide-protein docking and refinement using ATTRACT in CAPRI. *Proteins: Structure, Function, and Bioinformatics*. 2017 Mar; 85(3):391–8.
60. Chan AH, Wereszczynski J, Amer BR, Yi SW, Jung ME, McCammon JA, Clubb RT. Discovery of *Staphylococcus aureus* Sortase A Inhibitors Using Virtual Screening and the Relaxed Complex Scheme. *Chemical biology & drug design*. 2013 Oct; 82(4):418–28.
61. Decherchi S, Berteotti A, Bottegoni G, Rocchia W, Cavalli A. The ligand binding mechanism to purine nucleoside phosphorylase elucidated via molecular dynamics and machine learning. *Nature communications*. 2015 Jan 27; 6:6155. <https://doi.org/10.1038/ncomms7155> PMID: 25625196
62. Nguyen MN, Sen N, Lin M, Joseph TL, Vaz C, Tanavde V, Way L, Hupp T, Verma CS, Madhusudhan MS. Discovering putative protein targets of small molecules: A study of the p53 activator nutlin. *Journal of chemical information and modeling*. 2019 Feb 22; 59(4):1529–46. <https://doi.org/10.1021/acs.jcim.8b00762> PMID: 30794402
63. Buck M, Bouguet-Bonnet S, Pastor RW, MacKerell AD Jr. Importance of the CMAP correction to the CHARMM22 protein force field: dynamics of hen lysozyme. *Biophysical journal*. 2006 Feb 15; 90(4):L36–8. <https://doi.org/10.1529/biophysj.105.078154> PMID: 16361340
64. Zoete V, Cuendet MA, Grosdidier A, Michielin O. SwissParam: a fast force field generation tool for small organic molecules. *Journal of computational chemistry*. 2011 Aug; 32(11):2359–68. <https://doi.org/10.1002/jcc.21816> PMID: 21541964
65. Darden T, York D, Pedersen L. Particle mesh Ewald: An N log (N) method for Ewald sums in large systems. *The Journal of chemical physics*. 1993 Jun 15; 98(12):10089–92.
66. Hess B, Bekker H, Berendsen HJ, Fraaije JG. LINCS: a linear constraint solver for molecular simulations. *Journal of computational chemistry*. 1997 Sep; 18(12):1463–72.
67. Berendsen HJ, Postma JV, van Gunsteren WF, DiNola AR, Haak JR. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*. 1984 Oct 15; 81(8):3684–90.

68. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*. 1981 Dec; 52(12):7182–90.
69. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of chemical research*. 2000 Dec 19; 33(12):889–97. <https://doi.org/10.1021/ar000033j> PMID: 11123888
70. Srinivasan J, Cheatham TE, Cieplak P, Kollman PA, Case DA. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate– DNA helices. *Journal of the American Chemical Society*. 1998 Sep 23; 120(37):9401–9.
71. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences*. 2001 Aug 28; 98(18):10037–41.
72. Paissoni C, Spiliotopoulos D, Musco G, Spitaleri A. GMXPBSA 2.1: A GROMACS tool to perform MM/PBSA and computational alanine scanning. *Computer Physics Communications*. 2015 Jan 1; 186:105–7.
73. Duan L, Liu X, Zhang JZ. Interaction entropy: a new paradigm for highly efficient and reliable computation of protein–ligand binding free energy. *Journal of the American Chemical Society*. 2016 Apr 20; 138(17):5722–8. <https://doi.org/10.1021/jacs.6b02682> PMID: 27058988
74. Yang T, Wu JC, Yan C, Wang Y, Luo R, Gonzales MB, Dalby KN, Ren P. Virtual screening using molecular simulations. *Proteins: Structure, Function, and Bioinformatics*. 2011 Jun; 79(6):1940–51.
75. Wichapong K, Alard JE, Ortega-Gomez A, Weber C, Hackeng TM, Soehnlein O, Nicolaes GA. Structure-based design of peptidic inhibitors of the interaction between CC chemokine ligand 5 (CCL5) and human neutrophil peptides 1 (HNP1). *Journal of medicinal chemistry*. 2016 Feb 23; 59(9):4289–301. <https://doi.org/10.1021/acs.jmedchem.5b01952> PMID: 26871718
76. Harcourt BH, Lowe L, Tamin A, Liu X, Bankamp B, Bowden N, Rollin PE, Comer JA, Ksiazek TG, Hosain MJ, Gurley ES. Genetic characterization of Nipah virus, Bangladesh, 2004. *Emerging infectious diseases*. 2005 Oct; 11(10):1594. <https://doi.org/10.3201/eid1110.050513> PMID: 16318702
77. Yadav PD, Shete AM, Kumar GA, Sarkale P, Sahay RR, Radhakrishnan C, Lakra R, Pardeshi P, Gupta N, Gangakhedkar RR, Rajendran VR. Nipah Virus Sequences from Humans and Bats during Nipah Outbreak, Kerala, India, 2018. *Emerging infectious diseases*. 2019 May; 25(5):1003. <https://doi.org/10.3201/eid2505.181076> PMID: 31002049
78. Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K, Chetverin V, Church DM, DiCuccio M, Edgar R, Federhen S, Feolo M. Database resources of the national center for biotechnology information. *Nucleic acids research*. 2007 Nov 27; 36(suppl_1):D13–21.
79. Boutet E, Lieberherr D, Tognolli M, Schneider M, Bairoch A. Uniprotkb/swiss-prot. In *Plant bioinformatics 2007* (pp. 89–112). Humana Press.
80. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*. 2004 Mar 1; 32(5):1792–7. <https://doi.org/10.1093/nar/gkh340> PMID: 15034147
81. Yin HS, Paterson RG, Wen X, Lamb RA, Jardetzky TS. Structure of the uncleaved ectodomain of the paramyxovirus (hPIV3) fusion protein. *Proceedings of the National Academy of Sciences*. 2005 Jun 28; 102(26):9288–93.
82. Yin HS, Wen X, Paterson RG, Lamb RA, Jardetzky TS. Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. *Nature*. 2006 Jan; 439(7072):38. <https://doi.org/10.1038/nature04322> PMID: 16397490
83. Lamb RA, Jardetzky TS. Structural basis of viral invasion: lessons from paramyxovirus F. *Current opinion in structural biology*. 2007 Aug 1; 17(4):427–36. <https://doi.org/10.1016/j.sbi.2007.08.016> PMID: 17870467
84. Bossart KN, Broder CC. Developments towards effective treatments for Nipah and Hendra virus infection. *Expert review of anti-infective therapy*. 2006 Feb 1; 4(1):43–55. <https://doi.org/10.1586/14787210.4.1.43> PMID: 16441208
85. Harrison SC. Viral membrane fusion. *Nature structural & molecular biology*. 2008 Jul; 15(7):690.
86. Nguyen MN, Sim AY, Wan Y, Madhusudhan MS, Verma C. Topology independent comparison of RNA 3D structures using the CLICK algorithm. *Nucleic acids research*. 2016 Sep 14; 45(1):e5-. <https://doi.org/10.1093/nar/gkw819> PMID: 27634929
87. Dutch RE, Jardetzky TS, Lamb RA. Virus membrane fusion proteins: biological machines that undergo a metamorphosis. *Bioscience reports*. 2000 Dec 1; 20(6):597–612. <https://doi.org/10.1023/a:1010467106305> PMID: 11426696
88. Jardetzky TS, Lamb RA. Virology: a class act. *Nature*. 2004 Jan; 427(6972):307. <https://doi.org/10.1038/427307a> PMID: 14737155

89. Colman PM, Lawrence MC. The structural biology of type I viral membrane fusion. *Nature Reviews Molecular Cell Biology*. 2003 Apr; 4(4):309. <https://doi.org/10.1038/nrm1076> PMID: 12671653
90. Lommerse JP, Price SL, Taylor R. Hydrogen bonding of carbonyl, ether, and ester oxygen atoms with alkanol hydroxyl groups. *Journal of computational chemistry*. 1997 Apr 30; 18(6):757–74.
91. Earp LJ, Delos SE, Park HE, White JM. The many mechanisms of viral membrane fusion proteins. In *Membrane trafficking in viral replication 2004* (pp. 25–66). Springer, Berlin, Heidelberg.
92. Porotto M, Carta P, Deng Y, Kellogg GE, Whitt M, Lu M, Mungall BA, Moscona A. Molecular determinants of antiviral potency of paramyxovirus entry inhibitors. *Journal of virology*. 2007 Oct 1; 81(19):10567–74. <https://doi.org/10.1128/JVI.01181-07> PMID: 17652384
93. Dhawanjewar AS, Roy AA, Madhusudhan MS. A knowledge-based scoring function to assess the stability of quaternary protein assemblies. *bioRxiv*. Cold Spring Harbor Laboratory; 2019 Feb 28;562520. Available from: <https://www.biorxiv.org/content/10.1101/562520v2>
94. Roy AA, Dhawanjewar AS, Sharma P, Singh G, Madhusudhan MS. Protein Interaction Z Score Assessment (PIZSA): an empirical scoring scheme for evaluation of protein–protein interactions. *Nucleic acids research*. 2019 May 22; 47(W1):W331–W337. <https://doi.org/10.1093/nar/gkz368> PMID: 31114890
95. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. The FoldX web server: an online force field. *Nucleic acids research*. 2005 Jul 1; 33(suppl_2):W382–8.
96. Steffen DL, Xu K, Nikolov DB, Broder CC. Henipavirus mediated membrane fusion, virus entry and targeted therapeutics. *Viruses*. 2012; 4(2):280–308. <https://doi.org/10.3390/v4020280> PMID: 22470837
97. Chambers P, Pringle CR, Easton AJ. Heptad repeat sequences are located adjacent to hydrophobic regions in several types of virus fusion glycoproteins. *Journal of General Virology*. 1990 Dec 1; 71(12):3075–80.
98. Xu K, Rajashankar KR, Chan YP, Himanen JP, Broder CC, Nikolov DB. Host cell recognition by the henipaviruses: crystal structures of the Nipah G attachment glycoprotein and its complex with ephrin-B3. *Proceedings of the National Academy of Sciences*. 2008 Jul 22; 105(29):9953–8.
99. Bowden TA, Aricescu AR, Gilbert RJ, Grimes JM, Jones EY, Stuart DI. Structural basis of Nipah and Hendra virus attachment to their cell-surface receptor ephrin-B2. *Nature structural & molecular biology*. 2008 Jun; 15(6):567.
100. Cyclopent-1-ene-1,2-dicarboxylic acid | C7H8O4—PubChem [Internet]. [cited 2019 Apr 16]. Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/324857#section=WIPO-IPC>
101. Bicyclo[2.2.1]hepta-2,5-diene-2,3-dicarboxylic acid | C9H8O4—PubChem [Internet]. [cited 2019 Apr 16]. Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/139991#section=WIPO-IPC>