CrossMark
← click for updates

RESEARCH NOTE

Illustrating and homology modeling the proteins of the Zika virus [version 1; referees: 2 approved with reservations]

Sean Ekins[1-3], John Liebler[4], Bruno J. Neves[5], Warren G. Lewis[6], Megan Coffee[7], Rachelle Bienstock[8], Christopher Southan[9], Carolina H. Andrade[5]

[1]Collaborations in Chemistry, Fuquay-Varina, NC, USA
[2]Collaborations Pharmaceuticals Inc., Fuquay-Varina, NC, USA
[3]Collaborative Drug Discovery Inc, Burlingame, CA, USA
[4]Art of the Cell, Guilford, CT, USA
[5]LabMol - Laboratory for Molecular Modeling and Drug Design, Faculty of Pharmacy, Federal University of Goias, GO, Brazil
[6]Department of Medicine, Washington University School of Medicine, St Louis, MO, USA
[7]The International Rescue Committee, New York, NY, USA
[8]RJB Computational Modeling LLC, Chapel Hill, NC, USA
[9]Centre for Integrative Physiology, University of Edinburgh, Edinburgh, UK

Abstract

The Zika virus (ZIKV) is a flavivirus of the family *Flaviviridae*, which is similar to dengue virus, yellow fever and West Nile virus. Recent outbreaks in South America, Latin America, the Caribbean and in particular Brazil have led to concern for the spread of the disease and potential to cause Guillain-Barré syndrome and microcephaly. Although ZIKV has been known of for over 60 years there is very little in the way of knowledge of the virus with few publications and no crystal structures. No antivirals have been tested against it either *in vitro* or *in vivo*. ZIKV therefore epitomizes a neglected disease. Several suggested steps have been proposed which could be taken to initiate ZIKV antiviral drug discovery using both high throughput screens as well as structure-based design based on homology models for the key proteins. We now describe preliminary homology models created for NS5, FtsJ, NS4B, NS4A, HELICc, DEXDc, peptidase S7, NS2B, NS2A, NS1, E stem, glycoprotein M, propeptide, capsid and glycoprotein E using SWISS-MODEL. Eleven out of 15 models pass our criteria for selection. While a ZIKV glycoprotein E homology model was initially described in the immature conformation as a trimer, we now describe the mature dimer conformer which allowed the construction of an illustration of the complete virion. By comparing illustrations of ZIKV based on this new homology model and the dengue virus crystal structure we propose potential differences that could be exploited for antiviral and vaccine design. The prediction of sites for glycosylation on this protein may also be useful in this regard. While we await a cryo-EM structure of ZIKV and eventual crystal structures of the individual proteins, these homology models provide the community with a starting point for structure-based design of drugs and vaccines as well as a for computational virtual screening.

Open Peer Review

Referee Status: [?] [?]

	Invited Referees	
	1	**2**
REVISED **version 2** published 01 Sep 2016		
version 1 published 03 Mar 2016	[?] report	[?] report

1 **Andras Fiser**, Albert Einstein College of Medicine USA

2 **Bruno Botta**, Sapienza University of Rome Italy

Discuss this article

Comments (0)

This article is included in the Zika & Arbovirus

Outbreaks channel.

Competing interests: S.E. works for Collaborations in Chemistry, Collaborations Pharmaceuticals, Inc. and Collaborative Drug Discovery, Inc.

Introduction

All flaviviruses are spherical and contain a genome of approximately 11kb that functions as mRNA and encodes a polyprotein that leads to 10 proteins[1]. Examples include dengue virus, yellow fever and West Nile virus[2]. The recent pandemic of ZIKV occurring in South America, Latin America, the Caribbean and in particular Brazil spread by the *Aedes* mosquito has awakened dormant interest in this flavivirus which is a mild dengue-like disease[3]. However several documented cases of Guillain-Barré syndrome and other neurologic conditions represent important complications of the disease. In recent weeks the extent of the disease has also become apparent as new discoveries and announcements are made almost daily. Though clearly we have a considerable number of significant gaps in our knowledge which need addressing[4].

The most concerning issue however is microcephaly observed in women who had ZIKV during pregnancy. There have been multiple cases of ZIKV found in fetal or newborn brain tissue that had signs of prenatal damage. The virus seems to have neurotropism in fetal brains, which may account for the presumed association between the infection and microcephaly[5,6]. The fetus in the recent case study had microcephaly with calcifications and ZIKV was found in the brain[6]. The ZIKV strain was identified as from French Polynesia (GenBank accession number KJ776791) and several polymorphisms were noted in the NS1, NS4B and FtsJ like methyltransferase regions. While the findings are not absolute proof that ZIKV causes microcephaly, the evidence from this case report strengthens the linkage[7]. Experts involved in the decision on the World Health Organization determined Public Health Emergency of International Concern (PHEIC) recommended the need for more research into the microcephaly link and need for an animal model to be developed. This group also interestingly called for open data sharing[8]. Early work 45 years ago in inoculated newborn mice showed that ZIKV had neurological effects, enlarging astroglial cells and destroying pyriform cells. At the same time virus formation within the endoplasmic reticulum was also visualized[9]. We are not aware of any studies of effects of ZIKV on human brain or brain cells. Localization of such viruses to the brain is not unusual for flaviviruses i.e. West Nile virus and this tropism may arise from viral binding to glycosaminoglycans, as has been observed for dengue virus in human microvascular endothelial cells[10]. Heparan sulfate and the C-type lectin DC-SIGN (dendritic cell-specific intercellular adhesion molecule 3-grabbing nonintegrin) are well characterized attachment structures for flaviviruses on cells. Interfering with glycan binding is one potential approach to preventing virus entry. Another is to acidify the endosome as has been demonstrated *in vitro* with chloroquine for dengue virus infection[2]. Several entry and adhesion factors, including DC-SIGN, Tyro3, and AXL as well as others, have been shown to permit ZIKV entry in human skin cells[11].

The routes for transmission of ZIKV besides mosquito are of some concern. Recent US CDC guidance to pregnant women describes precautions against sexual transmission of ZIKV[12] and that the virus can persist for up to 12 weeks[13]. Possible ZIKV transmission through blood transfusion in French Polynesia was described by detecting the virus in 3% of asymptomatic blood donors[14]. Given how widespread ZIKV has become, there is a risk of depleting the blood supply, if donation after potential virus exposure is deferred. Methods have also been developed to inactivate ZIKV in plasma using amotosalen and UVA illumination[15]. There are issues with detection of ZIKV as a false positive dengue NS1 antigen test in a traveler to Switzerland was found to have the virus later. Therefore, cross-reactivity appears to be an issue in detection[16] and this also suggests the need for better diagnostics to be developed.

Structural knowledge of the ZIKV proteins may allow us to understand exposed epitopes which will facilitate the development of specific diagnostic reagents that differentiate it from dengue and other flaviviruses. Furthermore, open sharing of the three-dimensional arrangement of viral surface proteins could allow the mapping of potential neutralizing epitopes, guiding efforts to rationally design effective vaccines. We recently developed a preliminary model for ZIKV glycoprotein E based on the dengue virus glycoprotein E late stage fusion intermediate as a trimer[4]. We now provide homology models of the glycoprotein E based on a dimer structure as well as attempts at modeling the other proteins in ZIKV. We have also investigated the likely glycosylation sites of the ZIKV envelope glycoprotein. Glycosylation may obstruct the binding of antibodies, or block access to potential underlying peptide antigens, so glycans may be an important consideration in diagnostic and vaccine development. Previously, glycosylation analysis using several computational tools predicted mammalian *N*-linked glycosylation at Asn-154 in most ZIKV strains[17], a site known to be important in other flaviviruses. *N*-glycosylation of the dengue envelope glycoprotein at two sites has been shown to mediate interactions with DC-SIGN[18]. Insect cell expression of dengue shows both high-mannose and complex glycans[19]. However, flavivirus glycosylation in model systems such as insect cell culture and mammalian tumor cell lines may not represent the true infective insect and mammalian glycoprotein. Mammalian *O*-linked glycosylation was predicted at Thr-245 and Thr-381 in some isolates[17]. Yet, reliable prediction is hampered by a lack of *O*-glycosylation consensus sequences and preferences for *O*-glycosylation driven by structural characteristics[20].

Understanding the three-dimensional structure of antigenic ZIKV proteins may help accelerate the development of antibodies for diagnostics and rationally designed vaccines. In addition, the comparison of the assembled surface glycoprotein of ZIKV with that of dengue virus may help understand the accessible epitopes for the development of anti-flaviral vaccines in general. There is considerable prior work including structure-based design and virtual screening for dengue, yellow fever and other related flaviviruses to develop antivirals to target envelope glycoproteins[21–26], guanylyltransferase[27], capsid protein NS3 helicase, NS2B-NS3 protease and NS5 polymerase[28,29], as well as whole cell screens[30] which have produced many molecules potentially useful against ZIKV *in vitro*. Early work has only tested a small number of FDA-approved drugs against ZIKV including (EC$_{50}$ in parenthesis) interferon (34.3 IU/ml), ribavirin (143 ug/ml), 6-azauridine (1.5 ug/ml) and glycyrrhizin (384 ug/ml)[31]. A recent paper by Hamel *et al.*, from 2015 also showed interferon inhibited ZIKV replication in primary skin fibroblasts[11].

However, the use of compounds against ZIKV should take into account the treatment of pregnant women, and many of the potential options are unsuitable for use in pregnancy because of toxicity and/or teratogenicity. Despite limited human data, the available data in animal models suggests caution. Azauridine is highly toxic to the fetus in model systems (for example 32). Ribavirin is not recommended for use in pregnancy due to embryotoxic and teratogenic effects[33]. Interferon is a potential abortifacient[34]. We also need to consider the treatment of fetus as well as children (that might become infected after birth) and the relatively small subsection of FDA approved drugs that are approved for pediatric use[35]. Therefore, alternative potential drugs for ZIKV are needed. The risks of medication use in pregnancy are notable. In particular, these include teratogenicity concerns. There is also the issue that the disease does not usually pose a direct risk to pregnant women themselves, so it's important any drug, which will not be improving their own health, does not damage their health. Pregnancy can also create increased risks and liver problems, a particular concern with any new drug as it can also affect drug distribution. In pregnancy, a ZIKV infection at 13 weeks gestation was coupled with persistent virus in a fetus at 32 weeks[6]. Treatment of symptomatic pregnant women may reduce risks of transmission to the fetus. A potential drug for ZIKV could also protect fetuses from damage by reducing transmission in the general population. If the drug appeared to reduce the duration of symptoms (which though mild can be annoying) and in turn reduced viral load, reducing the chance of transmission, this could benefit them. For example, cholera patients are often given antibiotics to reduce transmission. Also those with the flu are prescribed Tamiflu/Oseltamivir to reduce the duration of symptoms, minimize the severity of symptoms, but its use might also reduce transmission of flu in the general population. Since ZIKV has been found in semen many weeks after symptoms resolve[12], treatment of male partners may reduce viral load and reduce the long term risk of transmission.

To help accelerate drug discovery through computational analysis, we have now developed homology models of ZIKV proteins that may serve as potential drug and vaccine targets. To complement high-throughput screening efforts, we could perform virtual screening against the proteins in ZIKV. While there are crystal structures for proteins from dengue[36,37], yellow fever, West Nile virus and other flaviviruses[38–44] there are (to date) none for ZIKV. Therefore we are limited to generating homology models, although the close evolutionary relationships between flavivirus and their component proteins and genomes represents a valid approach[45].

Methods
Protein sequence alignment
As a prelude to modeling we assessed what 3D structural data had significant identity scores to a representative ZIKV polyprotein. For this we chose UniProtKB Q32ZE1_9FLAV. While we have requested the promotion of this entry to the Swiss-Prot expert review level it has been selected as the representative sequence for the UniRef90_Q32ZE1 entry that currently clusters 108 ZIKV individual sequence entries at 90% (or above) amino acid identity. We then performed a BLAST search of this against Protein Data Bank (PDB) sequence entries. Protein BLAST analysis was also

performed for each ZIKV protein sequence[46] to identify the closest proteins[47] and understand potential evolution.

Homology modeling
The amino acid sequences of ZIKV strain (GenBank accession number KJ776791[48]) were retrieved from the GenBank database[49] and used as targets for homology modelling using the SWISS-MODEL server[50,51]. The latter performed the target-template sequence alignment after searching the putative X-ray template proteins in PDB for generating the 3D models for all target sequences. The best homology models were selected according to Global Model Quality Estimation (GMQE) and QMEAN statistical parameters. GMQE is a quality estimation which combines properties from the target-template alignment. The quality estimate ranges between 0 and 1 with higher values for better models. QMEAN4 scoring function consisting of a linear combination of four structural descriptors as described elsewhere in more detail[52,53]. The pseudo energies returned from the four descriptors are related to what we would expect from high resolution X-ray structures of similar size using a Z-score scheme. Further, built models were exported to the SAVES server Version 4[54] and their overall stereochemical quality, including backbone torsional angles through the Ramachandran plot, was checked according to PROCHECK[55]. Lastly, each model was refined by an energy minimization protocol, using the Smart Minimizer algorithm in Discovery Studio version 4.1 (Biovia, San Diego, CA).

Site of glycosylation prediction
Mammalian *N*-glycosylation sites were predicted for glycoprotein E by submitting the sequence to web-based tools namely N-GlycoSite[56], GlycoEP[57,58] and NetNGlyc Version 1.0[59].

Illustration for Zika virion and animation
The Zika virion illustrations were created by combining the homology model of the envelope ZIKV glycoprotein E with the symmetry data from the dengue virus envelope. PDB ID:1K4R[60] contains the coordinates for three copies of the protein subunit of the dengue virus envelope, along with the symmetry data necessary to create the 180-subunit icosahedral structure of the complete viral envelope. The PyMOL Molecular Graphics System, Version 1.7.6.0. Schrödinger, LLC. was used to export the surface models of the three proteins in .obj format. Then they were imported into Lightwave 3D (NewTek, San Antonio, TX) where the symmetry data was used to instance copies of the model into the icosahedral envelope. The entire structure was copied several times and lighting applied as a surfacing effect to create a visually pleasing composition, and the image rendered out.

The next step was to import into Pymol the homology model of the ZIKV envelope protein which was homology modeled using PDB ID: 3P54 (from Japanese Encephalitis Virus) as a template. A surface model of this protein was exported from Pymol as an .obj, and imported into Lightwave in place of the dengue model, using the same symmetry operators to create the envelope array. Everything else about the picture was left the same (color, composition, lighting, etc...) to allow the structural differences to be more apparent, and that image rendered out as well.

The last step was to overlay the detailed area of the two images and create an animated gif to flip back and forth between the two images of ZIKV and dengue, again to allow the differences to be more clearly seen. The structure of the Zika virion could be explored in a similar manner, using known data from other flaviviruses as a guide.

ZIKV glycoprotein E homology model conformation comparison
The immature[4] and mature (this study) homology models for glycoprotein E were compared using the 'align and superimpose' proteins protocol in Discovery Studio Version 4.1 (Biovia, San Diego, CA).

Results
Sequence alignment across flaviviruses
A BLAST search of ZIKV polyprotein against PDB sequence entries shows the highest scoring matches with 55–70% sequence identity (Supplementary material S1). Protein BLAST analysis of the individual ZIKV protein sequences show that many of the

proteins are similar to the same protein from Spondweni virus in 12 out of 15 cases (Table 1). Exceptions were: FtsJ which was closer to the Murray Valley encephalitis virus, NS1 which was more similar to dengue virus 3, and glycoprotein E which was closest to the dengue virus 1 protein. These results are in general accordance with whole sequence analyses[45].

Homology modeling
The SWISS-MODEL server was used to generate alignments (Supplementary material S2) and homology models for all Zika proteins (Table 2, Figure 1, Supplementary material S3). First, we selected suitable template protein structures in PDB, observing the following criteria: the template should have a high coverage (i.e., > 65% of target aligned to template) and sequence identity >30%. Then, we used GMQE and QMEAN4 scoring function as an initial criteria to discriminate good from bad models. Acceptable alignment values and higher GMQE and QMEAN4 scores were obtained during modeling, suggesting statistically acceptable homology models were generated for 11 proteins: NS5, FtsJ, HELICc, DEXDc, peptidase S7, NS1, E stem, glycoprotein M, propeptide,

Table 1. Protein BLAST search results - closest non-ZIKV proteins.

| Protein	FASTA	Closest sequence	Coverage	E value	Identity (%)
NS5	AHZ13508.1 (2,772..3,412)	RNA-dependent RNA polymerase NS5 [Spondweni virus]	100	0	77
FtsJ	AHZ13508.1 (2,575..2,746)	Chain A, Crystal Structure Of The Murray Valley Encephalitis Virus Ns5 2'-O Methyltransferase Domain In Complex With Sah (Monoclinic Form 1)	99	1e91	76
NS4A	AHZ13508.1 (2,124..2,267)	polyprotein [Spondweni virus]	100	3e68	75
HELICc	AHZ13508.1 (1,859..1,975)	nonstructural protein NS3 [Spondweni virus]	100	5e61	86
Peptidase S	AHZ13508.1 (1,520..1,670)	nonstructural protein NS3 [Spondweni virus]	99	2e81	82
DEXDc	AHZ13508.1 (1,683..1,828)	nonstructural protein NS3 [Spondweni virus]	99	2e79	83
NS2B	AHZ13508.1 (1,376..1,502)	nonstructural protein NS2B [Spondweni virus]	100	1e56	76
NS2A	AHZ13508.1 (1,158..1,372)	nonstructural protein NS2A [Spondweni virus]	100	1e64	55
NS1	AHZ13508.1 (796..1,148)	polyprotein [Dengue virus 3]	100	2e136	56
E stem	AHZ13508.1 (698..794)	polyprotein [Spondweni virus]	100	1e38	67
Glycoprotein M	AHZ13508.1 (216..290)	membrane glycoprotein m [Spondweni virus]	100	1e31	67
Propeptide	AHZ13508.1 (126..214)	protein pr [Spondweni virus]	100	2e35	61
Capsid	AHZ13508.1 (6..122)	anchored capsid protein C [Spondweni virus]	96	4e41	68
Glycoprotein E	AHZ13508.1 (291..592) AHZ13508.1 (601..693)	envelope protein E [Dengue virus 1]	100	3e165	57
NS4B	AHZ13508.1 (2,270..2,514)	nonstructural protein NS4B [Spondweni virus]	100	2e130	82

Table 2. Summary of ZIKV homology model statistics. The global and per-residue model quality has been assessed using the QMEAN scoring function[53]. For improved performance, weights of the individual QMEAN terms have been trained specifically for SWISS-MODEL[50,51,71–73]. GMQE = Global Model Quality Estimation, QMEAN4 is a scoring function consisting of a linear combination of four structural descriptors as described elsewhere in more detail[52,53].

Protein	Coverage	Sequence Identity	GMQE	MEAN4	PROCHECK analysis			
					Most favored regions	Additional allowed regions	Generously allowed regions	Disallowed regions
NS5	100%	53%	0.89	-2.80	66.3%	30.9%	2.3%	0.9%
FtsJ	99%	54%	0.94	-1.23	69.5%	27.7%	2.1%	0.7%
HELICc	100%	55%	0.94	-1.60	69.3%	27.7%	2.0%	1.0%
DEXDc	100%	54%	0.93	-0.92	70.3%	27.3%	1.6%	0.8%
Peptidase S7	100%	53%	0.92	-0.25	67.5%	28.3%	4.2%	0.0%
NS1	99%	48%	0.77	-4.09	58.4%	36.3%	3.3%	2.0%
E Stem	100%	46%	0.74	-8.02	62.5%	27.5%	6.3%	3.8%
Glycoprotein M	100%	40%	0.73	-6.90	60.6%	28.8%	4.5%	6.1%
Propeptide	87%	47%	0.74	-0.70	53.7%	43.3%	1.5%	1.5%
Capsid	65%	42%	0.50	-3.54	69.5%	27.1%	3.4%	0.0%
Glycoprotein E	99%	47%	0.81	-3.76	63.5%	31.3%	3.7%	1.4%
NS4A	20%	29%	0.07	-1.75	—	—	—	—
NS2B	37%	48%	0.21	-0.49	—	—	—	—
NS2A	16%	31%	0.04	-3.10	—	—	—	—
NS4B	14%	30%	0.03	-2.81	—	—	—	—

capsid, and glycoprotein E (Table 2, Figure 1 and Figure 2). The Ramachandran plots for these 11 proteins provide further evidence of their acceptability (Figure 2). On the other hand, because of low GMQE scores and of low coverage observed in X-ray template proteins available in the PDB, homology models for NS4B, NS4A, NS2B, and NS2A proteins appeared to have limitations regarding active sites and epitopes and they could not be validated.

The best NS5 homology model was built using the full-length Japanese encephalitis virus NS5 as a template (PDB ID: 4K6M)[61]. The homology model generated for the FtsJ protein was built using the crystal structure of the West Nile virus methyltransferase (PDB ID: 2OY0)[62]. The best HELICc model was built using dengue virus helicase/nucleoside triphosphatase catalytic domain (PDB ID: 2BHR)[63] whereas the DEXDc model was built using the structure of the Murray Valley encephalitis virus RNA helicase (PDB ID: 2V8O)[64]. The peptidase S7 model was built using the West Nile virus Ns2B-Ns3 protease (PDB ID: 2YOL)[65] and the NS1 model was built using West Nile virus non-structural protein 1 as template (PDB ID: 4O6D)[66]. In addition, the E stem model was built using the cryo-electron microscopy (cryo-EM) structure of dengue virus capsid protein heterotetramer (PDB ID: 3J2P)[67] whereas the glycoprotein M model was built using the cryo-EM structure of dengue virus as a template (PDB ID: 3J27)[67]. The propeptide model

was built using the crystal structure of the precursor membrane protein-envelope protein heterodimer from the dengue 2 virus at low pH (PDB ID: 3C5X)[68] and the capsid model was built using the core (C) protein from West Nile virus, subtype Kunjin (PDB ID: 1SFK)[69]. Finally, the best glycoprotein E model for the mature protein was built using the following sequence taken from the polyprotein, where the part corresponding to E is from residues 291-592, while the IG-like domain III is from residues 601-693. The model used the crystal structure of the Japanese encephalitis virus envelope protein, strain SA-14-14-2 as a template (PDB ID: 3P54)[70]. After building of homology models, we performed an additional validation in order to explore stereochemical quality of dihedral angles phi against psi of amino acid residues in modeled structures and identify sterically allowed regions for these angles using PROCHECK analysis. The results shown in Table 2 and Figure 2 reveal that 58.4−70.3% residues of the modeled proteins are within the most favored regions (red), 27.1−43.3% residues of modeled proteins are within the additional allowed regions (yellow), 1.5−6.3% residues of modeled proteins are within the generously allowed regions (beige), and only 0.0−6.1% residues of modeled proteins are within the disallowed regions (white). These results showed that the overall stereochemical properties of the generated models were highly reliable and the models could be useful to future molecular modeling studies.
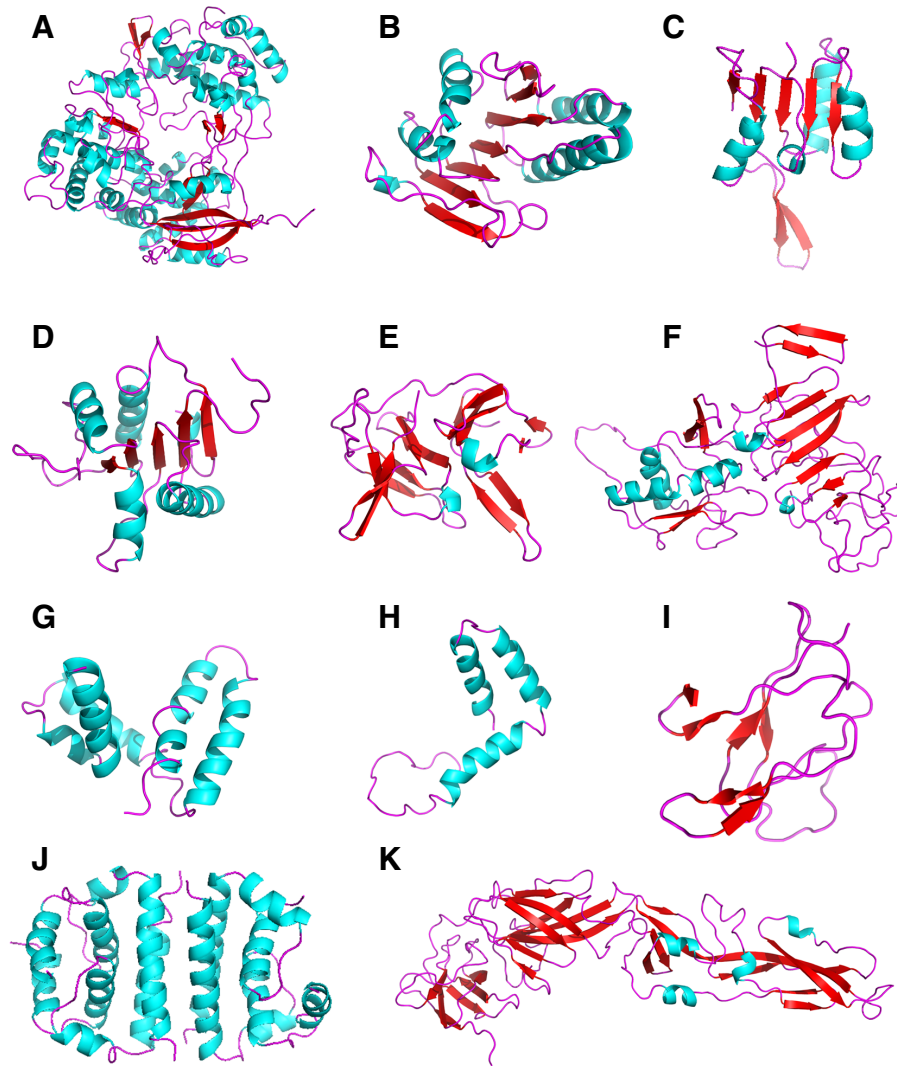
Figure 1. Selected ZIKV NS5 (**A**), FtsJ (**B**), HELICc (**C**), DEXDc (**D**), Peptidase S7 (**E**), NS1 (**F**), E Stem (**G**), Glycoprotein M (**H**), Propeptide (**I**), Capsid (**J**), and Glycoprotein E (**K**) homology models (minimized proteins) that had good sequence coverage with template proteins developed with SWISS-MODEL.

Site of glycosylation prediction

Several web-based tools were used for *N*-glycosylation site predictions as it provides a more thorough approach. N-GlycoSite[56] suggested N154 as a single N-glycosylation site matching the N-X-S/T/C consensus sequence. The same site was identified by GlycoEP using BPP settings (binary profile of patterns)[57,58] giving a score of 0.65/1.00. NetNGlyc[59] also gave the same predicted site, with a jury agreement of 6/9.

Zika virion compared to dengue virion

A qualitative analysis of the Zika virion (which was constructed based on the dengue virion) can be compared to the dengue cryo-EM virion (Figure 3) and indicates that Zika appears to have slightly more raised 'pimples' on the surface. The glycoprotein E dimer in ZIKV also has a narrow 'letter-box' groove while the dengue virion has a bigger 'pore' between the intersection of 5 dimers (5 fold axis). These differences are considerably more apparent in the animation (Supplementary material S4). It is important to note that the differences may also be artefacts of the homology modeling approach and template used for modeling ZIKV glycoprotein E.

ZIKV glycoprotein E homology model conformation comparison

The homology models developed using two different templates namely the immature protein which was based on the dengue crystal structure 4gsx as a template[50,71–73] and the mature protein which was based on PDB ID:3P54 from Japanese encephalitis
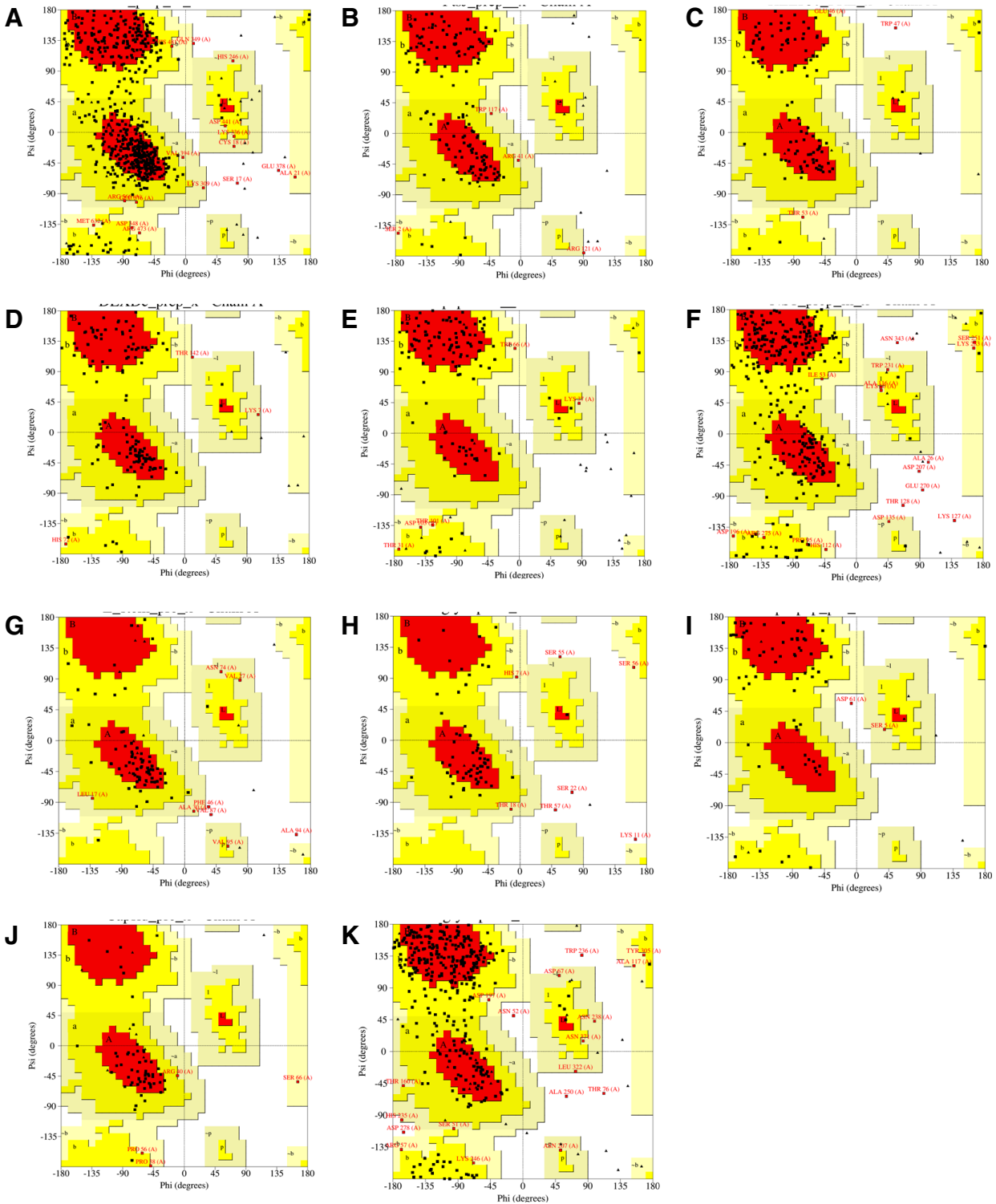
Figure 2. Ramachandran plots for ZIKV NS5 (**A**), FtsJ (**B**), HELICc (**C**), DEXDc (**D**), Peptidase S7 (**E**), NS1 (**F**), E Stem (**G**), Glycoprotein M (**H**), Propeptide (**I**), Capsid (**J**), and Glycoprotein E (**K**) obtained by PROCHECK, showing the dihedral angles Psi and Phi of amino acid residues. Red represents most favored regions; yellow represents additional allowed regions; beige represents generously allowed regions; and white areas are disallowed regions.
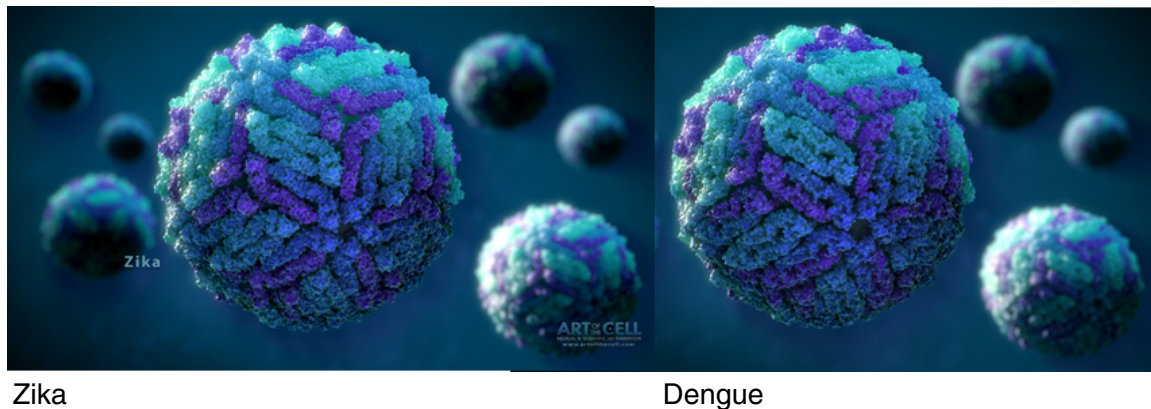
Figure 3. Comparison of Zika and dengue virion illustrations.

virus showed a large difference (RMSD 13.47Å) (Figure 4). These proteins also demonstrate differences around the pocket used centered on the residues 270-277.

Discussion

The genus Flavivirus consists of 70 viruses many of which can cause severe human disease. There have been few sequence analyses of ZIKV previously in comparison to other flaviviruses. The genus Flavivirus produces a monophyletic tree with ZIKV being closest to Spondweni virus[74] while mosquito borne, tick borne and no-vector viruses cluster separately[45]. A BLAST analysis of all the ZIKV proteins in this study suggests for 12 of 15, their closest protein is in Spondweni virus (Table 1). More often strain sequences are compared within ZIKV and these showed variations in the NS5 gene[75] and glycoprotein E[17]. This is important as it would suggest perhaps targeting other proteins would have less issue with resistance or variability due to the strain of ZIKV.

If we are to address ZIKV in the short term while we await a vaccine we need to rapidly identify an antiviral, and preferably one that can be used against other related flaviviruses. Ideally we would need to treat pregnant women and provide them with prophylaxis that was safe to them and their fetus. Such an antiviral could also be used to reduce transmission in the population in general (by reducing viral load and symptoms and/or duration). As noted a decade ago and is still is true today, no antiviral drug is approved for any flavivirus to date[76]. It has been suggested that one of the ways to target these viruses is to interfere with the NS2B/NS3 protease complex[76]. Understanding of flavivirus proteins and other RNA viruses has benefited from the EU funded project VIZIER[77], in particular several West Nile virus, dengue virus and other flavivirus structures of NS3 or NS5 were solved during this project and allosteric inhibitor sites were identified on NS5[78]. Multiple pharmaceutical companies have worked on this target for HCV leading to clinical candidates like IDX320[79], danoprevir (ITMN-191/R7227)[80], GS-9256[81] and others[82,83]. The only HCV protease targeting FDA approved drug is simeprevir, TMC435[84,85] and its use is avoided in pregnancy. Other HCV protease compounds are in clinical trials or submitted for FDA approval including Ledipasvir (formerly GS-5885)[86]. Testing these molecules against ZIKV *in vitro* would be useful.
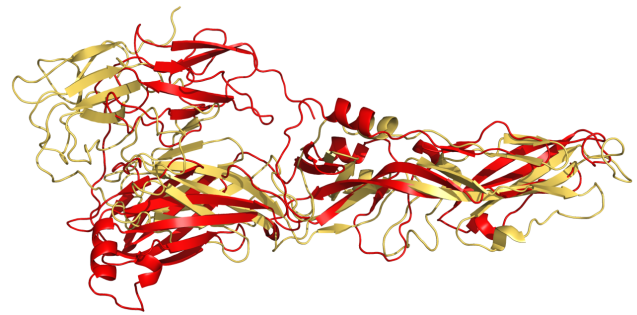
Figure 4. Overlap of ZIKV homology models for glycoprotein E, Yellow = mature conformation (this study) compared with the immature conformation (red)[4].

We recently described 6 steps which could be taken to kick start research on ZIKV[4], one of which was to develop homology models for ZIKV proteins that are similar to those targeted by molecules that are also active against the dengue virus. Such an approach would then enable docking of compound libraries of known antivirals, FDA approved drugs or other compounds[4]. Ideally generating homology models with a single tool may not be enough. In particular, for those proteins with low sequence identity the use of servers and methods that use threading may be worthwhile (e.g. I-TASSER[87–89]). However these methods are generally only accessible to academics while others are required to license the technologies. This is ironic as these technologies were developed in most cases with NIH and NSF funds. An alternative commercial homology modeling approach (MODELLER) was also used and generated a NS5 homology model and the top hit was also the Japanese encephalitis virus RdRp domain (PDB ID: 4HDH) compared with PDB ID: 4K6M from SWISS-MODEL[61]. 4HDH also includes the ATP and zinc metal where the catalytic centers are. The dengue virus 3 polymerase (PDB ID: 4HHJ)[90] has very high sequence homology and comes up as a potential target in MODELLER, which illustrates that all these viral RNA dependent polymerases are very similar.

While it is likely that the eventual availability of crystal structures of ZIKV proteins would improve the results of docking, the homology

models described here (Figure 1, Figure 2, Supplementary material S2) represent a starting point that can be used to help prioritize compounds for testing as described previously[4]. Proteins with templates above 25–40% sequence identity might suggest the proteins are related while below this is a twilight zone. Homology modeling is thought to fill in the gaps between proteins with x-ray structures and those with none[91]. Experimental testing of homology models and crystal structures indicate that a similar enrichment rate can be achieved when identifying active compounds in a set decoys[92]. Others have also described homology models that may be an excellent alternative when crystal structures are unavailable for human GPCRs[93,94], and have led to the first identification of inhibitors of the *Mycobacterium tuberculosis* Topoisomerase I after virtual screening[95,96] prior to the crystal structure becoming available[97]. Certainly there are still considerable challenges using homology models such as prediction of the correct binding pose[98] but there are plenty of success stories[98–100]. While databases of homology models exist like MODBASE[101] and SWISS-MODEL[50,51,71–73] neither of these have any ZIKV protein homology models at the time of writing. There are many structural genomics initiatives and yet it would seem there are few if any continuing the work of VIZIER working on flaviviruses or emerging viruses.

Availability of structures are important as the structure of the ZIKV glycoprotein could be useful for design of antibodies selective for the virus which will be critical for the development of diagnostics, and understanding antibody binding also for the use of IV immunoglobulin in pregnancy and the organization of the epitopes on viral proteins may facilitate early work in vaccine development. There are further implications for understanding the antibody binding epitopes, which are sometimes shared between different flaviviruses. Broadly protective vaccines for flaviviruses may allow the simultaneous targeting of ZIKV and related viruses such as dengue[102]. Understanding glycosylation is therefore important. To date Asn-154 is mentioned in Faye *et al.*,[17] as a glycosylation site, as are Thr-170 (mucin type *O*-linked glycosylation) and other mucin sites at Thr-245 and Thr-381. Other probably O-GlcNAC attachment sites (Ser-142, Ser-227, Thr-231, Ser-304, Thr-366, and Thr-381) were also predicted. Our analysis of *N*-glycosylation with 3 different websites suggest Asn-154 also as a likely site of *N*-glycosylation in agreement with dengue virus[18].

Cryo-EM has been used to show how glycoprotein E dimers arrange on the surface of virions for flaviviruses including tick-borne encephalitis virus[103], West Nile virus[104], dengue virus 1[105] and dengue virus 4[106] (Supplementary material S3). The early work on the tick-borne encephalitis virus[103] suggested 30 dimers on the surface and also pointed to how the glycoprotein E dimers can reorganize under low pH to form a trimer. The packing of the dimers in the dengue virion is different to tick-borne encephalitis with the glycoprotein E dimers showing 30 'herringbone rafts' each containing three dimers to result in 180 copies of the protein[67,107]. West Nile virus again has a different arrangement with 60 trimers shown in the structure of the immature virus[104] (Supplementary material S4). Even between the dengue serotypes 1, 2 and 4 for which there are cryo-EM structures[105,106] it is apparent while the rafts are very similar (as are the sequence identities [60%]) there is a different charge distribution of the surface of each. Dengue serotype 2 had larger continuous patches of positive charges which was proposed

to enable improved binding to heparan sulfate. This might also be the case for ZIKV in that the charge pattern is again different and could be key for vaccine development. The availability of virion structures makes it feasible to understand structure function of the complete virus such as assessment of membrane curvature and how organization of membrane proteins affects this[43].

A model of the Zika virion was constructed as an illustration using the homology model of the glycoprotein E dimer (Figure 3). While the combined protein sequence of glycoprotein E and the immunoglobulin like domain is closest to dengue virus 1 (57 percent identity, Table 1) the closest template was for the crystal structure of the Japanese encephalitis virus envelope protein (53.12 percent identity, Table 2). This would suggest the virion should more closely resemble that of dengue virus 1, while producing a homology model based on a more distant virus might not be ideal. The homology model of glycoprotein E developed for the mature conformation in this study is significantly different from that developed previously for the immature conformation (Figure 4). The proposed binding site centered around residues 270-277 appears shallower in the mature conformation and this would certainly affect the kinds of molecules that it could interact with. It might also point to the need to interfere with the immature conformation as preferable versus the mature conformation. Ultimately perhaps this model of the Zika virion could help us understand how drugs could access the virus. Viruses affecting pregnancy, like say Varicella which causes microcephaly and other developmental problems[106,107], are often treated with IV immunoglobulin, i.e. antibodies, as well as antivirals to reduce the effect of the virus (or to avoid infection if given soon after exposure). The models could help us design combination approaches possibly targeting multiple proteins that might prevent drug resistance from occurring also.

Does having the homology models and the virion illustration help understand function? Well, the surface charge pattern might be inferred from the homology model and could be compared with dengue and other filoviruses for which there are cryo-EM structures. This may in turn present opportunities for vaccine design by indicating accessible surfaces and properties, allowing mapping of epitopes, design of accessible fragments and peptides for vaccine/diagnostic design. Vaccines themselves might be the only way to avoid the inevitable, otherwise, simply reducing the spread of ZIKV would just delay it. Women ultimately may just want to 'get it over with' and have ZIKV before they get pregnant and hope there is lasting immunity.

In summary, in the absence of crystal structures for any of the proteins comprising the ZIKV, we are left to attempt to construct homology models which we have done using the freely available SWISS-MODEL server. Further preparation of these models required freely available and commercial tools. In the case of the ZIKV glycoprotein E homology model, this has the added benefit of enabling the construction of a full virion. By comparing the Zika virion to the existing structures for other flaviviruses we can see similarities and differences on the surface (Supplementary material S4, Supplementary material S5). This relatively crude approach could help to understand how we might develop antivirals and vaccines against it. In addition we now provide homology models as a starting point for (small and large scale) docking studies and

further evaluation which may complement other modeling efforts for ZIKV[110]. Ultimately the results of their use can be compared with using ZIKV crystal structures once generated.

Author contributions

All authors contributed to the collaborative writing of this project. SE conceived and designed the experiments. SE, JL, BJN, WGL, CS and CHA carried out the research.

Competing interests

S.E. works for Collaborations in Chemistry, Collaborations Pharmaceuticals, Inc. and Collaborative Drug Discovery, Inc.

Supplementary material

Supplementary material S1 (Alignments of ZIKV to other proteins), S2 (Alignment of individual ZIKV proteins for homology models), S4 (ZIKV versus dengue virion animation), and S5 (Published flavivirus cryo-EM structures [not to scale]).

Supplementary material S3. PDB files for ZIKV homology models.

References

1. Mlera L, Melik W, Bloom ME: **The role of viral persistence in flavivirus biology.** *Pathog Dis.* 2014; **71**(2): 137–63.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

2. Pierson TC, Kielian M: **Flaviviruses: braking the entering.** *Curr Opin Virol.* 2013; **3**(1): 3–12.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

3. Fauci AS, Morens DM: **Zika Virus in the Americas--Yet Another Arbovirus Threat.** *N Engl J Med.* 2016; **374**(7): 601–4.
   PubMed Abstract | **Publisher Full Text**

4. Ekins S, Mietchen D, Coffee M, *et al.*: **Open drug discovery for the Zika virus [version 1; referees: awaiting peer review].** *F1000Res.* 2016; **5**: 150.
   Publisher Full Text

5. Martines RB, Bhatnagar J, Keating MK, *et al.*: ***Notes from the Field*: Evidence of Zika Virus Infection in Brain and Placental Tissues from Two Congenitally Infected Newborns and Two Fetal Losses - Brazil, 2015.** *MMWR Morb Mortal Wkly Rep.* 2016; **65**(6): 159–60.
   PubMed Abstract | **Publisher Full Text**

6. Mlakar J, Korva M, Tul N, *et al.*: **Zika Virus Associated with Microcephaly.** *N Engl J Med.* 2016.
   PubMed Abstract | **Publisher Full Text**

7. Rubin EJ, Greene MF, Baden LR: **Zika Virus and Microcephaly.** *N Engl J Med.* 2016.
   PubMed Abstract | **Publisher Full Text**

8. Heymann DL, Hodgson A, Sall AA, *et al.*: **Zika virus and microcephaly: why is this situation a PHEIC?** *Lancet.* 2016; **387**(10020): 719–21.
   PubMed Abstract | **Publisher Full Text**

9. Bell TM, Field EJ, Narang HK: **Zika virus infection of the central nervous system of mice.** *Arch Gesamte Virusforsch.* 1971; **35**(2): 183–93.
   PubMed Abstract | **Publisher Full Text**

10. Vervaeke P, Alen M, Noppen S, *et al.*: **Sulfated *Escherichia coli* K5 polysaccharide derivatives inhibit dengue virus infection of human microvascular endothelial cells by interacting with the viral envelope protein E domain III.** *PLoS One.* 2013; **8**(8): e74035.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

11. Hamel R, Dejarnac O, Wichit S, *et al.*: **Biology of Zika Virus Infection in Human Skin Cells.** *J Virol.* 2015; **89**(17): 8880–96.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

12. Anon. **Zika and Sexual Transmission.** 2016.
    Reference Source

13. Oduyebo T, Petersen EE, Rasmussen SA, *et al.*: **Update: Interim Guidelines for Health Care Providers Caring for Pregnant Women and Women of Reproductive Age with Possible Zika Virus Exposure - United States, 2016.** *MMWR Morb Mortal Wkly Rep.* 2016; **65**(5): 122–7.
    PubMed Abstract | **Publisher Full Text**

14. Musso D, Nhan T, Robin E, *et al.*: **Potential for Zika virus transmission through blood transfusion demonstrated during an outbreak in French Polynesia, November 2013 to February 2014.** *Euro Surveill.* 2014; **19**(14). pii: 20761.
    PubMed Abstract | **Publisher Full Text**

15. Aubry M, Richard V, Green J, *et al.*: **Inactivation of Zika virus in plasma with amotosalen and ultraviolet A illumination.** *Transfusion.* 2016; **56**(1): 33–40.
    PubMed Abstract | **Publisher Full Text**

16. Gyurech D, Schilling J, Schmidt-Chanasit J, *et al.*: **False positive dengue NS1 antigen test in a traveller with an acute Zika virus infection imported into Switzerland.** *Swiss Med Wkly.* 2016; **146**: w14296.
    PubMed Abstract

17. Faye O, Freire CC, Iamarino A, *et al.*: **Molecular evolution of Zika virus during its emergence in the 20th century.** *PLoS Negl Trop Dis.* 2014; **8**(1): e2636.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

18. Alen MM, Kaptein SJ, De Burghgraeve T, *et al.*: **Antiviral activity of carbohydrate-binding agents and the role of DC-SIGN in dengue virus infection.** *Virology.* 2009; **387**(1): 67–75.
    PubMed Abstract | **Publisher Full Text**

19. Lei Y, Yu H, Dong Y, *et al.*: **Characterization of N-Glycan Structures on the Surface of Mature Dengue 2 Virus Derived from Insect Cells.** *PLoS One.* 2015; **10**(7): e0132122.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

20. Nishikawa I, Nakajima Y, Ito M, *et al.*: **Computational prediction of *O*-linked glycosylation sites that preferentially map on intrinsically disordered regions of extracellular proteins.** *Int J Mol Sci.* 2010; **11**(12): 4991–5008.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

21. Li Z, Khaliq M, Zhou Z, *et al.*: **Design, synthesis, and biological evaluation of antiviral agents targeting flavivirus envelope proteins.** *J Med Chem.* 2008; **51**(15): 4660–71.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

22. Poh MK, Yip A, Zhang S, *et al.*: **A small molecule fusion inhibitor of dengue virus.** *Antiviral Res.* 2009; **84**(3): 260–6.
    PubMed Abstract | **Publisher Full Text**

23. Wang QY, Patel SJ, Vangrevelinghe E, *et al.*: **A small-molecule dengue virus entry inhibitor.** *Antimicrob Agents Chemother.* 2009; **53**(5): 1823–31.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

24. Mayhoub AS, Khaliq M, Kuhn RJ, *et al.*: **Design, synthesis, and biological evaluation of thiazoles targeting flavivirus envelope proteins.** *J Med Chem.* 2011; **54**(6): 1704–14.
PubMed Abstract | **Publisher Full Text**

25. Schmidt AG, Lee K, Yang PL, *et al.*: **Small-molecule inhibitors of dengue-virus entry.** *PLoS Pathog.* 2012; **8**(4): e1002627.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

26. Zhou Z, Khaliq M, Suk JE, *et al.*: **Antiviral compounds discovered by virtual screening of small-molecule libraries against dengue virus E protein.** *ACS Chem Biol.* 2008; **3**(12): 765–75.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

27. Stahla-Beek HJ, April DG, Saeedi BJ, *et al.*: **Identification of a novel antiviral inhibitor of the flavivirus guanylyltransferase enzyme.** *J Virol.* 2012; **86**(16): 8730–9.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

28. Behnam MA, Nitsche C, Boldescu V, *et al.*: **The Medicinal Chemistry of Dengue Virus.** *J Med Chem.* 2016.
PubMed Abstract | **Publisher Full Text**

29. Vincetti P, Caporuscio F, Kaptein S, *et al.*: **Discovery of Multitarget Antivirals Acting on Both the Dengue Virus NS5-NS3 Interaction and the Host Src/Fyn Kinases.** *J Med Chem.* 2015; **58**(12): 4964–75.
PubMed Abstract | **Publisher Full Text**

30. Shum D, Smith JL, Hirsch AJ, *et al.*: **High-content assay to identify inhibitors of dengue virus infection.** *Assay Drug Dev Technol.* 2010; **8**(5): 553–70.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

31. Crance JM, Scaramozzino N, Jouan A, *et al.*: **Interferon, ribavirin, 6-azauridine and glycyrrhizin: antiviral compounds active against pathogenic flaviviruses.** *Antiviral Res.* 2003; **58**(1): 73–9.
PubMed Abstract | **Publisher Full Text**

32. Raska K Jr, Zedeck MS, Welch AD: **Relationship between the metabolic effects and the pregnancy-interrupting property of 6-azauridine in mice.** *Biochem Pharmacol.* 1966; **15**(12): 2136–8.
PubMed Abstract | **Publisher Full Text**

33. Kochhar DM, Penner JD, Knudsen TB: **Embryotoxic, teratogenic, and metabolic effects of ribavirin in mice.** *Toxicol Appl Pharmacol.* 1980; **52**(1): 99–112.
PubMed Abstract | **Publisher Full Text**

34. Entrican G: **Immune regulation during pregnancy and host-pathogen interactions in infectious abortion.** *J Comp Pathol.* 2002; **126**(2–3): 79–94.
PubMed Abstract | **Publisher Full Text**

35. Blatt J, Farag S, Corey SJ, *et al.*: **Expanding the scope of drug repurposing in pediatrics: the Children's Pharmacy Collaborative.** *Drug Discov Today.* 2014; **19**(11): 1696–8.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

36. Klein DE, Choi JL, Harrison SC: **Structure of a dengue virus envelope protein late-stage fusion intermediate.** *J Virol.* 2013; **87**(4): 2287–93.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

37. Modis Y, Ogata S, Clements D, *et al.*: **A ligand-binding pocket in the dengue virus envelope glycoprotein.** *Proc Natl Acad Sci U S A.* 2003; **100**(12): 6986–91.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

38. Wu J, Bera AK, Kuhn RJ, *et al.*: **Structure of the Flavivirus helicase: implications for catalytic activity, protein interactions, and proteolytic processing.** *J Virol.* 2005; **79**(16): 10268–77.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

39. Geiss BJ, Thompson AA, Andrews AJ, *et al.*: **Analysis of flavivirus NS5 methyltransferase cap binding.** *J Mol Biol.* 2009; **385**(5): 1643–54.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

40. Zhang Y, Corver J, Chipman PR, *et al.*: **Structures of immature flavivirus particles.** *EMBO J.* 2003; **22**(11): 2604–13.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

41. Brecher M, Chen H, Li Z, *et al.*: **Identification and Characterization of Novel Broad-Spectrum Inhibitors of the Flavivirus Methyltransferase.** *ACS Infect Dis.* 2015; **1**(8): 340–349.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

42. Akey DL, Brown WC, Konwerski JR, *et al.*: **Use of massively multiple merged data for low-resolution S-SAD phasing and refinement of flavivirus NS1.** *Acta Crystallogr D Biol Crystallogr.* 2014; **70**(Pt 10): 2719–29.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

43. Zhang W, Kaufmann B, Chipman PR, *et al.*: **Membrane curvature in flaviviruses.** *J Struct Biol.* 2013; **183**(1): 86–94.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

44. Aleshin AE, Shiryaev SA, Strongin AY, *et al.*: **Structural evidence for regulation and specificity of flaviviral proteases and evolution of the *Flaviviridae* fold.** *Protein Sci.* 2007; **16**(5): 795–806.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

45. Kuno G, Chang GJ, Tsuchiya KR, *et al.*: **Phylogeny of the genus *Flavivirus*.** *J Virol.* 1998; **72**(1): 73–83.
PubMed Abstract | **Free Full Text**

46. Anon. **BLAST.**
Reference Source

47. Altschul SF, Gish W, Miller W, *et al.*: **Basic local alignment search tool.** *J Mol Biol.* 1990; **215**(3): 403–10.
PubMed Abstract | **Publisher Full Text**

48. Baronti C, Piorkowski G, Charrel RN, *et al.*: **Complete coding sequence of Zika virus from a French polynesia outbreak in 2013.** *Genome Announc.* 2014; **2**(3): pii: e00500-14.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

49. Benson DA, Cavanaugh M, Clark K, *et al.*: **GenBank.** *Nucleic Acids Res.* 2013; **41**(Database issue): D36–42.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

50. Biasini M, Bienert S, Waterhouse A, *et al.*: **SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information.** *Nucleic Acids Res.* 2014; **42**(Web Server issue): W252–8.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

51. Bordoli L, Kiefer F, Arnold K, *et al.*: **Protein structure homology modeling using SWISS-MODEL workspace.** *Nat Protoc.* 2009; **4**(1): 1–13.
PubMed Abstract | **Publisher Full Text**

52. Benkert P, Tosatto SC, Schomburg D: **QMEAN: A comprehensive scoring function for protein model quality assessment.** *Proteins.* 2008; **71**(1): 261–77.
PubMed Abstract | **Publisher Full Text**

53. Benkert P, Künzli M, Schwede T: **QMEAN server for protein model quality estimation.** *Nucleic Acids Res.* 2009; **37**(Web Server issue): W510–4.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

54. Anon. **The Structure Analysis and Verification Server.**
Reference Source

55. Anon. **PROCHECK and PROCHECK-NMR.**
Reference Source

56. Anon. **N-GlycoSite.**
Reference Source

57. Anon. **GlycoEP.**
Reference Source

58. Chauhan JS, Rao A, Raghava GP: ***In silico* platform for prediction of N-, O- and C-glycosites in eukaryotic protein sequences.** *PLoS One.* 2013; **8**(6): e67008.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

59. Anon. **NetNGlyc.**
Reference Source

60. Kuhn RJ, Zhang W, Rossmann MG, *et al.*: **Structure of dengue virus: implications for flavivirus organization, maturation, and fusion.** *Cell.* 2002; **108**(5): 717–725.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

61. Lu G, Gong P: **Crystal Structure of the full-length Japanese encephalitis virus NS5 reveals a conserved methyltransferase-polymerase interface.** *PLoS Pathog.* 2013; **9**(8): e1003549.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

62. Zhou Y, Ray D, Zhao Y, *et al.*: **Structure and function of flavivirus NS5 methyltransferase.** *J Virol.* 2007; **81**(8): 3891–903.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

63. Xu T, Sampath A, Chao A, *et al.*: **Structure of the Dengue virus helicase/ nucleoside triphosphatase catalytic domain at a resolution of 2.4 Å.** *J Virol.* 2005; **79**(16): 10278–88.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

64. Mancini EJ, Assenberg R, Verma A, *et al.*: **Structure of the Murray Valley encephalitis virus RNA helicase at 1.9 Angstrom resolution.** *Protein Sci.* 2007; **16**(10): 2294–300.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

65. Hammamy MZ, Haase C, Hammami M, *et al.*: **Development and characterization of new peptidomimetic inhibitors of the West Nile virus NS2B-NS3 protease.** *ChemMedChem.* 2013; **8**(2): 231–41.
PubMed Abstract | **Publisher Full Text**

66. Akey DL, Brown WC, Dutta S, *et al.*: **Flavivirus NS1 structures reveal surfaces for associations with membranes and the immune system.** *Science.* 2014; **343**(6173): 881–5.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

67. Zhang X, Ge P, Yu X, *et al.*: **Cryo-EM structure of the mature dengue virus at 3.5-Å resolution.** *Nat Struct Mol Biol.* 2013; **20**(1): 105–10.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

68. Li L, Lok SM, Yu IM, *et al.*: **The flavivirus precursor membrane-envelope protein complex: structure and maturation.** *Science.* 2008; **319**(5871): 1830–4.
PubMed Abstract | **Publisher Full Text**

69. Dokland T, Walsh M, Mackenzie JM, *et al.*: **West Nile virus core protein; tetramer structure and ribbon formation.** *Structure.* 2004; **12**(7): 1157–63.
PubMed Abstract | **Publisher Full Text**

70. Luca VC, AbiMansour J, Nelson CA, *et al.*: **Crystal structure of the Japanese encephalitis virus envelope protein.** *J Virol.* 2012; **86**(4): 2337–46.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

71. Arnold K, Bordoli L, Kopp J, *et al.*: **The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling.** *Bioinformatics.* 2006; **22**(2): 195–201.
PubMed Abstract | **Publisher Full Text**

72. Kiefer F, Arnold K, Künzli M, *et al.*: **The SWISS-MODEL Repository and associated resources.** *Nucleic Acids Res.* 2009; **37**(Database issue): D387–92.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

73. Guex N, Peitsch MC, Schwede T: **Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective.** *Electrophoresis.* 2009; **30**(Suppl 1): S162–73.
PubMed Abstract | **Publisher Full Text**

74. Kokernot RH, Smithburn KC, Muspratt J, *et al.*: **Studies on arthropod-borne viruses of Tongaland. VIII. Spondweni virus, an agent previously unknown, isolated from *Taeniorhynchus (Mansonioides) uniformis*.** *S Afr J Med Sci.* 1957; **22**(2–3): 103–12.
PubMed Abstract

75. Tognarelli J, Ulloa S, Villagra E, *et al.*: **A report on the outbreak of Zika virus on Easter Island, South Pacific, 2014.** *Arch Virol.* 2016; **161**(3): 665–8.
PubMed Abstract | **Publisher Full Text**

76. Bessaud M, Pastorino BA, Peyrefitte CN, *et al.*: **Functional characterization of the NS2B/NS3 protease complex from seven viruses belonging to different groups inside the genus *Flavivirus*.** *Virus Res.* 2006; **120**(1–2): 79–90.
PubMed Abstract | **Publisher Full Text**

77. Anon. **VIZIER.**
Reference Source

78. Bollati M, Alvarez K, Assenberg R, *et al.*: **Structure and functionality in flavivirus NS-proteins: perspectives for drug design.** *Antiviral Res.* 2010; **87**(2): 125–48.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

79. Parsy CC, Alexandre FR, Bidau V, *et al.*: **Discovery and structural diversity of the hepatitis C virus NS3/4A serine protease inhibitor series leading to clinical candidate IDX320.** *Bioorg Med Chem Lett.* 2015; **25**(22): 5427–36.
PubMed Abstract | **Publisher Full Text**

80. Jiang Y, Andrews SW, Condroski KR, *et al.*: **Discovery of danoprevir (ITMN-191/R7227), a highly selective and potent inhibitor of hepatitis C virus (HCV) NS3/4A protease.** *J Med Chem.* 2014; **57**(5): 1753–69.
PubMed Abstract | **Publisher Full Text**

81. Sheng XC, Casarez A, Cai R, *et al.*: **Discovery of GS-9256: a novel phosphinic acid derived inhibitor of the hepatitis C virus NS3/4A protease with potent clinical activity.** *Bioorg Med Chem Lett.* 2012; **22**(3): 1394–6.
PubMed Abstract | **Publisher Full Text**

82. Schoenfeld RC, Bourdet DL, Brameld KA, *et al.*: **Discovery of a novel series of potent non-nucleoside inhibitors of hepatitis C virus NS5B.** *J Med Chem.* 2013; **56**(20): 8163–82.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

83. Kazmierski WM, Hamatake R, Duan M, *et al.*: **Discovery of novel urea-based hepatitis C protease inhibitors with high potency against protease-inhibitor-resistant mutants.** *J Med Chem.* 2012; **55**(7): 3021–6.
PubMed Abstract | **Publisher Full Text**

84. Rosenquist Å, Samuelsson B, Johansson PO, *et al.*: **Discovery and development of simeprevir (TMC435), a HCV NS3/4A protease inhibitor.** *J Med Chem.* 2014; **57**(5): 1673–93.
PubMed Abstract | **Publisher Full Text**

85. Cummings MD, Lin TI, Hu L, *et al.*: **Discovery and early development of TMC647055, a non-nucleoside inhibitor of the hepatitis C virus NS5B polymerase.** *J Med Chem.* 2014; **57**(5): 1880–92.
PubMed Abstract | **Publisher Full Text**

86. Link JO, Taylor JG, Xu L, *et al.*: **Discovery of ledipasvir (GS-5885): a potent, once-daily oral NS5A inhibitor for the treatment of hepatitis C virus infection.** *J Med Chem.* 2014; **57**(5): 2033–46.
PubMed Abstract | **Publisher Full Text**

87. Zhang Y: **I-TASSER.**
Reference Source

88. Zhang Y: **I-TASSER server for protein 3D structure prediction.** *BMC Bioinformatics.* 2008; **9**: 40.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

89. Yang J, Yan R, Roy A, *et al.*: **The I-TASSER Suite: protein structure and function prediction.** *Nat Methods.* 2015; **12**(1): 7–8.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

90. Noble CG, Lim SP, Chen YL, *et al.*: **Conformational Flexibility of the Dengue Virus RNA-Dependent RNA Polymerase Revealed by a Complex with an Inhibitor.** *J Virol.* 2013; **87**(9): 5291–5295.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

91. Flower DR: **Bioinformatics for vaccinology.** Chichester: Wiley-Blackwell, 2008.
Publisher Full Text

92. Du H, Brender JR, Zhang J, *et al.*: **Protein structure prediction provides comparable performance to crystallographic structures in docking-based virtual screening.** *Methods.* 2015; **71**: 77–84.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

93. Levoin N, Calmels T, Krief S, *et al.*: **Homology Model Versus X-ray Structure in Receptor-based Drug Design: A Retrospective Analysis with the Dopamine D3 Receptor.** *ACS Med Chem Lett.* 2011; **2**(4): 293–7.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

94. Carlsson J, Coleman RG, Setola V, *et al.*: **Ligand discovery from a dopamine D$_3$ receptor homology model and crystal structure.** *Nat Chem Biol.* 2011; **7**(11): 769–78.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

95. Godbole AA, Ahmed W, Bhat RS, *et al.*: **Targeting *Mycobacterium tuberculosis* topoisomerase I by small-molecule inhibitors.** *Antimicrob Agents Chemother.* 2015; **59**(3): 1549–57.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

96. Godbole AA, Ahmed W, Bhat RS, *et al.*: **Inhibition of *Mycobacterium tuberculosis* topoisomerase I by m-AMSA, a eukaryotic type II topoisomerase poison.** *Biochem Biophys Res Commun.* 2014; **446**(4): 916–20.
PubMed Abstract | **Publisher Full Text**

97. Tan K, Cao N, Cheng B, *et al.*: **Insights from the Structure of *Mycobacterium tuberculosis* Topoisomerase I with a Novel Protein Fold.** *J Mol Biol.* 2016; **428**(1): 182–93.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

98. Nguyen ED, Norn C, Frimurer TM, *et al.*: **Assessment and challenges of ligand docking into comparative models of G-protein coupled receptors.** *PLoS One.* 2013; **8**(7): e67302.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

99. da Silveira NJ, Arcuri HA, Bonalumi CE, *et al.*: **Molecular models of NS3 protease variants of the Hepatitis C virus.** *BMC Struct Biol.* 2005; **5**: 1.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

100. Lee TV, Johnson RD, Arcus VL, *et al.*: **Prediction of the substrate for nonribosomal peptide synthetase (NRPS) adenylation domains by virtual screening.** *Proteins.* 2015; **83**(11): 2052–66.
PubMed Abstract | **Publisher Full Text**

101. Anon. **PEPI-TiDP23-C103: First-in-Human Study to Examine the Safety, Tolerability, and Plasma Pharmacokinetics of Increasing Single and Repeated Oral Doses of TMC558445 and of a Combined Single Day Dosing of Oral TMC558445 and Oral TMC310911 and Also Oral Darunavir.** 2009.
Reference Source

102. Keck ZY, Enterlein SG, Howell KA, *et al.*: **Macaque Monoclonal Antibodies Targeting Novel Conserved Epitopes within Filovirus Glycoprotein.** *J Virol.* 2015; **90**(1): 279–91.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

103. Ferlenghi I, Clarke M, Ruttan T, *et al.*: **Molecular organization of a recombinant subviral particle from tick-borne encephalitis virus.** *Mol Cell.* 2001; **7**(3): 593–602.
PubMed Abstract | **Publisher Full Text**

104. Zhang Y, Kaufmann B, Chipman PR, *et al.*: **Structure of immature West Nile virus.** *J Virol.* 2007; **81**(11): 6141–5.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

105. Kostyuchenko VA, Zhang Q, Tan JL, *et al.*: **Immature and mature dengue serotype 1 virus structures provide insight into the maturation process.** *J Virol.* 2013; **87**(13): 7700–7.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

106. Kostyuchenko VA, Chew PL, Ng TS, *et al.*: **Near-atomic resolution cryo-electron microscopic structure of dengue serotype 4 virus.** *J Virol.* 2014; **88**(1): 477–82.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

107. Zhang Y, Zhang W, Ogata S, *et al.*: **Conformational changes of the flavivirus E glycoprotein.** *Structure.* 2004; **12**(9): 1607–18.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

108. Scheffer IE, Baraitser M, Brett EM: **Severe microcephaly associated with congenital varicella infection.** *Dev Med Child Neurol.* 1991; **33**(10): 916–20.
PubMed Abstract | **Publisher Full Text**

109. Deasy NP, Jarosz JM, Cox TC, *et al.*: **Congenital varicella syndrome: cranial MRI in a long-term survivor.** *Neuroradiology.* 1999; **41**(3): 205–7.
PubMed Abstract | **Publisher Full Text**

110. Gatherer D: **Zika virus protein structure homology modelling**. 2016.
Reference Source

111. Anon. **EMDataBank.**
Reference Source

Open Peer Review

Current Referee Status: [?] [?]

Version 1

Referee Report 02 August 2016

[?] **Bruno Botta**

Department of Chemistry and Pharmaceutical Technology, Sapienza University of Rome, Rome, Italy

The work of Sean Ekins and colleagues describes an attempt to predict the 3D structure of a number of Zika Virus proteins by homology modeling.

Infections by Zika virus are seriously preoccupying the population and governments of endemic countries, with particular attention to pregnant women. Indeed, the most serious complications of infections by Zika virus have been observed in pregnant women (i.e. microencephalic fetus). However, the evidence that pregnant women are most affected by Zika virus infection, set some difficulties in developing focused therapeutic strategies, as nicely discussed by the authors. However, in this respect, which is the scenario proposed by authors for a candidate inhibitor of Zika virus replication. Why this molecule should be useful, and how pregnant women could deal with the administration of the drug?

Overall, the work is well written and organized in a clear and rational way, even if no experimental validation of the work carried out *in silico* is proposed. Accordingly, conclusions of this work are merely speculative. My major concern is on the usability of the protein structures for drug design or virtual screening, as claimed by the authors in the abstract and discussion. Indeed, sequence identity between template and target sequences is rather low in some cases (< 60 %). Moreover, the atomistic detail of the active site and its conformation may vary noticeably in the models depending on the program used, the force field, the level of refinement (algorithm, steps, solvent model, …) and the quality of the template structure. All these variables should be taken into consideration by the authors, and the refinement of 3D models should be at least attempted or discussed in deeper details. In my personal opinion, details on the conservation rate, sequence identity and structural similarity of the binding sites would be more helpful for drug design purposes.

Finally, an experimental validation of at least one of the structures modeled by authors (i.e. the most promising target for drug designing studies) should be provided. In fact, analysis with PROCHECK or the use of numerical scores is not sufficient to validate a 3D structure generated by homology modeling. There is a high risk that molecular docking towards a low-resolution structure could provide unrealistic results. The paper can be accepted after revision.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Author Response 25 Aug 2016

Sean Ekins, Collaborations in Chemistry, USA

Response: Dear Prof Botta, thank you for taking the time to review and constructively comment on this manuscript.

Infections by Zika virus are seriously preoccupying the population and governments of endemic countries, with particular attention to pregnant women. Indeed, the most serious complications of infections by Zika virus have been observed in pregnant women (i.e. microencephalic fetus). However, the evidence that pregnant women are most affected by Zika virus infection, set some difficulties in developing focused therapeutic strategies, as nicely discussed by the authors. However, in this respect, which is the scenario proposed by authors for a candidate inhibitor of Zika virus replication. Why this molecule should be useful, and how pregnant women could deal with the administration of the drug?

Response – If we understand the question correctly - any drug for Zika should be safe for all including pregnant women. If they have a Zika infection it would be important to eradicate the virus as soon as possible so that the fetus perhaps is not exposed to the virus.

Overall, the work is well written and organized in a clear and rational way, even if no experimental validation of the work carried out *in silico* is proposed.

Response: This work was initiated in Feb 2016, months before the first crystal structures and experimental in vitro systems were published. Our goal was to build protein models that we would eventually use and which others could use in parallel, to start docking-based virtual screening and identify VS and experimental hits, to initiate the drug discovery cascade.

Accordingly, conclusions of this work are merely speculative.

Response: A modeling study, by definition is speculative and does not claim otherwise. However, in good faith, we provided models which did not exist. The fact that they have since gone on to form the basis of the OpenZika Project (manuscript submitted, http://openzika.ufg.br) and have been used to suggest compounds which are now being tested argues for their utility. Notwithstanding, we acknowledge that Zika protease, helicase and glycoprotein E experimental structures have now been elucidated and deposited in PDB.

My major concern is on the usability of the protein structures for drug design or virtual screening, as claimed by the authors in the abstract and discussion. Indeed, sequence identity between template and target sequences is rather low in some cases (< 60 %).

Response – As expected for the enzyme functions, the local active site identity and similarity scores are higher than indicated by the full-length identity %. Thus, a reasonable pocket model for docking is the primary goal, not the structural accuracy for the whole protein. We would also add that any team using our models would doubtless take a pragmatic approach that would include running various types of internal controls for virtual screening studies and confirmatory screening experiments (e.g. running the

Dengue virus templates alone and checking for enrichment of established Dengue actives against the purified proteins). They would also probably prioritize their effort on the models with highest template similarity. Note also that, the first reviewer Dr. Fiser accepts that our similarity values are high enough for reliable modelling.

Moreover, the atomistic detail of the active site and its conformation may vary noticeably in the models depending on the program used, the force field, the level of refinement (algorithm, steps, solvent model, …) and the quality of the template structure.

Response – We agree that this is one of the challenges with this approach, but there were no other options without a single crystal structure for Zika protein when we started this work and submitted it.

All these variables should be taken into consideration by the authors, and the refinement of 3D models should be at least attempted or discussed in deeper details. In my personal opinion, details on the conservation rate, sequence identity and structural similarity of the binding sites would be more helpful for drug design purposes.

Response: We agree with your point. We are now refining the 3D models using molecular dynamics simulations. However, this will be the subject of work on going, taking an in depth approach to each structure and comparison to crystal structures when available. Our main goal in this paper which we submitted in Feb 2016 (and made available early March) was to provide the scientific community with the first 3D models of Zika proteins.

Finally, an experimental validation of at least one of the structures modeled by authors (i.e. the most promising target for drug designing studies) should be provided. In fact, analysis with PROCHECK or the use of numerical scores is not sufficient to validate a 3D structure generated by homology modeling.

Response – We agree that the experimental validation of the protein structures is of upmost importance in drug discovery programs, but this was not the scope of this paper. Again, our main goal was to generate homology models for the key proteins, which none crystal was available at the moment of the submission, and these models could be taken to initiate ZIKV antiviral drug discovery using both high throughput screens as well as structure-based design. Moreover, these structures were the core to initiate the World Community Grid project called OpenZika (
https://www.worldcommunitygrid.org/research/zika/overview.do), which is a global research project to accelerate the discovery of an antiviral against the Zika virus. PROCHECK was originally developed to evaluate the stereochemical quality of experimentally determined protein structures (Laskowski R A, MacArthur M W, Moss D S, Thornton J M (1993). J. App. Cryst., 26, 283-291; Morris A L, MacArthur M W, Hutchinson E G & Thornton J M (1992). Proteins, 12, 345-364). However, PROCHECK has been widely used and reported in the literature to check the stereochemical quality of homology models, to provide numerical scores that can evaluate the statistical quality of the generated 3D protein models (See bellow the list of summarized most recent literature). It is a quality assessment criterion and is not related to experimental validation.

Papers reporting the use of PROCHECK for assessing the stereochemical quality of homology models:

J Biomol Struct Dyn. 2016 Apr 4:1-19 [DOI: 10.1080/07391102.2015.1117397].
Curr Cancer Drug Targets. 2015;15(9):822-35 [PMID: 26567883].
Interdiscip Sci. 2015 Aug 15. In Press [DOI: 10.1007/s12539-015-0121-z].
Onco Targets Ther. 2015 Jul 30;8:1923-30 [DOI: 10.2147/OTT.S84200].
Interdiscip Sci. 2016 Mar;8(1):41-52 [DOI: 10.1007/s12539-015-0269-6].
J Mol Model. 2015 Mar;21(3):37. [DOI: 10.1007/s00894-015-2586-4].

There is a high risk that molecular docking towards a low-resolution structure could provide unrealistic results.

Response – Yes we absolutely agree, but when we started this we had no crystal structures we had to start somewhere as a means to start to filter compounds rather than random HTS.
Our work has also correctly predicted the glycosylation site for glycoprotein E one of the first structures crystallized. Moreover, we have compared our models with the crystallographic structures, when they were made available. The RMSD values ranged from 0.72 to 1.8 Å, for the structures of NS1, NS3, glycoprotein E and NS2/NS3 proteins. We have also optimized the 3D models and the MolProbit scores ranged from 1.28 to 2.81 Å. MolProbit is a score that combines the clashscore, rotamer, and Ramachandran evaluations into a single score, normalized to be on the same scale as X-ray resolution. Therefore, we are optimistic that our models are reliable to start a docking-based virtual screening program in the search of a new antiviral drug, which we already initiated.

The paper can be accepted after revision.

Response – Thank you!

Competing Interests: No competing interests were disclosed.

Referee Report 11 April 2016

Andras Fiser
Albert Einstein College of Medicine, Bronx, NY, USA

The paper describes homology models of 15 proteins encoded in the genome of Zika virus that are built by the SWISSmodel web server. A glycolysation site was also identified. Using the models of Zika glycoprotein E a complete structural model of the virion was constructed.

The practical results of the work are the 11 high quality homology models that could be used in the future for structure based drug development. I believe an interested researcher will initiate time consuming follow up studies with these models only if there is a substantial added value.

These models passed quality assessment criteria according to the authors. However a number of

successful models (passing similarly well the same quality requirements) can be generated that will differ in small but essential details by e.g. side chain placement, or loop conformations. However, these differences can have a dramatic effect on the outcome of subsequent drug docking trials. These alternative models can be obtained when different softwares are used, as the authors allude to it, e.g. using Modeller vs SWISSModel, as these use different forcefields and restraints to generate models. Therefore, the authors should consider providing a more insightful result by running several different modeling programs or using alternative templates (see 4., below), and comparing the results, identifying similarly modeled parts of these models and providing a set of possible solutions for subsequent studies.

A battery of model quality checks were performed and additional energy minimization. However the overall high sequence identities between the target proteins and their respective templates (55% and up,) ensure that these models are highly reliable.

1. Therefore the extensive reporting on Ramachandran plots (Figure 3) is not adding much to the results, it can be transferred to supplementary material.

2. Similarly, Table 2 can be shrunk by eliminating many details of PROCHECK results.

3. However the information on templates should be reported in the table, instead listing them extensively in the text.

4. Template selection was purely based on sequence similarity requirements. Given how influential this step and how few cases require attention in this specific study, one could imagine to perform a more detailed quality check of potential templates. For instance, if several templates are available in the same sequence identity range, one should consider picking the one wit the best resolution etc. Figure 4 exposes this dilemma where two isoforms of the same template can results models with very significant differences.

5. Energy refinement rarely, if ever, improves model quality and its use here should be better justified.

6. Some statements require attention e.g in Abstract: "Eleven out of 15 models pass our criteria for selection". What selection? This must be referring to quality or accuracy of models, and should be rephrased accordingly.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Author Response 25 Aug 2016

Sean Ekins, Collaborations in Chemistry, USA

Dear Prof Fiser, thank you for taking the time to review and constructively comment on this manuscript.

The paper describes homology models of 15 proteins encoded in the genome of Zika virus that are built by the SWISSmodel web server. A glycolysation site was also identified. Using the models of Zika glycoprotein E a complete structural model of the virion was constructed.

The practical results of the work are the 11 high quality homology models that could be used in the future for structure based drug development. I believe an interested researcher will initiate time consuming follow up studies with these models only if there is a substantial added value.

Response: Thank you! We would also point out the utility of this work is not only reflected in the high level of accesses, downloads and citations, but also for initiating the World Community Grid as the OpenZika project, which was launched in May, 2016 (https://www.worldcommunitygrid.org/research/zika/overview.do)

These models passed quality assessment criteria according to the authors. However a number of successful models (passing similarly well the same quality requirements) can be generated that will differ in small but essential details by e.g. side chain placement, or loop conformations. However, these differences can have a dramatic effect on the outcome of subsequent drug docking trials. These alternative models can be obtained when different softwares are used, as the authors allude to it, e.g. using Modeller vs SWISSModel, as these use different force fields and restraints to generate models. Therefore, the authors should consider providing a more insightful result by running several different modeling programs or using alternative templates (see 4., below), and comparing the results, identifying similarly modeled parts of these models and providing a set of possible solutions for subsequent studies.

Response: We would agree that certainly an exhaustive analysis of different methods was not undertaken. That was not our goal. We attempted to build models that we could then put into practice and provide to the community for them to use knowing the possible caveats.

A battery of model quality checks were performed and additional energy minimization. However the overall high sequence identities between the target proteins and their respective templates (55% and up,) ensure that these models are highly reliable.

Response: Thank you!
1. Therefore the extensive reporting on Ramachandran plots (Figure 3) is not adding much to the results, it can be transferred to supplementary material.

   Response: Actually many publications use these plots and they may provide some insights on model quality so we have opted to retain them in the body of the manuscript, as the F1000Research journal does not have a limit for figures.

2. Similarly, Table 2 can be shrunk by eliminating many details of PROCHECK results.

   Response: This is an internet publication we are not constrained by pages or page charges and these numbers are informative to readers.

3. However the information on templates should be reported in the table, instead listing them extensively in the text.

   Response: We agree and we believe that it is neater and more accessible in a table. Therefore, we have added the information on templates in Table 2 and removed this information from the text.

4. Template selection was purely based on sequence similarity requirements. Given how influential this step and how few cases require attention in this specific study, one could imagine to perform a more detailed quality check of potential templates. For instance, if several templates are available in the same sequence identity range, one should consider picking the one with the best resolution etc. Figure 4 exposes this dilemma where two isoforms of the same template can results models with very significant differences.

Response: We totally agree – some decisions had to be made as a trade-off to get the models built in a reasonable time to put into use. We could have spent months building models by which time the x-ray structures would have been out. We had models of the proteins and the virion at least 1 month before the first cryo-EM structure. One month is a long time in drug discovery, in which we could be finding molecules for testing.

5. Energy refinement rarely, if ever, improves model quality and its use here should be better justified.

Response: We totally agree with you, Dr. Fiser. Energy minimization does not improve model quality. The word "refined"in the paper was a mistake and we have corrected it now. We are using some servers, such as KoBaMIN server, to perform structure refinement and to improve models quality, but this will be subject of another manuscript. Our main goal in this first paper was to provide the scientific community with homology models for the key proteins, which none crystal was available at the moment of the submission of this paper, and these models could be taken to initiate ZIKV antiviral drug discovery using both high throuput screens as well as structure-based design. Moreover, these structures were the core to initiate the World Community Grid project called OpenZika (https://www.worldcommunitygrid.org/research/zika/overview.do), which is a global research project to accelerate the discovery of an antiviral against the Zika virus.

6. Some statements require attention e.g in Abstract: "Eleven out of 15 models pass our criteria for selection". What selection? This must be referring to quality or accuracy of models, and should be rephrased accordingly.

Response: OK Thank you! Changed to - Eleven out of 15 models pass our model quality criteria for their further use.

Competing Interests: No competing interests were disclosed.