PeerJ

Structural models for the design of novel antiviral agents against Greek Goat Encephalitis

Louis Papageorgiou[1,2], Styliani Loukatou[1], Vassiliki Lila Koumandou[1], Wojciech Makałowski[3], Vasileios Megalooikonomou[4], Dimitrios Vlachakis[1] and Sophia Kossida[1]

[1] Computational Biology & Medicine Group, Biomedical Research Foundation, Academy of Athens, Athens, Greece
[2] Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Athens, Greece
[3] Institute of Bioinformatics, University of Münster, Münster, Germany
[4] Computer Engineering and Informatics Department, University of Patras, Greece

ABSTRACT

The Greek Goat Encephalitis virus (GGE) belongs to the *Flaviviridae* family of the genus *Flavivirus*. The GGE virus constitutes an important pathogen of livestock that infects the goat's central nervous system. The viral enzymes of GGE, helicase and RNA-dependent RNA polymerase (RdRP), are ideal targets for inhibitor design, since those enzymes are crucial for the virus' survival, proliferation and transmission. In an effort to understand the molecular structure underlying the functions of those viral enzymes, the three dimensional structures of GGE NS3 helicase and NS5 RdRP have been modelled. The models were constructed in silico using conventional homology modelling techniques and the known 3D crystal structures of solved proteins from closely related species as templates. The established structural models of the GGE NS3 helicase and NS5 RdRP have been evaluated for their viability using a repertoire of in silico tools. The goal of this study is to present the 3D conformations of the GGE viral enzymes as reliable structural models that could provide the platform for the design of novel anti-GGE agents.

Subjects Bioinformatics, Biotechnology, Computational Biology, Virology, Infectious Diseases
Keywords Greek Goat Encephalitis virus, Flaviviridae, RNA-dependent RNA polymerase, Homology modelling, Phylogenetic analysis, Drug design, Helicase, Computational biology

INTRODUCTION

Greek Goat Encephalitis (GGE) is an endemic virus in Greece which infects the central nervous system of goats and leads to neurological disorders. As an arbovirus (arthropod-borne virus), GGE can be transmitted through ticks. In spite of the association of GGE with Tick-borne encephalitis (TBE) and Louping ill (LI) virus (*Pavlidou et al., 2008*) which are zoonotic viruses (transmitted from animals to humans), there are no verified zoonosis cases yet regarding the GGE virus (*Vlachakis, Tsiliki & Kossida, 2013*). However, even though the GGE virus may not infect humans, the results of its epidemics are catastrophic, since national economies are closely related with livestock (*Ataide Dias, Mahon & Dore,*

Submitted 18 June 2014
Accepted 22 October 2014
Published 6 November 2014

Corresponding authors
Dimitrios Vlachakis, dvlachakis@bioacademy.gr
Sophia Kossida, skossida@bioacademy.gr

Academic editor
Maria Anisimova

Additional Information and Declarations can be found on page 15

DOI **10.7717/peerj.664**

2008). GGE is a member of the Flaviviridae virus family (*Papageorgiou et al., 2014*; *Pavlidou et al., 2008*), a family with several species causing lethal vertebrate infections, through arthropod vector transmissions (*Vlachakis & Kossida, 2013*). Numerous important human and animal pathogens are classified in the Flaviviridae virus family. GGE belongs to the genus Flavivirus and specifically the group of Mammalian tick-borne encephalitis viruses (*Vlachakis, Argiro & Kossida, 2013*). The GGE viral genome is a positive-sense RNA in a linear, single-stranded form, of about 10 kb length.

To date, no anti-GGE virus therapy is available, while many other fatal viruses of the Flaviviridae family also remain untreatable, so there is an urgent need for new antiviral drugs and vaccines (*Vlachakis, Tsaniras & Kossida, 2013*). Based on recent genomic studies, Flavivirus RNA replication is associated with the non-structural proteins, such as NS3 helicase and NS5 RdRP, encoded by the C-terminal part of the viral polyprotein (Fig. S1) (*Nulf & Corey, 2004*). Viral RNA-dependent RNA polymerase is encoded by the NS5 domain and catalyzes the replication of RNA. The viral helicase is encoded by the NS3 domain and is the enzyme responsible for the unwinding of the viral genetic material during replication. Therefore, the NS3 viral helicase and NS5 viral RdRP constitute ideal pharmacological targets, as they are key enzymes for the survival, propagation and eventual transmission of the virus (*Vlachakis et al., 2012*). We propose that structural research and the establishment of a comprehensive in silico platform can aid the de novo design of novel inhibitors (*Dalkas et al., 2013*).

METHODS

Database sequence search

NS3 helicase and NS5 RdRP sequence data from 87 species with fully sequenced genomes were collected from the NIAID Virus Pathogen Database and Analysis Resource (ViPR) (*Pickett et al., 2012*) (http://www.viprbrc.org), and the NCBI RefSeq database, as described in *Vlachakis, Koumandou & Kossida (2013)* In total, 7 *Pestivirus*, 8 *Hepacivirus*, 3 *Pegivirus* and 69 *Flavivirus* sequences were used, as shown in Table S1. In cases where the NS5 annotation was not available in RefSeq, the start and end positions of the NS5 sequence within the whole genome polyprotein was inferred from CLUSTALW alignments with closely related annotated species; only the amino acid sequence corresponding to NS5, determined in this way, was used for the multiple sequence alignments and phylogenetic analysis (Table S1).

Phylogenetic analysis

Alignments of the NS5 protein sequences (available upon request) were created using MUSCLE (*Edgar, 2004*), and visualized using Jalview (*Waterhouse et al., 2009*). Only unambiguous homologous regions were retained for phylogenetic analysis; manual masking and trimming was performed in MacClade (*Maddison & Maddison, 1989*). The NS5 alignment was combined with data about NS3 helicase from our previous work (*Vlachakis, Koumandou & Kossida, 2013*) to create the tree based on both sequences. For each species, the NS3 helicase and NS5 RdRP sequences were concatenated after alignment,

masking and trimming. This concatenated alignment was then tested with ProtTest (*Abascal, Zardoya & Posada, 2005*) to estimate the appropriate model of sequence evolution. Phylogenetic analysis was performed by Bayesian and Maximum likelihood methods as described in *Vlachakis, Koumandou & Kossida (2013)* with 100 bootstrap replicates. Trees were visualized using FigTree v1.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Molecular modelling

All calculations and visual constructions were performed on a six-core workstation running Windows 7 and using Molecular Operating Environment (MOE) version 2013.08 software package developed by Chemical Computing Group (Montreal, Canada).

Identification of template structures and sequence alignment

The amino acid sequence of the GGE genome polyprotein was retrieved from the conceptual translation of the virus genome sequence at the NCBI database (http://www.ncbi.nlm.nih.gov/) (genome sequence GenBank: DQ235153.1) (polyprotein sequence GenBank: ABB90677.1). The blastp algorithm (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to identify homologous structures by searching in the Protein Data Bank (PDB) (*Berman et al., 2002*; *Berman et al., 2013*). The multiple sequence alignment was performed using MOE. Conserved residues in the multiple alignment were visualized with WEBLOGO (http://weblogo.berkeley.edu/logo.cgi).

Homology modelling

The homology modelling of the Greek Goat Encephalitis viral RdRP and helicase were carried out using MOE. The selection of template crystal structures for homology modelling was based on the primary sequence identity and the crystal resolution. The MOE homology model method is separated into four main steps (*Vlachakis et al., 2013a*). First, a primary fragment geometry specification. Second, the insertion and deletions task. The third step is the loop selection and the side-chain packing and the last step is the final model selection and refinement.

Model optimization

Energy minimisation was done in MOE initially using the Amber99 (*Loukatou et al., 2014*; *Wang, Cieplak & Kollman, 2000*) force-field implemented into the same package, up to a root mean square deviation (RMSd) gradient of 0.0001 to remove the geometrical strain. The models were subsequently solvated with simple point charge (SPC) water using the truncated octahedron box extending to 7 Å from the model, and molecular dynamics were performed at 300 K, 1 atm with 2 fsecond step size for a total of ten nanoseconds, using the NVT ensemble in a canonical environment (NVT stands for Number of atoms, Volume and Temperature that remain constant throughout the calculation). The results of the molecular dynamics simulation were collected into a database by MOE for further analysis (*Vlachakis et al., 2014*).

Model evaluation

The produced models were initially evaluated within the MOE package by a residue packing quality function which depends on the number of buried non-polar side-chain groups and on hydrogen bonding. Moreover, the suite PROCHECK (*Laskowski et al., 1996*; *Papagelopoulos et al., 2014*) was employed to further evaluate the quality of the produced models. Finally, Verify3D (*Eisenberg, Luthy & Bowie, 1997*; *Luthy, Bowie & Eisenberg, 1992*) was used to evaluate whether the models of NS3 helicase and NS5 RdRP domains are similar to known protein structures of this viral family (*Vlachakis, Karozou & Kossida, 2013*).

RESULTS

Phylogenetic analysis

Alignment of the *Flaviviridae* NS5 protein sequences from a variety of *Flaviviridae* species highlights novel important conserved functional domains (Fig. S2), as described previously for NS3 helicase (*Vlachakis, Koumandou & Kossida, 2013*). Good conservation is evident throughout the whole length of the sequence, particularly between species that belong to the same genus. The annotated *Pestivurs* NS5 sequences (∼1215 aa) are significantly longer than the annotated NS5 sequences from *Flavivirus* (∼903 aa) and *Hepacivirus* or Pegivirus (∼1039 aa). *Pestivirus* NS5 shows various internal insertions mostly throughout the N-terminal half of the protein. Regions conserved across all three genera likely indicate important functional domains of the NS5 protein, and the alignment highlights 13 residues which are absolutely conserved between all species. Many of the conserved regions identified here have not been reported previously, and probably deserve further study.

When data for NS3 helicase and NS5 RdRP is combined for phylogenetic reconstruction across all 87 *Flaviviridae* species (Fig. 1), the results are similar to our previous analysis based on NS3 helicase (*Vlachakis, Koumandou & Kossida, 2013*). Clear separation of the different genera is seen. Within the *Flavivirus*, monophylletic groups for the tick-borne and insect-specific species are evident. TABV and the insect-specific *Flavivirus* appear basal to (closest to the origin of) the whole *Flavivirus* group. Mosquito-borne species diverge afterwards, followed by the tick-borne species, and species with no known vector (NKV). Most of the groupings in the tree are highly supported by posterior probability and bootstrap values by all three phylogenetic methods.

The tree highlights the phylogenetic distance between species for which structural information regarding NS3 helicase and NS5 RdRP is available. No structure is available from a tick-borne *Flavivirus* species for either NS3 helicase or NS5 RdRP, so our model proposed here for GGE fills a gap in structural information covering all *Flaviviridae* lineages. Given that *Flaviviridae* NS5 sequences for which structures are available come from phylogenetically disparate groups (see arrows in Fig. 1) alignment of these structures allows fine-tuning of the structural model, and highlights conserved domains (Fig. S3).

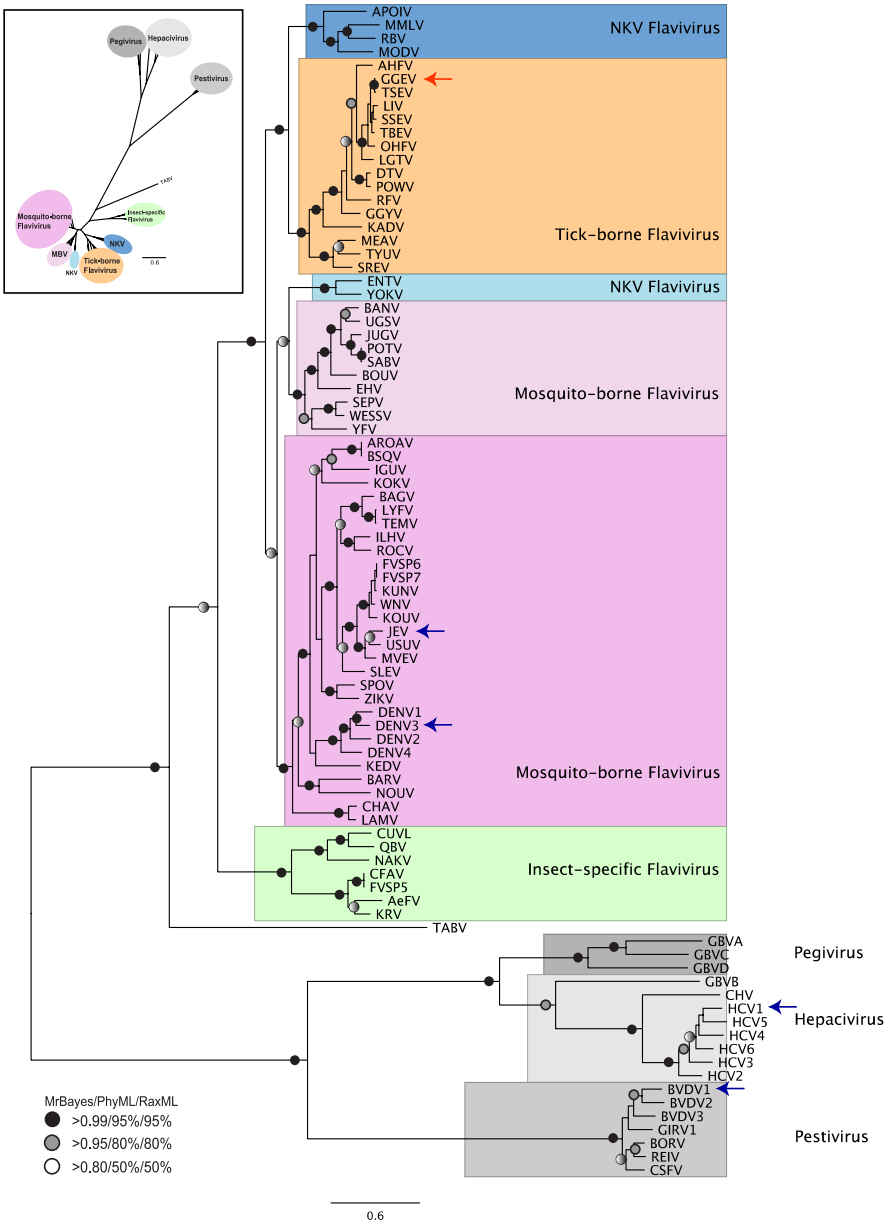

Figure 1 Phylogenetic reconstruction of *Flaviviridae* concatenated NS3 helicaseand NS5 RdRP protein sequences. The tree shown is the best Bayesian topology. Numerical values at the nodes of the tree (x/y/z) indicate statistical supportby MrBayes, PhyML and RAxML (posterior probability, bootstrap and bootstrap, respectively). Values for highly supported nodes have been replaced by symbols, as indicated. Nodes with at least 0.95 posterior probability and 80% bootstrap support were considered robust, and nodes with at least 0.80 posterior probability and 50% bootstrap support are indicated. The position of GGE is indicated by a red arrow. The species for which 3D structure of NS5 RdRP are available are indicated with a blue arrow. The yellow fever virus (YFV), for which the 3D structure of NS3 helicase is available, is indicated with a green arrow. Full details and accession numbers for all protein sequences used are in Table S1. The tree confidently separates the *Hepacivurs*, *Pestivirus*, *Pegivirus*and *Flavivirus* genera. Within the *Flavivirus*, TABV and insect-specific species appear basal, whereas Tick-borne species and species with no known vector (NKV) are the most derived.

Structural models of the Greek Goat Encephalitis NS3 helicase and NS5 RdRP

BLASTp against the PDB was used to identify the best available crystal structures as templates for homology modelling. Sequence alignments of the Greek goat encephalitis viral RdRP and viral helicase, to homologs from other *Flaviviridae*, identified several motifs that are important for biological functionality (*Vlachakis, Kontopoulos & Kossida, 2013*). A multiple sequence alignment (Fig. S3) was constructed including the Greek Goat Encephalitis NS5 viral RdRP, the crystal structure of Japanese Encephalitis NS5 RdRP with a conserved methyltransferase-polymerase interface (PDB code 4K6M) (*Lu & Gong, 2013*), the crystal structure of Dengue Virus NS5 RdRP bound to NITD-107 (PDB code 3VWS) (*Noble et al., 2013*), the crystal structure of Bovine Viral Diarrhea Virus NS5 RdRP (PDB code 1S48) (*Choi et al., 2004*) and the crystal structure of Hepatitis C NS5 RdRP with short RNA template strand (PDB code 1NB7) (*O'Farrell et al., 2003*).

The final choice of a template structure was not only based on the percent sequence identity, but also on the structure resolution. The Japanese Encephalitis virus RdRP and Yellow Fever helicase are suitable templates because both viruses belong to the same viral family with Greek Goat Encephalitis (*Flaviviridae*). The predicted secondary structure for the GGE virus RdRP was found to be similar to the actual structure of the JE virus RdRP; the same was true for the GGE viral helicase and the Yellow Fever viral helicase. The Japanese Encephalitis virus, Dengue virus, Hepatitis C virus and Greek Goat Encephalitis RdRPs all belong to the superfamily of RdRP and share eight common motifs within their domains. The *Flaviviridae* family helicases belong to the superfamily II of helicases and within their domains share seven common motifs. All RdRP and helicase invariant motifs have been conserved on the Greek Goat Encephalitis models (*Lu & Gong, 2013*).

The alignment between the RdRP sequence of the Greek Goat Encephalitis and the sequence of the Japanese Encephalitis virus RdRP (PDB 4K6M, resolution 2.6 Å) (*Lu & Gong, 2013*) template revealed 58% Identity and 73% similarity (Fig. 2). The alignment for the viral helicase between the sequence of the Greek Goat Encephalitis and the sequence of the Yellow Fever virus (PDB 1YKS, resolution 1.8 Å) (*Wu et al., 2005*) template revealed 50% Identity and 65% similarity (Fig. 3). These results allow for a conventional homology modelling approach to be considered. The two models of NS5 GGE RdRP and NS3 GGE helicase were first structurally superimposed and subsequently compared to their templates using the Swiss Pdb viewer (4.1.0) algorithm (*Guex & Peitsch, 1997*). The NS5 GGE RdRP model and the NS3 GGE helicase model exhibited an alpha-carbon RMSD of less than 0.78 and less than 1.35 angstroms, respectively. Furthermore, the models were evaluated in terms of their geometry with MOE and the Verify3D algorithm (*Eisenberg, Luthy & Bowie, 1997*). The compatibility of both the NS5 GGE RdRP model and the NS3 GGE helicase model, with their own amino acid sequence was assessed using the Verify3D algorithm. A structural class was delegated for each residue of every model based on its 3D environment and location (*Vlachakis et al., 2013b*). Each residue score value was calculated by the algorithm using a collection of reference structures. The scores for the GGE virus RdRP model ranged from +0.02 to +0.75 and for the GGE virus helicase ranged from

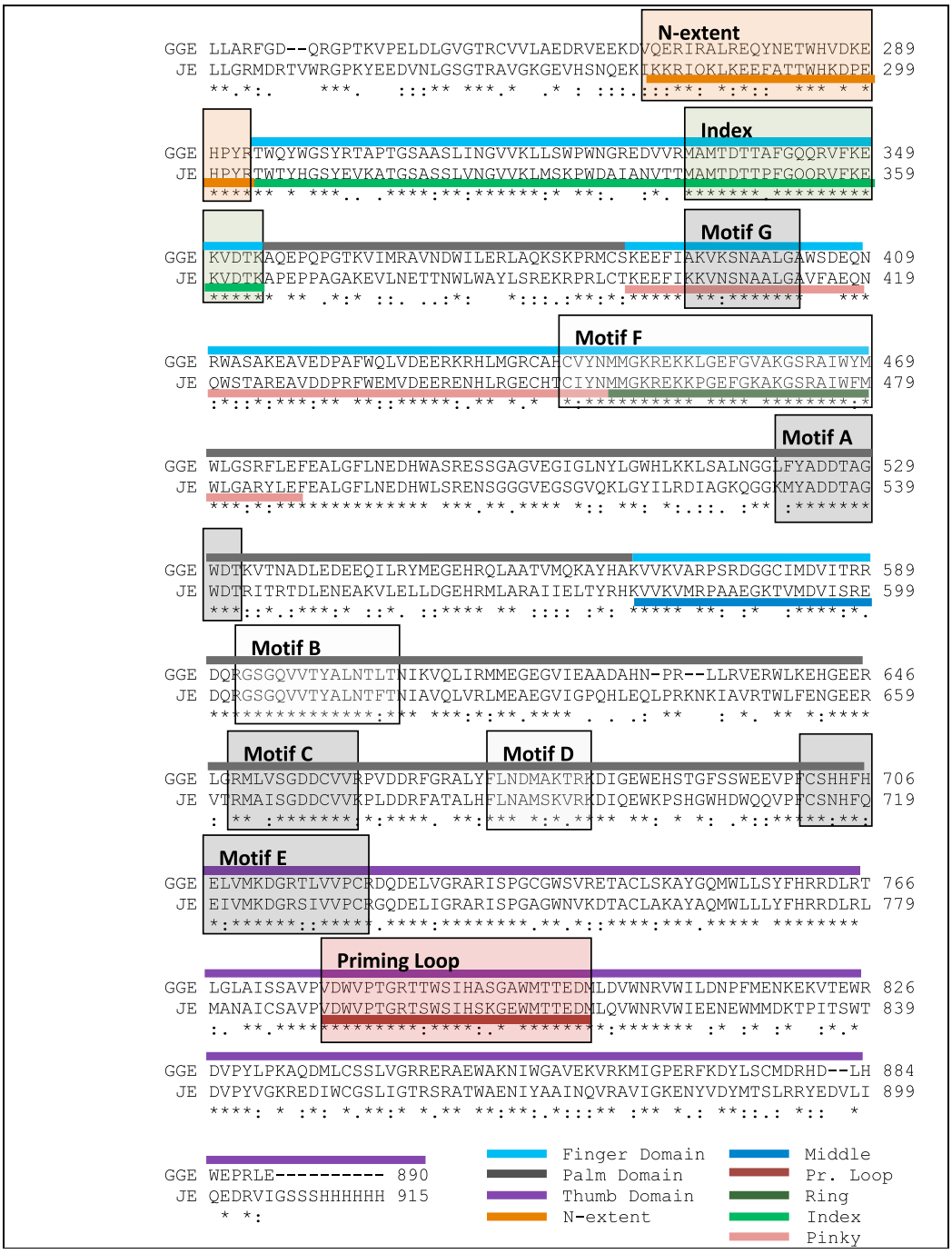

Figure 2 Sequence alignment between the GreekGoat Encephalitis viral NS5 RdRP and the corresponding sequence of the crystal structure of Japanese Encephalitis NS5 RdRP. All eight major conserved motifs of *Flaviviridae* have been marked (Motifs A–G).

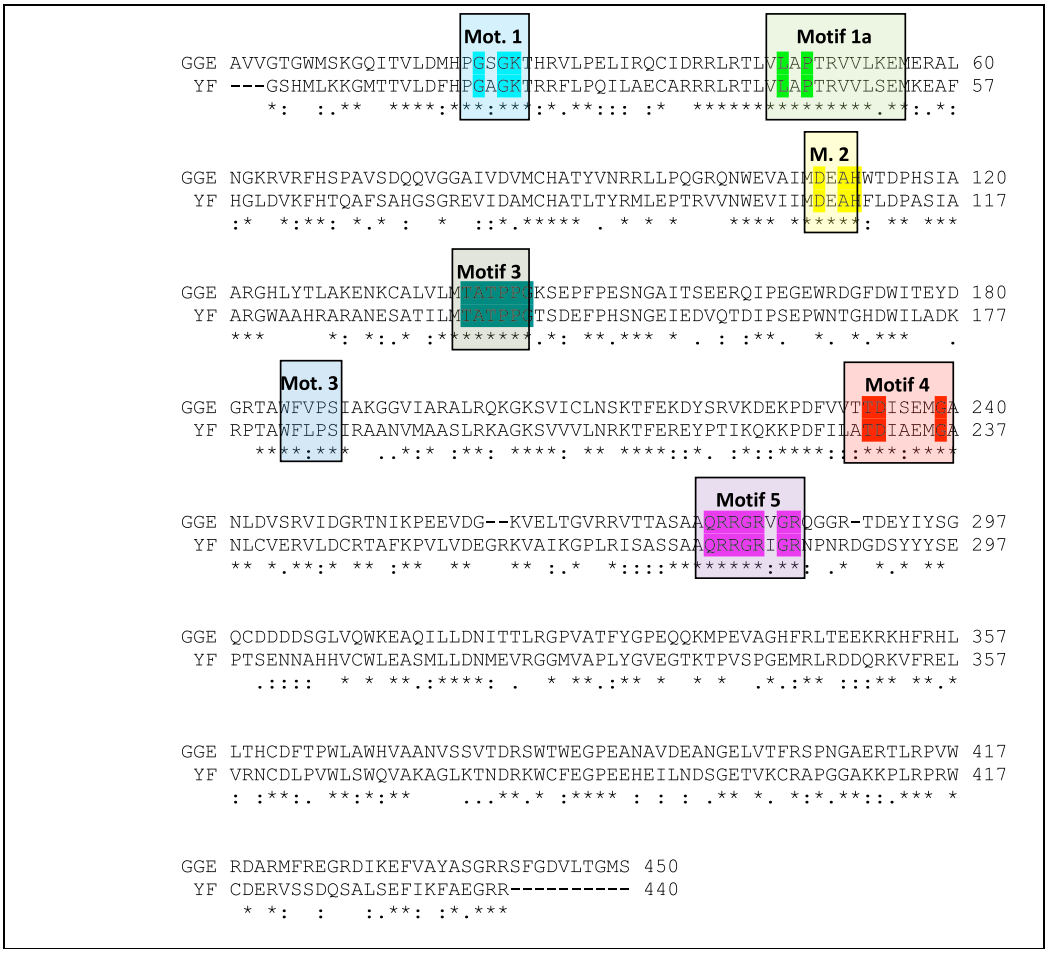

Figure 3 Sequence alignment between the Greek Goat Encephalitis viral NS3 helicase and the corresponding sequence of the crystal structure of Yellow Fever NS3 helicase. All seven major conserved motifs of *Flaviviridae* have been marked (M1, M1a, M2–M6).

−0.09 to +0.7. According to these results, the two models can be characterized as having high quality; scores below +0.1 are indicative of serious problems in the models (*Bujnicki, Rychlewski & Fischer, 2002*). It was concluded that no poorly scored stretches of amino acids were included neither in structured nor conserved regions of the NS3 helicase and NS5 RdRP models (Table S2 and Fig. S4).

Description of the Greek Goat Encephalitis helicase model

As predicted, from the sequence alignment of Greek Goat Encephalitis helicase (Fig. 3) and its secondary structure prediction (data not shown), the constructed model presented the structural features of known *Flaviviridae* helicases. In the Greek Goat Encephalitis helicase model (Fig. 4) the three distinct domains of helicases were structurally conserved, as well as the various motifs described by *Wu et al. (2005)*. The seven sequence motifs of the GGE helicase occur in domain 1 and 2 (Figs. 3, 4C and 4D) (*Gorbalenya & Koonin, 1993*). Helicase motif residues are involved in NTP binding and hydrolysis. Moreover, those

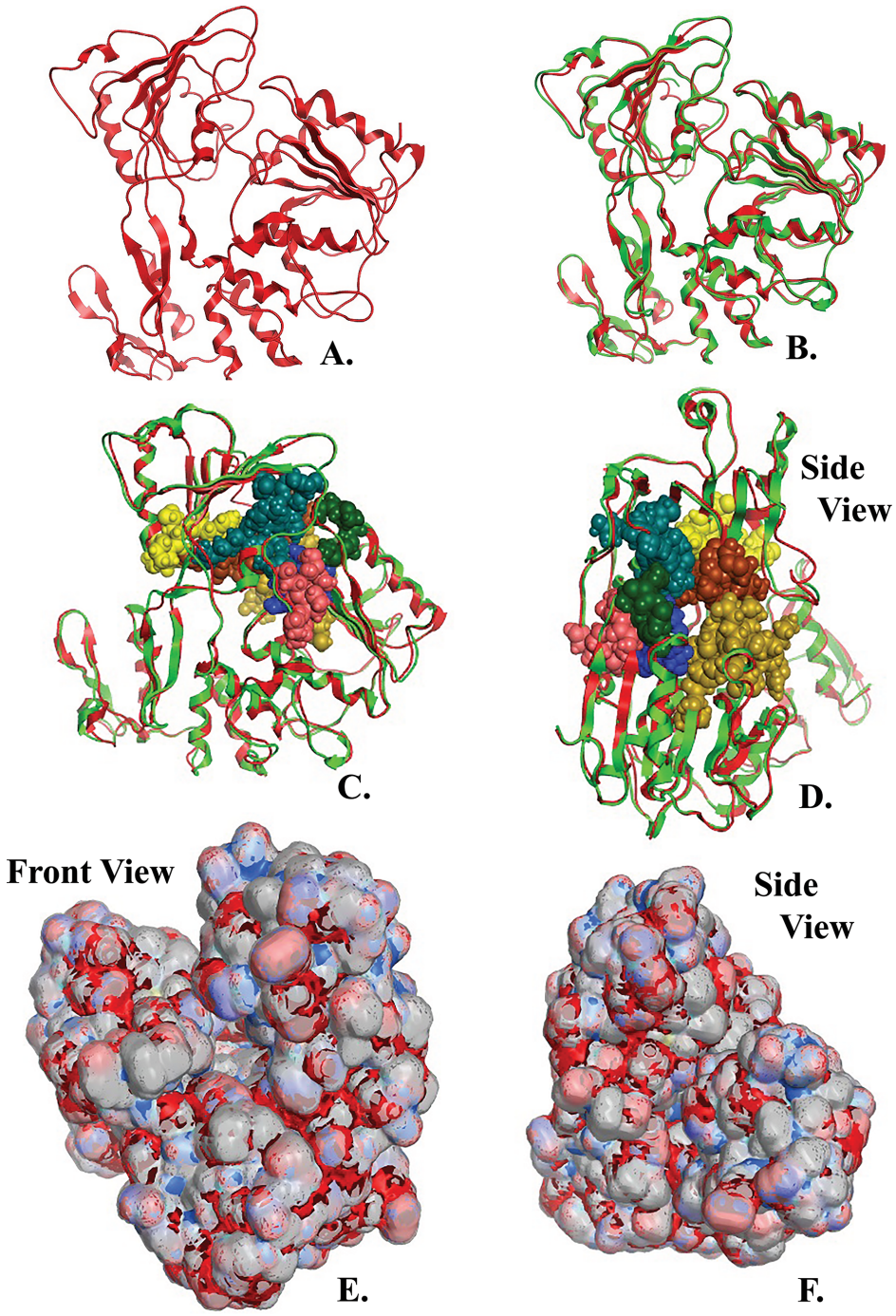

Figure 4 NS3 Model of the Greek Goat Encephalitisvirus helicase. (A) Ribbon representation of the produced Greek Goat Encephalitis virus helicase model. (B) Ribbon representation of the produced Greek Goat Encephalitis virus helicase model (colored Red) next to the corresponding Yellow Fever virus helicase (in green color). (C and D) The conserved motifs of the Yellow Fever virus helicase (in green color) next to the corresponding motifs on the Greek Goat Encephalitis helicase model (colored in Red). The major motifs have been color-coded according to the conventions of Fig. 3, and are shown in CPK format (Usual space filling) along with the rest of the helicase motifs. (E and F) Electrostatic surface potential for the Greek Goat Encephalitis helicase. Represented with red is the area of negative charge. Blue is the area of positive charge and white is the un-charged region.

motifs are involved in coupling the NTPase to RNA duplex unwinding by an unknown mechanism (*von Hippel & Delagoutte, 2003*). One of the most critical motifs in *Flaviviridae* helicases is the GxGKT/S motif 1 in domain 1, which in GGE is conserved in the same sequence position (loop) as in kinases. Motif 1 is also known as Walker A motif and its role involves binding of the β-phosphate of adenosine triphosphate (ATP) (*Saraste, Sibbald & Wittinghofer, 1990*). Based on site directed mutagenesis studies for Motif 1, the mutant protein is inactive. Another crucial conserved motif for the GGE helicase is the DExH motif 2. The DExH motif also known as Walker B motif mediates the binding of the Mg2+-ATP substrate. The aspartate (Asp 170) was found to bind the Mg2+ and create the best conditions for nucleophilic attack through the optimum orientation of ATP (*Ruff et al., 1991*). Furthermore, motif 2 is expected to bind a bivalent cation associated with nucleotide γ-phosphate. Finally, another critical motif is the QRxGRxGR motif 5 in domain 2 which is right across the inter-domain cleft from the DExH motif 2. Motif 5 is important to the *Flaviviridae* helicase function as it is involved in nucleic acid binding (*Niedenzu et al., 2001*).

The Greek Goat Encephalitis helicase model showed a very similar topology to its template, Yellow Fever helicase (Figs. 4B–4D). In general, the *Flaviviridae* viral helicases contain three domains, which are separated by two channels. Domains 1 and 3 have more interactions with each other than they have with domain 2. During the RNA chain unwinding, domain 2 is expected to undergo crucial movement rather than the other two domains. The channel between domains 1–2 and 3 accommodates the ssRNA during the viral RNA unwinding process. The helicase binds RNA in the specific arginine-rich site contained in the second domain (*Luo et al., 2008*). Specific contacts between domains 1 and 2 form the helicase motifs. In the NTPase catalytic site these contacts are expected to change, as the domains link together. Motifs 1a and 6 are in direct contact as the two domains are topologically closer. The linker connecting domains 1 and 2 is motif 3. Finally motif 7 forms direct contacts with motifs 1, 2, and 3. The non-structural protein 3 (NS3) helicase has been reported to interact with other viral replication proteins including non-structural protein 5 (NS5) (*Filomatori et al., 2006*).

Description of the Greek Goat Encephalitis RdRP model

The Greek Goat Encephalitis RdRP secondary structure (data not shown) was predicted from its sequence alignment (Fig. 2). The Greek Goat Encephalitis NS5 model produced of ~900 amino acids contained two main regions, the N-terminal S-adenosyl-L-methyltransferase dependent methyltransferase (MTase) region and the C-terminal RdRP region (Figs. 5A, 5B and 5D). The MTase region bears a 5′ cap structure and is equally important as RdRP (*Geiss et al., 2009*). The MTase plays key roles in the capping process and is believed to catalyze the methylation steps and act as a guanylytrasferase to form the linkage (*Ray et al., 2006*). In the Greek Goat Encephalitis model a linker connects the MTase region and the RdRP region (Figs. 5A, 5B, 5D, 6A1 and 6B). This connection is presented in the Japanese Encephalitis NS5 RdrRP, which is the template model of GGE as well as in other previous studies (*Lu & Gong, 2013*).

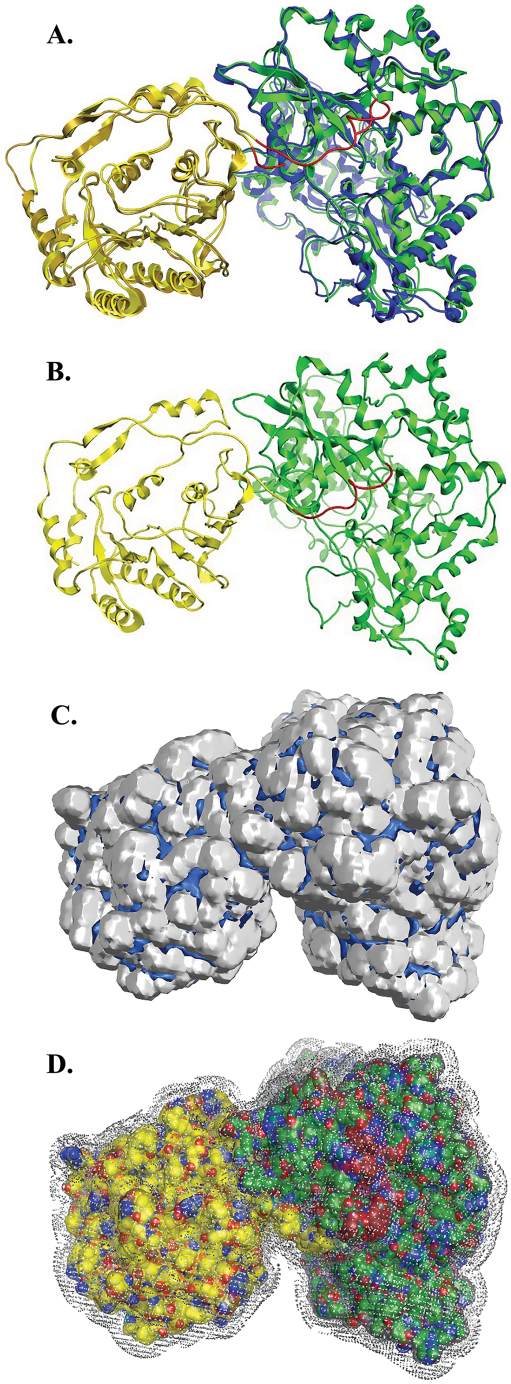

Figure 5 NS5 Model of the Greek GoatEncephalitis virus RdRP. (A) Ribbon representation of the produced Greek Goat Encephalitis virus RdRP and N-terminal S-adenosyl-L-methyltransferase dependent methyltransferase (MTase) side view model (colored Yellow and green) next to the corresponding Japanese Encephalitis virus RdRP and MTase (in gold and blue color). The linker is shown in red. (B) Ribbon representation of the produced Greek Goat Encephalitis virus RdRP (colored green) and MTase (colored yellow) side view complex model. The linker is shown in red. (C and D) Electrostatic surface potential side view for the Greek Goat Encephalitis RdRP and MTase complex.

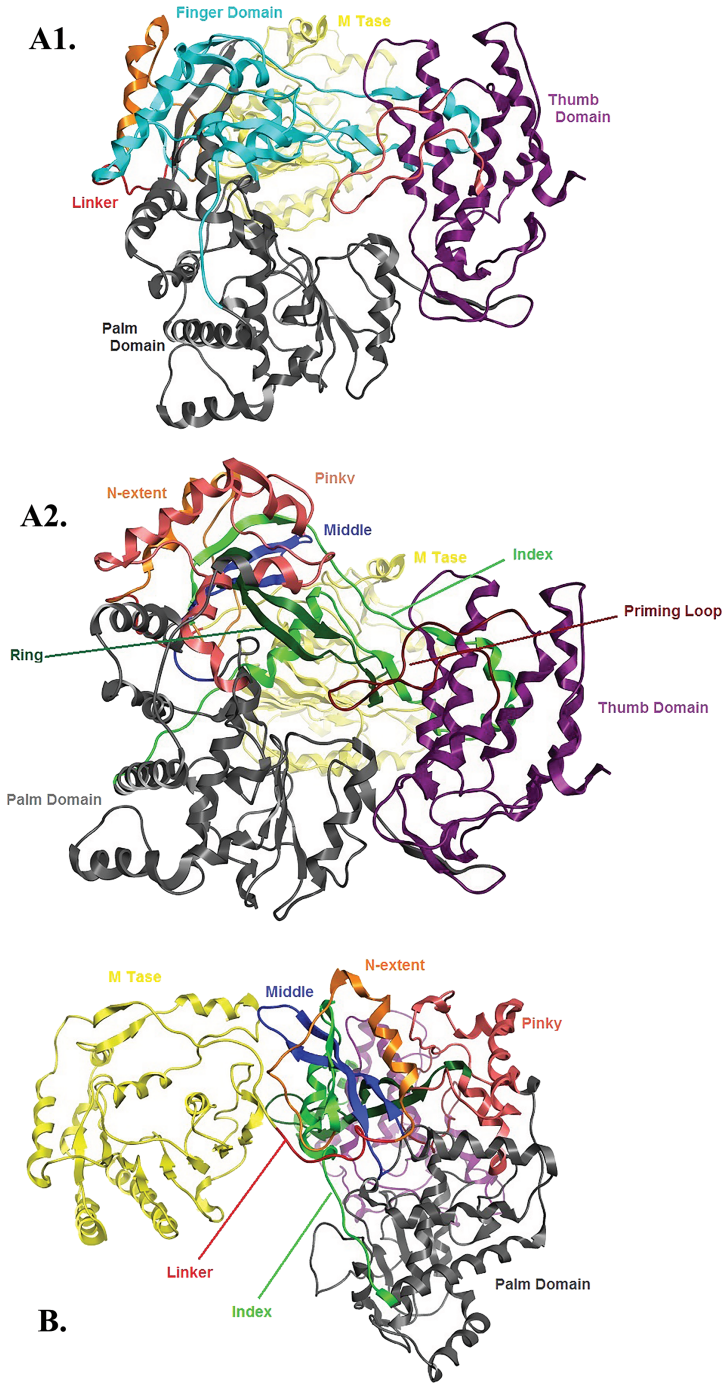

Figure 6 Global views of Greek Goat Encephalitis NS5 RdRP shown in different orientations. (A1) Ribbon representation of the produced Greek Goat Encephalitisvirus RdRP model, looking into the RdRP front channel (putative dsRNAchannel). The three main domains of RdRP are shown: thumb (colored violet), palm (gray) and finger (light blue), with the MTase region at the back (yellow). The linker is shown in red. (A2) Ribbon representation of the produced Greek Goat Encephalitis virus RdRP model with more details about the finger domain subdomains, looking into the RdRP front channel (putative dsRNA channel); (B) Sideview of Ribbon representation of the produced Greek Goat Encephalitis virus RdRP model with more details about the finger domain subdomains. The linker between MTase and RdRP is shown in red.

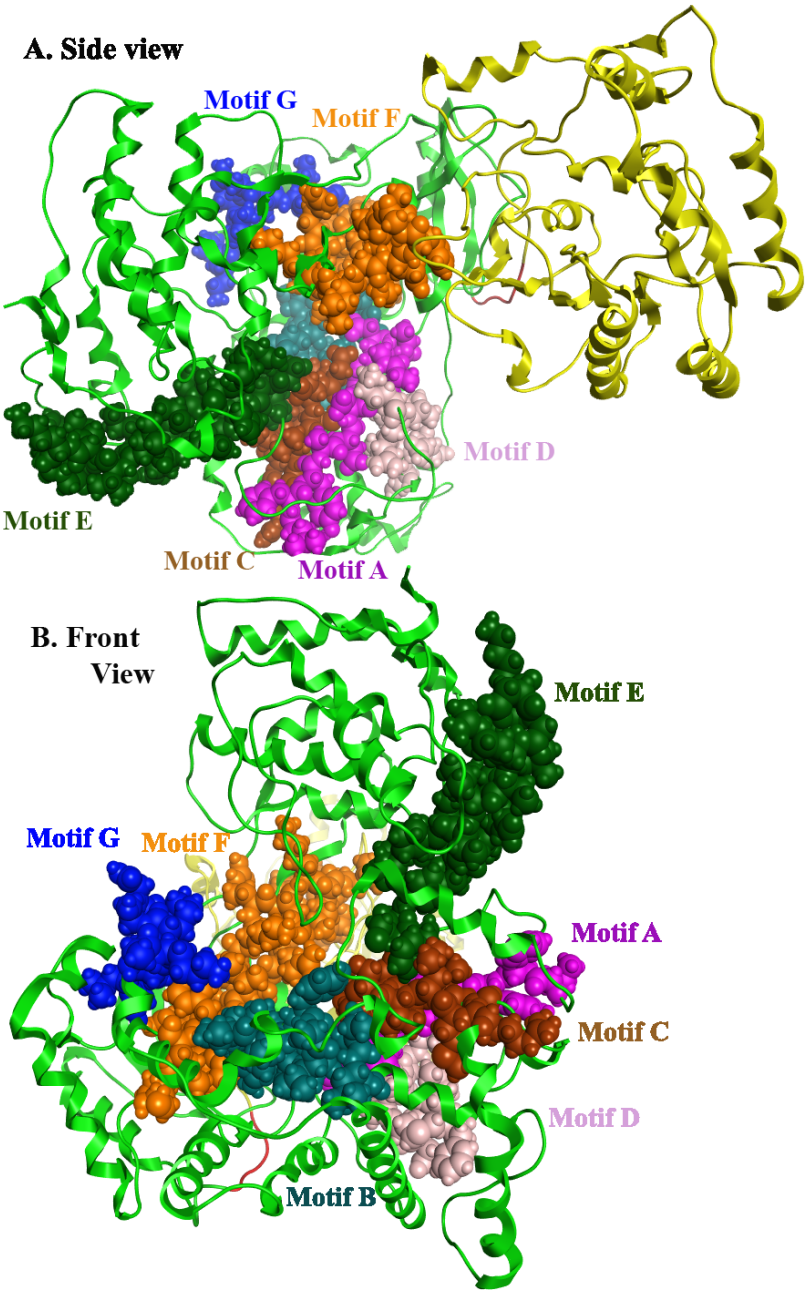

Figure 7 Global views of Greek Goat Encephalitis NS5 RdRP shown in different orientations. The major motifs have been color-coded according to the conventions of Fig. 2, and are shown in CPK format (Usual space filling).

Flaviviridae non structural protein 5 (NS5) comprises eight conserved motifs (*Lu & Gong, 2013*) (Fig. 7). The three domains of Greek Goat Encephalitis RdRP were found to be structurally conserved, as well as the various motifs and regions of *Flaviviridae* NS5 (Fig. 2). As with other viral *Flaviviridae* RdRP, Greek Goat Encephalitis RdRP is separated into the N-terminal extension, the main polymerase and the priming loop. The main

polymerase adopts a shape analogous to a cupped right hand and contains the finger and thumb domains, which rise on the sides of the palm domain (Fig. 6A1). The most conserved part of viral RdRP is the palm domain, which contains the motifs A, B, C and D (*Lu & Gong, 2013*). In motifs A and B there are two catalytic aspartic acid residues (D536 and D668) that are conserved in all single-subunit possessive polymerases (*Gong & Peersen, 2010*). The finger domain contains the motifs F and G. Despite the fact that these motifs have not been resolved, these RdRP structures exhibit high similarity to the NS5B that are represented in the other two genera of the *Flaviviridae* family (*Gong & Peersen, 2010*). The thumb domain of Greek Goat Encephalitis viral RdRP is relatively divergent and contains only Motif E. The size of the thumb domain is larger than in other viral *Flaviviridae* RdRPs because it carries additional elements (C-terminal extension) that facilitate *de novo* initiation.

The Greek Goat encephalitis RdRP finger domain is separated into individual subdomains according to Japanese Encephalitis (*Lu & Gong, 2013*) and Powassan Encephalitis RdRP (*Thompson & Peersen, 2004*). In the finger domain four different subdomains have been marked, the index finger, the middle finger, the ring finger and the pinky finger (Figs. 6A2 and 6B). The index finger comprises a nuclear localization signal that coincides with the recommended binding site of the non-structural protein 3 (NS3) (*Rawlinson et al., 2009*). Moreover the thumb domain interacts with the tip of the index finger and forms a unique encircled active site. The second and the third strands of the fingers domain are in the middle finger area. No specific function has been related with the middle finger area yet. The forth and the fifth strands of the fingers domain are described as ring finger and this area includes the NTP binding of motif F. Finally the pinky finger contains motif G, which is relatively bulky and forms one side of the dsRNA channel (*Gong & Peersen, 2010*).

CONCLUSIONS

The 3D models of the Greek Goat Encephalitis viral enzymes were designed using homology modelling. The X-ray crystal structure of the viral Yellow Fever helicase was employed as a template for the Greek Goat Encephalitis viral helicase and the Japanese Encephalitis RNA dependent RNA polymerase as a template for the Greek Goat Encephalitis RNA dependent RNA polymerase. The models were evaluated and display high conservation of the functional domains previously characterized in other *Flaviviridae* species. We therefore nominate our Greek Goat Encephalitis enzymes models to be suitable for advanced, *in silico* de-novo drug design experiments. In the future this drug discovery process may lead to the development of potential inhibitor molecules.

ACKNOWLEDGEMENTS

We are indebted to the CamGrid computational resource on which the phylogenetic analyses were performed. LP is grateful to Prof. Maroulis Dimitrios for his continuous support throughout his PhD research.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was partially supported by The BIOEXPLORE research project, which falls under the Operational Program "Education and Lifelong Learning" and is co-financed by the European Social Fund (ESF) and National Resources and ESF and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF)—Research Funding Program: "Thales" Investing in knowledge society through the European Social Fund. The authors are members of the BM1006 COST Action, SeqAhead: Next Generation Sequencing Data Analysis Network. VLK is funded by a Marie Curie Intra European Fellowship within the 7th European Community Framework Programme. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
BIOEXPLORE research project.
ESF and National Resources.
Greek national funds.
COST Action: BM1006.
Marie Curie Intra European Fellowship.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Louis Papageorgiou performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables.
- Styliani Loukatou performed the experiments, analyzed the data, wrote the paper.
- Vassiliki Lila Koumandou performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Wojciech Makałowski and Sophia Kossida conceived and designed the experiments, reviewed drafts of the paper.
- Vasileios Megalooikonomou conceived and designed the experiments, analyzed the data, reviewed drafts of the paper.
- Dimitrios Vlachakis conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.664#supplemental-information.

REFERENCES

Abascal F, Zardoya R, Posada D. 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**:2104–2105 DOI 10.1093/bioinformatics/bti263.

Ataide Dias R, Mahon G, Dore G. 2008. EU sheep and goat population in December 2007 and production forecasts for 2008. *Eurostat, European Commission* **67**.

Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C. 2002. The Protein Data Bank. *Acta Crystallographica Section D Biological Crystallography* **58**:899–907 DOI 10.1107/S0907444902003451.

Berman HM, Coimbatore Narayanan B, Di Costanzo L, Dutta S, Ghosh S, Hudson BP, Lawson CL, Peisach E, Prlic A, Rose PW, Shao C, Yang H, Young J, Zardecki C. 2013. Trendspotting in the Protein Data Bank. *FEBS Letters* **587**:1036–1045 DOI 10.1016/j.febslet.2012.12.029.

Bujnicki J, Rychlewski L, Fischer D. 2002. Fold-recognition detects an error in the Protein Data Bank. *Bioinformatics* **18**:1391–1395 DOI 10.1093/bioinformatics/18.10.1391.

Choi KH, Groarke JM, Young DC, Kuhn RJ, Smith JL, Pevear DC, Rossmann MG. 2004. The structure of the RNA-dependent RNA polymerase from bovine viral diarrhea virus establishes the role of GTP in de novo initiation. *Proceedings of the National Academy of Sciences of the United States of America* **101**:4425–4430 DOI 10.1073/pnas.0400660101.

Dalkas GA, Vlachakis D, Tsagkrasoulis D, Kastania A, Kossida S. 2013. State-of-the-art technology in modern computer-aided drug design. *Briefings in Bioinformatics* **14**:745–752 DOI 10.1093/bib/bbs063.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**:1792–1797 DOI 10.1093/nar/gkh340.

Eisenberg D, Luthy R, Bowie JU. 1997. VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods in Enzymology* **277**:396–404.

Filomatori CV, Lodeiro MF, Alvarez DE, Samsa MM, Pietrasanta L, Gamarnik AV. 2006. A 5′ RNA element promotes dengue virus RNA synthesis on a circular genome. *Genes and Development* **20**:2238–2249 DOI 10.1101/gad.1444206.

Geiss BJ, Thompson AA, Andrews AJ, Sons RL, Gari HH, Keenan SM, Peersen OB. 2009. Analysis of flavivirus NS5 methyltransferase cap binding. *Journal of Molecular Biology* **385**:1643–1654 DOI 10.1016/j.jmb.2008.11.058.

Gong P, Peersen OB. 2010. Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America* **107**:22505–22510 DOI 10.1073/pnas.1007626107.

Gorbalenya AE, Koonin EV. 1993. Helicases: amino acid sequence comparisons and structure-function relationships. *Current Opinion in Structural Biology* **3(3)**:412–429 DOI 10.1016/S0959-440X(05)80116-2.

Guex N, Peitsch MC. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**:2714–2723 DOI 10.1002/elps.1150181505.

Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. 1996. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR* **8**:477–486 DOI 10.1007/BF00228148.

Loukatou S, Papageorgiou L, Fakourelis P, Filtisi A, Polychronidou E, Bassis I, Megalooikonomou V, Makalowski W, Vlachakis D, Kossida S. 2014. Molecular dynamics simulations through GPU video games technologies. *Journal of Molecular Biochemistry* **3**:64–71.

Lu G, Gong P. 2013. Crystal structure of the full-length Japanese encephalitis virus NS5 reveals a conserved methyltransferase-polymerase interface. *PLoS Pathogens* **9**:e1003549 DOI 10.1371/journal.ppat.1003549.

Luo D, Xu T, Watson RP, Scherer-Becker D, Sampath A, Jahnke W, Yeong SS, Wang CH, Lim SP, Strongin A, Vasudevan SG, Lescar J. 2008. Insights into RNA unwinding and ATP hydrolysis by the flavivirus NS3 protein. *EMBO Journal* **27**:3209–3219 DOI 10.1038/emboj.2008.232.

Luthy R, Bowie JU, Eisenberg D. 1992. Assessment of protein models with three-dimensional profiles. *Nature* **356**:83–85 DOI 10.1038/356083a0.

Maddison WP, Maddison DR. 1989. Interactive analysis of phylogeny and character evolution using the computer program MacClade. *Folia Primatologica* **53**:190–202 DOI 10.1159/000156416.

Niedenzu T, Roleke D, Bains G, Scherzinger E, Saenger W. 2001. Crystal structure of the hexameric replicative helicase RepA of plasmid RSF1010. *Journal of Molecular Biology* **306**:479–487 DOI 10.1006/jmbi.2000.4398.

Noble CG, Lim SP, Chen YL, Liew CW, Yap L, Lescar J, Shi PY. 2013. Conformational flexibility of the Dengue virus RNA-dependent RNA polymerase revealed by a complex with an inhibitor. *Journal of Virology* **87**:5291–5295 DOI 10.1128/JVI.00045-13.

Nulf CJ, Corey D. 2004. Intracellular inhibition of hepatitis C virus (HCV) internal ribosomal entry site (IRES)-dependent translation by peptide nucleic acids (PNAs) and locked nucleic acids (LNAs). *Nucleic Acids Research* **32**:3792–3798 DOI 10.1093/nar/gkh706.

O'Farrell D, Trowbridge R, Rowlands D, Jager J. 2003. Substrate complexes of hepatitis C virus RNA polymerase (HC-J4): structural evidence for nucleotide import and de-novo initiation. *Journal of Molecular Biology* **326**:1025–1035 DOI 10.1016/S0022-2836(02)01439-0.

Papagelopoulos N, Vlachakis D, Filtisi A, Fakourelis P, Papageorgiou L, Megalooikonomou V, Kossida S. 2014. State of the art GPGPU applications in bioinformatics. *International Journal of Systems Biology and Biomedical Technologies* **2**:24–48 DOI 10.4018/ijsbbt.2013100103.

Papageorgiou L, Vlachakis D, Koumandou VL, Papagelopoulos N, Kossida S. 2014. Computer-aided drug design and biological evaluation of novel anti-Greek Goat Encephalitis agents. *International Journal of Systems Biology and Biomedical Technologies* **2**:1–16 DOI 10.4018/ijsbbt.2013100101.

Pavlidou V, Gerou S, Diza E, Antoniadis A, Papa A. 2008. Genetic study of the distribution of Greek goat encephalitis virus in Greece. *Vector-borne and Zoonotic Diseases* **8**:351–354 DOI 10.1089/vbz.2007.0215.

Pickett BE, Sadat EL, Zhang Y, Noronha JM, Squires RB, Hunt V, Liu M, Kumar S, Zaremba S, Gu Z, Zhou L, Larson CN, Dietrich J, Klem EB, Scheuermann RH. 2012. ViPR: an open bioinformatics database and analysis resource for virology research. *Nucleic Acids Research* **40**:D593–D598 DOI 10.1093/nar/gkr859.

Rawlinson SM, Pryor MJ, Wright PJ, Jans DA. 2009. CRM1-mediated nuclear export of dengue virus RNA polymerase NS5 modulates interleukin-8 induction and virus production. *Journal of Biological Chemistry* **284**:15589–15597 DOI 10.1074/jbc.M808271200.

Ray D, Shah A, Tilgner M, Guo Y, Zhao Y, Dong H, Deas TS, Zhou Y, Li H, Shi PY. 2006. West Nile virus 5′-cap structure is formed by sequential guanine N-7 and ribose 2′-O methylations by nonstructural protein 5. *Journal of Virology* **80**:8362–8370 DOI 10.1128/JVI.00814-06.

Ruff M, Krishnaswamy S, Boeglin M, Poterszman A, Mitschler A, Podjarny A, Rees B, Thierry JC, Moras D. 1991. Class II aminoacyl transfer RNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with tRNA(Asp). *Science* **252**:1682–1689 DOI 10.1126/science.2047877.

Saraste M, Sibbald PR, Wittinghofer A. 1990. The P-loop–a common motif in ATP- and GTP-binding proteins. *Trends in Biochemical Sciences* **15**:430–434 DOI 10.1016/0968-0004(90)90281-F.

Thompson AA, Peersen OB. 2004. Structural basis for proteolysis-dependent activation of the poliovirus RNA-dependent RNA polymerase. *EMBO Journal* **23**:3462–3471 DOI 10.1038/sj.emboj.7600357.

Vlachakis D, Argiro A, Kossida S. 2013. An update on virology and emerging viral epidemics. *Journal of Molecular Biochemistry* **2**:80–84.

Vlachakis D, Bencurova E, Papangelopoulos N, Kossida S. 2014. Current state-of-the-art molecular dynamics methods and applications. *Advances in Protein Chemistry and Structural Biology* **94**:269–313.

Vlachakis D, Karozou A, Kossida S. 2013. 3D molecular modelling study of the H7N9 RNA-dependent RNA polymerase as an emerging pharmacological target. *Influenza Research and Treatment* **2013**:645348 DOI 10.1155/2013/645348.

Vlachakis D, Kontopoulos DG, Kossida S. 2013. Space constrained homology modelling: the paradigm of the RNA-dependent RNA polymerase of dengue (type II) virus. *Computational and Mathematical Methods in Medicine* **2013**:108910.

Vlachakis D, Kossida S. 2013. Molecular modeling and pharmacophore elucidation study of the Classical Swine Fever virus helicase as a promising pharmacological target. *PeerJ* **1**:e85 DOI 10.7717/peerj.85.

Vlachakis D, Koumandou VL, Kossida S. 2013. A holistic evolutionary and structural study of flaviviridae provides insights into the function and inhibition of HCV helicase. *PeerJ* **1**:e74 DOI 10.7717/peerj.74.

Vlachakis D, Pavlopoulou A, Tsiliki G, Komiotis D, Stathopoulos C, Balatsos NA, Kossida S. 2012. An integrated in silico approach to design specific inhibitors targeting human poly(a)-specific ribonuclease. *PLoS ONE* **7**:e51113 DOI 10.1371/journal.pone.0051113.

Vlachakis D, Tsaniras SC, Feidakis C, Kossida S. 2013a. Molecular modelling study of the 3D structure of the biglycan core protein, using homology modelling techniques. *Journal of Molecular Biochemistry* **2**:85–93.

Vlachakis D, Tsaniras SC, Kossida S. 2013. Current viral infections and epidemics of flaviviridae; lots of grief but also some hope. *Journal of Molecular Biochemistry* **3(1)**:144–149.

Vlachakis D, Tsiliki G, Kossida S. 2013. *3D molecular modelling of the helicase enzyme of the endemic, zoonotic Greek Goat Encephalitis virus.* Vol. 383. Springer, 165–171 DOI 10.1007/978-3- 642-41013-0_17.

Vlachakis D, Tsiliki G, Pavlopoulou A, Roubelakis MG, Tsaniras SC, Kossida S. 2013b. Antiviral stratagems against HIV-1 using RNA interference (RNAi) technology. *Evolutionary Bioinformatics Online* **9**:203–213 DOI 10.4137/EBO.S11412.

von Hippel PH, Delagoutte E. 2003. Macromolecular complexes that unwind nucleic acids. *Bioessays* **25**:1168–1177 DOI 10.1002/bies.10369.

Wang J, Cieplak P, Kollman P. 2000. How well does a restrained electrostatic potential (resp) model perform in calculating conformational energies of organic and biological molecules. *Journal of Computational Chemistry* **21**:1049–1071 DOI 10.1002/1096-987X(200009)21:12<1049::AID-JCC3>3.0.CO;2-F.

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**:1189–1191 DOI 10.1093/bioinformatics/btp033.

Wu J, Bera AK, Kuhn RJ, Smith JL. 2005. Structure of the Flavivirus helicase: implications for catalytic activity, protein interactions, and proteolytic processing. *Journal of Virology* **79**:10268–10277 DOI 10.1128/JVI.79.16.10268-10277.2005.