Analysis of methyltransferase (MTase) domain from Zika virus (ZIKV)

Sarah Afaq1,6,5, Akhtar Atiya2, Arshi Malik1, Afaf S. Alwabli3, Dhafer A. Alzahrani3, Habeeb M. Al-Solami3, Othman Alzahrani1, Qamre Alam5, Mohammad Azhar Kamal6,7, Aala A. Abulfaraj8, Alawiah M. Alhebshi9 & Mohammed Tarique10,5

1Department of Clinical Biochemistry, College of Medicine, King Khalid University, Abha, Kingdom of Saudi Arabia; 2Department of Pharmacognosy, College of Pharmacy, King Khalid University, Abha, Kingdom of Saudi Arabia; 3Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah 21589, Kingdom of Saudi Arabia; 4Department of Biology, Faculty of Science, University of Tabuk, Tabuk, Kingdom of Saudi Arabia; 5Medical Genomics Research Department, King Abdullah International Medical Research Center, King Saud bin Abdulaziz University for Health Sciences, Ministry of National Guard Health Affairs, Riyadh, Kingdom of Saudi Arabia; 6University of Jeddah, Faculty of Science, Department of Biochemistry, Jeddah, Kingdom of Saudi Arabia; 7University of Jeddah Center for Science and Medical Research (UJC-SMR), Jeddah, Kingdom of Saudi Arabia; 8Department of Biological Sciences, College of Sciences and Arts-Rabigh Campus, King Abdulaziz University, Jeddah 21589, Saudi Arabia; 9Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah 21589, Kingdom of Saudi Arabia; 10Center for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, Jamia Nagar, New Delhi-110025, India; Dr. Sarah Afaq - Email: safaq@kku.edu.sa; Dr. Mohammad Tarique - E-mail id: tariqueaims@gmail.com; *Equal contribution; *Corresponding author

Received October 22, 2019; Revised February 15, 2020, Accepted February 20, 2020; Published March 31, 2020

DOI: 10.6026/97320630016229

Declaration on official E-mail:
The corresponding author declares that official e-mail from their institution is not available for all authors

Declaration on Publication Ethics:
The authors state that they adhere with COPE guidelines on publishing ethics as described elsewhere at https://publicationethics.org/. The authors also undertake that they are not associated with any other third party (governmental or non-governmental agencies) linking with any form of unethical issues connecting to this publication. The authors also declare that they are not withholding any information that is misleading to the publisher in regard to this article.

Note:
The editorial board and the Publisher has taken reasonable steps where possible to check and evaluate the data provided by the authors in this report.
Abstract:
A comprehensive analysis of methyltransferase (MTase) from Zika virus (ZIKV) is of interest in the development of drugs and biomarkers in the combat and care of ZIKA fever with impulsive joint pain and conjunctivitis. MTase sequence is homologous in several viral species. We analyzed the MTase domain from ZIKV using Bioinformatics tools such as SMART, PROSITE, PFAM, PANTHER, and InterProScan to glean insights on the sequence to structure to function data. We document inclusive information on MTase from ZIKV for application in the design of drugs and biomarkers to fight against the disease.

Keywords: ZIKV, methyltransferase, beta turn, α-helix, SMART, Prosite, Pfam and InterProScan

Background:
The Flavivirus Zika virus (ZIKV), which was announced as a Public Health Emergency by the World Health Organization (WHO) on February first, 2016. ZIKV is a virus from the Flaviviridae family and the Aedes mosquitoes acts as carriers [1-3]. The ZIKV genome is at first converted into a solitary precursor protein assembled into nonstructural proteins (1, 2A, 2B, 3, 4A, 4B and 5) [4]. The NS5 is the most significant and most monitored protein that contains methyltransferase (MTase) at the N-terminal region and an RNA-dependent RNA polymerase (RdRP) at C terminal region. It function in the replication of viral genome with RdRP and with methyltransferase domain separately [5]. Crystal structures of the MTase domain demonstrated that it possess the characteristic α/β overlay as found in the Dengue virus (DENV) [6], Japanese encephalitis (JEV) infection [7], or the ongoing structures of Zika virus (ZIKV) (5KQR) [8]. A sequence similarity among different flaviviruses is around 60%. The MTase domain NS5 protein consists of three subdomains. Initially, the C-terminal end has the conserved MTase crease framed by 7 strands β-sheet encompassed by 4 α-helices. In some structures an SAH (S-adenosyl-L-homocysteine) molecule is discovered bound to this domain [9]. The second subdomain contains a helix-turn-helix theme, a β-strand and a α-helix structure at its N-terminals end. This domain was proposed to organize the GTP (guanosine-5’-triphosphate) moiety of 7-methylguanosine-GTP amid the 2'-O-ribose methylation as observed in the crystal structures bound to m7Gppp-RNA (7-methylguanosine cap at the 5’ end of mRNA) [9]. The 3rd subdomain is situated between the two previous ones and is made out of an α-helix and two β strands [10]. Therefore, it is of interest to document broad information on MTase from ZIKV for application in the fight against the disease.

Methodology:
Sequence and conserved domain analysis MTase:
The sequence of the MTase domain was retrieved from the NCBI genome database followed by protein BLAST (BLASTp) analysis. The sequence was further subject to SMART, Prosite, Pfam, PANTHER, and InterProScan as described elsewhere [11-13].

Analysis of predicted secondary structure:
The secondary structures were assigned using PSIPRED available at http://bioinf.cs.ucl.ac.uk/psipred.

Figures 1: The detail domain organization of Zika virus (ZIKV). The conserved sequences of two important domains (MTase and RdRP) are written inside the boxes and highlighted. The text in purple and orange color box refers to the names of conserved domains and the numbers refer to the amino acids sequence. In the box important active site amino acid highlighted in bold with different colors.

Epitope prediction:
Epitope prediction was completed using the tool at http://tools.immuneepitope.org as described elsewhere [14].
Structure analysis

The sequence was further analyzed for structural features such as beta turns, helices and disallowed regions using tools as described elsewhere [15-19].

Figures 2: The protein sequences of MTase domain of Zika virus. A. The sequence was submitted to the server at http://bioinf.cs.ucl.ac.uk/psipred/. The graph represents the strand, helix and coil. B. Conservation of sequence within the specific MTase domain motifs. The peak of the amino acids residues reflects the level of retention at a position, and tall letters represents higher retention.

Figures 3: Prediction of structure (secondary) using server at http://bioinf.cs.ucl.ac.uk/psipred/. The protein sequence of MTase domain of Zika virus was submitted to the server and the secondary structures was determined. The graph represents the structures of MTase domain of Zika virus. Epitope peptide (6) boxed in different color the numbers refer to the amino acids sequence.
Figures 4: The plots for turns demonstrate a Ramachandaran plot with residues i+1 (brown circle) and i+2 (green square) plotted on it. The following is a graphic plot of the turn with the four amino acid residues and marked C alpha (i) C alpha (i+3) distance. A red arrow, if present, indicates that residue I donate a hydrogen bond to residue i+3. The numbers of residue and type of turn are demonstrated over the Ramachandaran plot.

Figures 5: The Helical haggle, and net color diagrams represent the organization of the amino acid residues in every helix. The amino acid residues are in green color for hydrophobic, blue color for polar and red color for charged amino acid. Haggles and nets accepted the helical estimation of 3.6 residues per turn.
Table 1: The turns are doled out to one of 9 classes based on the phi, psi edges of the Ramachandran plot occupied by residues i + 1 and i + 2 of the turn.

| No. | Turn | Sequence | Type | Residue(i+1) | Residue(i+2) | Phi(i+1) | Psi(i+1) | Phi(i+2) | Psi(i+2) | Turn | Identity |
|-----|------|----------|------|--------------|--------------|----------|----------|----------|----------|------|-----------|
| 1   | Lys28-His32 | KGHI | H | -64.5 | 116.6 | -71.1 | 70.6 | 8.5 | -5.5 | No |
| 2   | Val49-Gly53 | VAYT | IV | 1 | -64.4 | - | 15.2 | 55.6 | 6.4 | Yes |
| 3   | Gly98-Val102 | GYLQ | VIII | 11.0 | -58.3 | -59.3 | 116.6 | 132.4 | 171.2 | 0.0 | Yes |
| 4   | Arg99-Gly102 | DGLG | IV | -72.8 | 125.2 | -98.7 | - | 51.6 | - | 5.6 | No |
| 5   | Cys166-Gly181 | CCRG | H | -60.9 | - | 145.8 | 125.7 | 66.0 | 5.8 | No |
| 6   | Asn240-Val247 | IV | -56.0 | -44.2 | - | -75.5 | 11.8 | 72.5 | 5.4 | No |
| 7   | Gly187-Ile190 | GNIH | III | -60.6 | 126.3 | 170.0 | 83.1 | 5.4 | - | 5.6 | No |
| 8   | Ser159-Pro162 | SYQW | II | 93.8 | 255.6 | 307.0 | 185.9 | 11.5 | - | 5.6 | No |
| 9   | Lys277-Val280 | KGIV | IV | -96.1 | 173.4 | 6.4 | 71.5 | 24.8 | - | 6.9 | Yes |
| 10  | Cys184-Leu191 | CDIT | VIII | -93.2 | 37.6 | 178.6 | - | 124.9 | -36.8 | 6.7 | Yes |

No. of beta turns in the Mtase domain from ZIKV

Number of beta turns in chain 17; *Asterisked motifs correspond to those illustrated in the motif plots (Figure 4).

Table 3: Helices in the Mtase domain from ZIKV

| No. | Start | End | Type | Residue | Longevity | Unwinding | Residue | Phi | Deviation from Ideal | Sequence |
|-----|-------|-----|------|---------|-----------|-----------|---------|-----|----------------------|----------|
| 1   | Lys9 | Ala13 | H | H | 15.62 | 1.3 | 3.9 | 5.4 | 2.3 | LLKNNWAK | 2 | 1.3 | 1.3 | LEKNNWAK |
| 2   | Ala12 | Val16 | I | H | 12.52 | 1.8 | 3.7 | 5.0 | 11.2 | ALEKYYE | 2 | 1.3 | 1.3 | LEKNNWAK |
| 3   | Gly18 | Gly22 | H | H | 15.72 | 1.3 | 3.6 | 5.4 | 8.8 | GAKLRE | 1 | 1.3 | 1.3 | LEKNNKAK |
| 4   | Gln28 | Gly32 | H | H | 14.15 | 1.3 | 3.6 | 5.5 | 7.2 | GWSYVAA | 1 | 1.3 | 1.3 | LEKNNKAK |
| 5   | Gln34 | His38 | H | H | 11.26 | 1.3 | 3.6 | 5.5 | 7.2 | GWSYVAA | 1 | 1.3 | 1.3 | LEKNNKAK |
| 6   | Thr13 | His17 | G | G | 15.72 | 1.3 | 3.6 | 5.4 | 8.8 | GAKLRE | 1 | 1.3 | 1.3 | LEKNNKAK |
| 7   | Val7 | Ile11 | H | H | 19.28 | 1.3 | 3.6 | 5.5 | 13.4 | HNPEALE | 2 | 1.4 | 1.4 | LEKNNKAK |
| 8   | Ser18 | Tyr20 | H | H | 28.05 | 1.4 | 3.6 | 5.2 | 6.7 | KTMMMTLRE | 1 | 1.3 | 1.3 | LEKNNKAK |
| 9   | Ser18 | Tyr20 | H | H | 28.05 | 1.4 | 3.6 | 5.2 | 6.7 | KTMMMTLRE | 1 | 1.3 | 1.3 | LEKNNKAK |

Number of helix in chain 17; *Asterisked motifs correspond to those illustrated in the motif plots.

Results and Discussion:
The sequence analysis and domain organization of Mtase domain and (264 amino acids) and RdRp area (149 amino acids) is shown using SMART, Prosite, Pfam, PANTHER, and InterProScan in Figure 1. There is three crucial amino acid arrangement of Mtase domain are in charge of dynamic site restricting which are mRNA top official (K), mRNA top authoritative; using of carbonyl oxygen (L), S-adenosyl-L-methionine (S) and Essential for 2’-O- methyltransferase action (K). Amino acid consensus logo based analysis of the Mtase domain with strands, α-helix, and coil is shown in Figure 2A. Different residues at the same location are scaled on the basis of residue frequency as shown in Figure 2B. Secondary structure and antigenic determinant of the Mtase domain is shown in Figure 3. The major epitope peptides are six that are highlighted in color boxes (Figure 3). Data on beta turns in the Mtase is given in Table 1 and Table 2. Data on helices in the Mtase domain is given in Table 3. Thus, we document inclusive information on Mtase from ZIKV for application through a comprehensive understanding in the design of drugs and biomarkers to fight against the disease caused by the virus.

Conclusion:
We document preliminary information from a comprehensive analysis on Mtase from ZIKV using Bioinformatics tools such as SMART, PROSITE, PFAM, PANTHER, and InterProScan to glean insights on the sequence to structure to function data for combat and care of ZIKI fever.

Conflict of Interests:
There is no conflict of interests among the authors regarding the present publication.

Acknowledgments:
The authors want to thank the Almanac Life Science India Pvt. Ltd. for valuable suggestion in this analysis.

References:
[1] Chen LH & Hamer DH, *Annals of internal medicine* 2016 164:613 [PMID: 26832396].
[2] Triulzi M., *The Lancet Infectious Diseases* 2016 16:156 [PMID: 26723756].
[3] Makar J et al. *New England Journal of Medicine* 2016 374: 951-958 [PMID: 26859296].
[4] Roze B et al. *Eurosurveillance* 2016 21:30154 [PMID: 26967758].
[5] Lindenbach B.D et al. *Journal of Virology* 2007 81:8905 [PMID: 17581983].
[6] Lim SP et al. *Journal of Biological Chemistry* 2011 286:6233 [PMID: 21147775].
[7] Lu G & Gong P, *PLoS pathogens* 2013 9:e1003549 [PMID: 23950717].
[8] Coloma J et al. *Cell reports* 2016 16:3097 [PMID: 27633330].
Articles published in BIOINFORMATION are open for relevant post publication comments and criticisms, which will be published immediately linking to the original article for FREE of cost without open access charges. Comments should be concise, coherent and critical in less than 1000 words.
