BIOINFORMATION
Discovery at the interface of physical and biological sciences

OPEN ACCESS GOLD

BIOMEDICAL INFORMATICS

www.bioinformation.net
Volume 16(3)

Research Article

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

Sarah Afaq[1*,$], Akhtar Atiya[2], Arshi Malik[1], Afaf S. Alwabli[3], Dhafer A. Alzahrani[3], Habeeb M. Al-Solami[3], Othman Alzahrani[4], Qamre Alam[5], Mohammad Azhar Kamal[6,7], Aala A. Abulfaraj[8], Alawiah M. Alhebshi[9] & Mohammed Tarique[10,*,$]

[1]Department of Clinical Biochemistry, College of Medicine, King Khalid University, Abha, Kingdom of Saudi Arabia; [2]Department of Pharmacognosy, College of Pharmacy, King Khalid University, Abha, Kingdom of Saudi Arabia; [3]Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah 21589, Kingdom of Saudi Arabia; [4]Department of Biology, Faculty of Science, University of Tabuk, Tabuk, Kingdom of Saudi Arabia; [5]Medical 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; [6]University of Jeddah, Faculty of Science, Department of Biochemistry, Jeddah, Kingdom of Saudi Arabia; [7]University of Jeddah Center for Science and Medical Research (UJC-SMR), Jeddah, Kingdom of Saudi Arabia; [8]Department of Biological Sciences, College of Sciences and Arts-Rabigh Campus, King Abdulaziz University, Jeddah 21589, Saudi Arabia; [9]Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah 21589, Kingdom of Saudi Arabia; [10]Center 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: tariqueaiims@gmail.com; [$]Equal contribution; *Corresponding author

Arshi Malik - E-mail: amalik@kku.edu.sa; Akhtar Atiya - E-mail: akhan@kku.edu.sa; Afaf S. Alwabli - E-mail: afafalwabli@yahoo.com; Dhafer A. Alzahrani - E-mail: dalzahrani@kau.edu.sa; Habeeb M. Al-Solami - E-mail: hmalsolami@kau.edu.sa; Othman Alzahrani - E-mail: o-alzahrani@ut.edu.sa; Qamre Alam - E-mail: alamqa@ngha.med.sa; Mohammad Azhar Kamal - E-mail: makamal@uj.edu.sa; Aala A. Abulfaraj - E-mail: aaabulfaraj@kau.edu.sa; Alawiah M. Alhebshi- E-mail: aalhebshi@kau.edu.sa; Mohammed Tarique-E-mail: tariqueaiims@gmail.com

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 reasonble steps where possible to check and evaluate the data provided by the authors in this report.

BIOINFORMATION
Discovery at the interface of physical and biological sciences

OPEN ACCESS GOLD

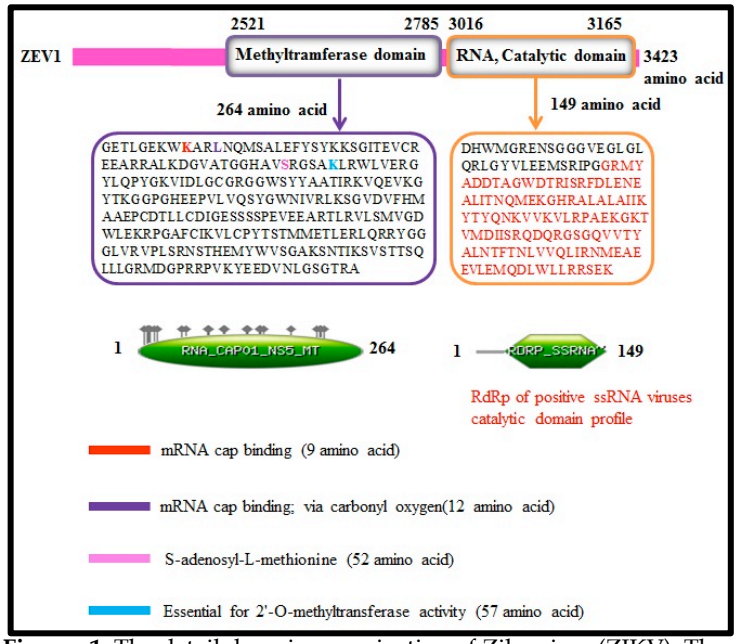
Abstract:
A comprehensive analysis of methyltransferase (MTase) from Zika virus (ZIKV) is of interest in the development of drugs and biomarkers in the combat and care of ZIKA fever with impulsive joint pain and conjunctivitis. MTase sequence is homologous in several viral species. We analyzed the MTase 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 methyltranferase 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 sub-domain 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.



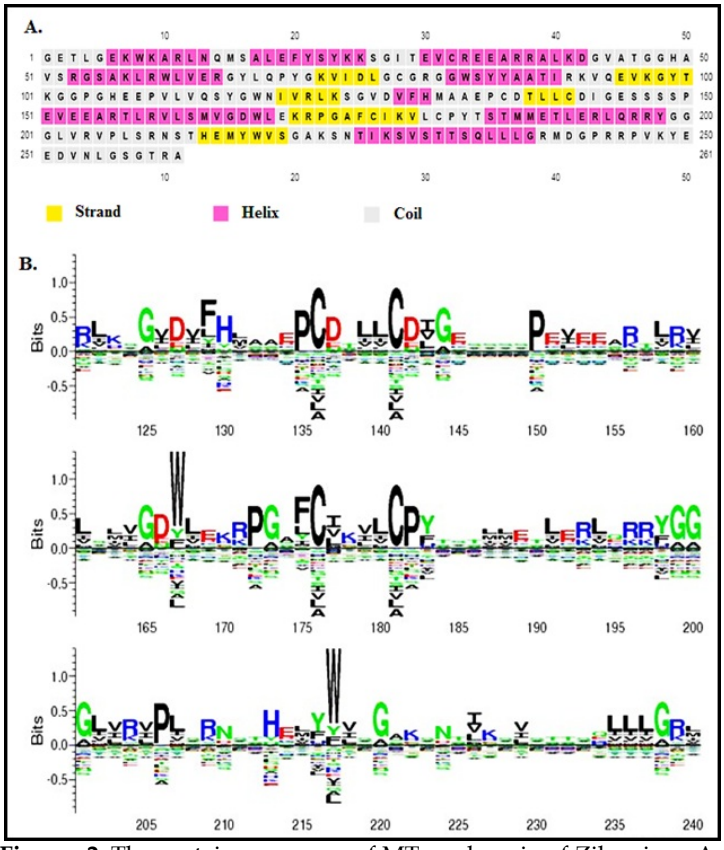
Figures 1: The detail domain organization of Zika virus (ZIKV). The conserved sequences of two important domains (MTase and RdRP) are written inside the boxes and highlighted. The text in purple and orange color box refers to the names of conserved domains and the numbers refer to the amino acids sequence.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

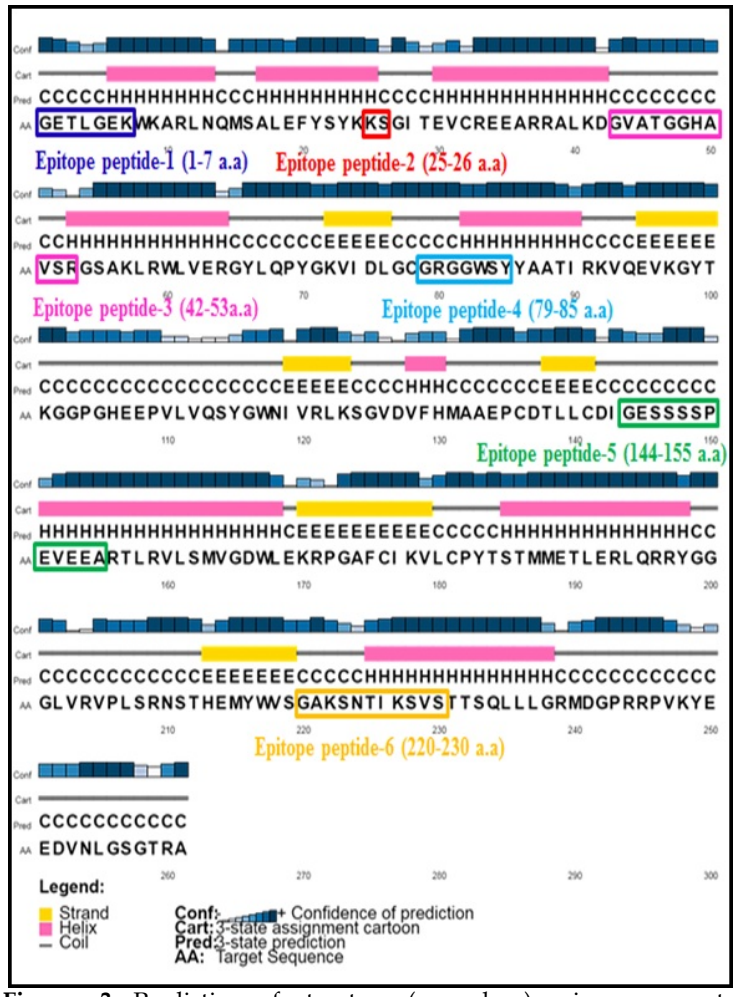
BIOINFORMATION
Discovery at the interface of physical and biological sciences

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]**

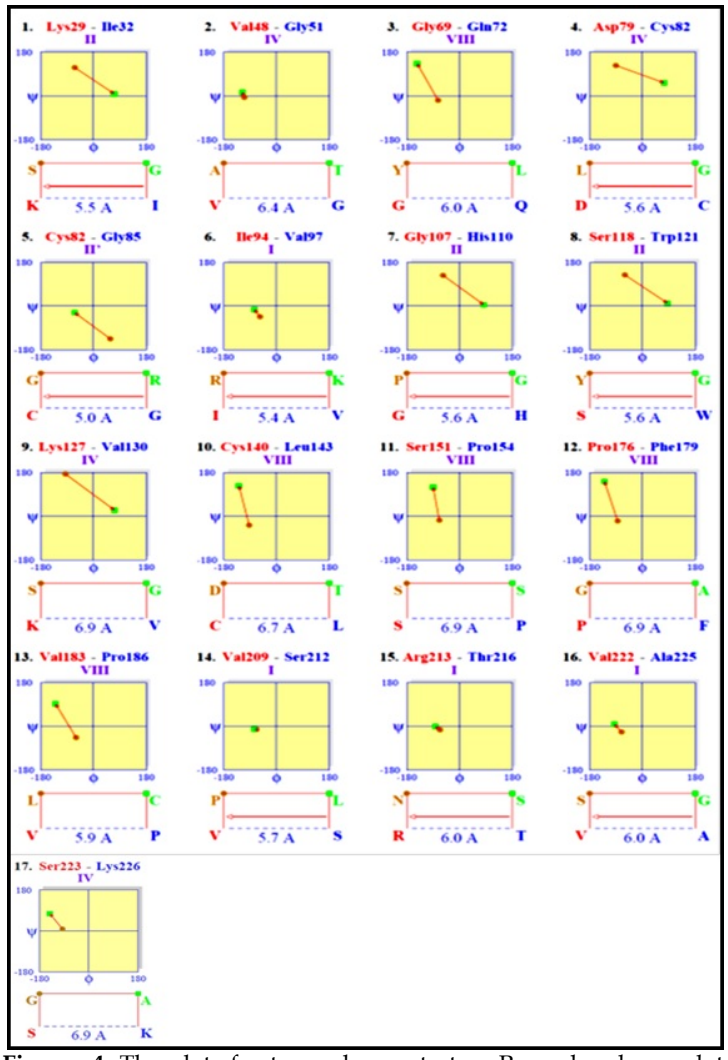


Figures 2: The protein sequences of MTase domain of Zika virus. A. The sequence was submitted to the server at http://bioinf.cs.ucl.ac.uk/psipred/. The graph represents the strand, helix and coil. B. Conservation of sequence within the specific MTase domain motifs. The peak of the amino acids residues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

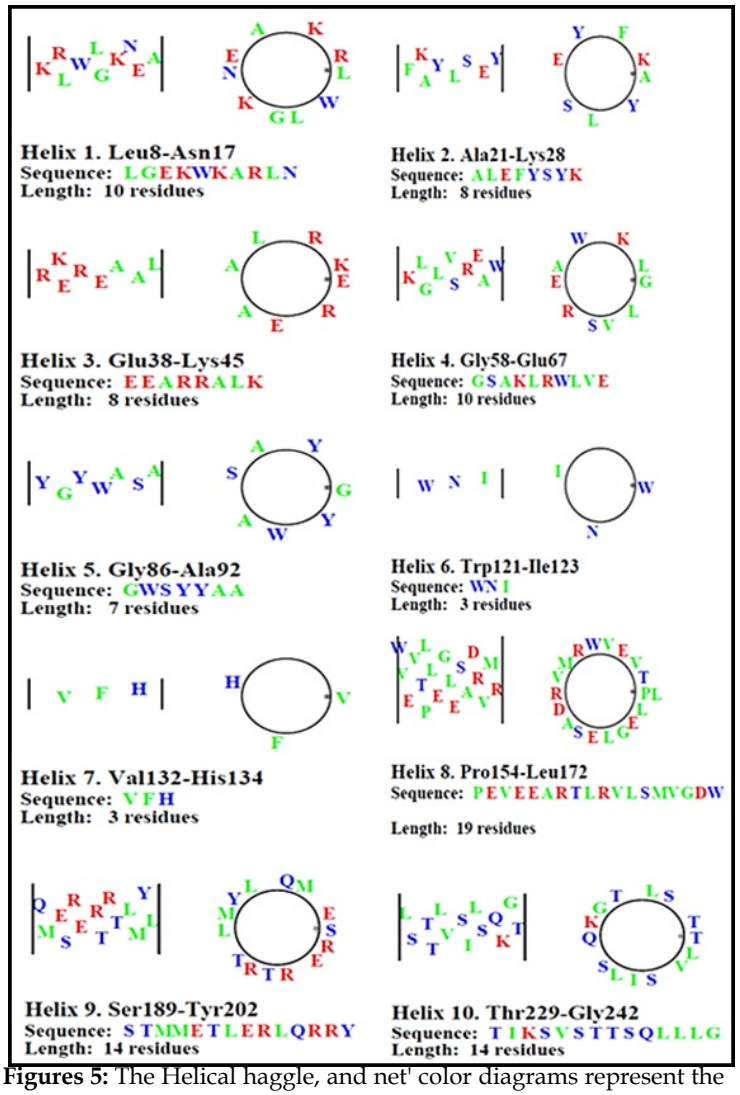


Figures 3: Prediction of structure (secondary) using server at http://bioinf.cs.ucl.ac.uk/psipred/. The protein sequence of MTase domain of Zika virus was submitted to the server and the secondary structures was determined. The graph represents the structures of MTase domain of Zika virus. Epitope peptide (6) boxed in different color the numbers refer to the amino acids sequence.

BIOINFORMATION
Discovery at the interface of physical and biological sciences



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.

BIOINFORMATION
Discovery at the interface of physical and biological sciences

Table 1: The turns are doled out to one of 9 classes based on the phi, psi edges of buildups i+1 and i+2. The perfect plots for every one of the turn types are as per the following:

Type	Phi(i+1)	Psi(i+1)	Phi(i+2)	Psi(i+2)	
I	-60	-30	-90	0	
II	-60	120	80	0	
VIII	-60	-30	-120	120	
I′	60	30	90	0	
II′	60	-120	-80	0	
VIa1	-60	120	-90	0	cis-proline(i+2)
VIa2	-120	120	-60	0	cis-proline(i+2)
VIb	-132	135	-75	160	cis-proline(i+2)
IV	Turns excluded from all the above categories				

The nomenclature describes the regions of the Ramachandaran plot occupied by residues i + 1 and i + 2 of the turn.

Table 2: Beta turns in the MTase domain from ZIKV

No.	Turn	Sequence*	Turn Type	Residue i+1			Residue i+2			I to i+3 CA-dist	H bond
				Phi	Psi	Chi1	Phi	Psi	Chi1		
1.*	Lys29-Ile32	KSGI	II	-64.5	116.6	-77.1	70.6	8.5	-	5.5	No
2.*	Val48-Gly51	VATG	IV	-	-6.4	-	-	15.2	55.6	6.4	Yes
3.*	Gly69-Gln72	GYLQ	VIII	110.1	-18.3	-58.3	116.6	132.4	171.2	6.0	Yes
4.*	Asp79-Cys82	DLGC	IV	-72.8	125.2	-59.7	-	53.6	-	5.6	No
5.*	Cys82-Gly85	CGRG	II	-90.9 57.0	- 134.9	-	143.9 74.3 -64.3	-27.5	-66.0	5.0	No
6.	Ile94-Val97	IRKV	I	-56.0	-44.2	-	-75.7	-13.8	-73.5	5.4	No
7.	Gly107-His110	GPGH	II	-56.0	126.3	173.0	83.1	5.4	-	5.6	No
8.	Ser118-Trp121	SYGW	II	59.8	128.4	-30.7	85.9	11.5	-	5.6	No
9.	Lys127-Val130	KSGV	IV	-96.1	173.4	-	71.5	24.4	-	6.9	Yes
10.	Cys140-Leu143	CDTL	VIII	-93.2	-37.6	178.6 - 168.8 -58.9	- 128.1	124.9	-56.8	6.7	Yes
11.	Ser151-Pro154	SSSP	VIII	-68.7	-16.9	57.6	-89.7	119.5	162.8	6.9	Yes
12.	Pro176-Phe154	PGAF	VIII	-84.4	-19.9	-	-	143.1	-	6.9	Yes
13.	Val183-Pro186	VLCP	VIII	-59.7	-46.5	169.0	129.7	91.6	175.5	5.9	Yes
14.	Val209-Ser212	VPLS	I	-68.9	-13.5	19.0	-	-11.7	-63.1	5.7	No
15.	Arg213-Thr216	RNST	I	-66.4	-14.5	-62.1	129.5 -75.5 -82.5	-0.9	51.0	6.0	No
16.	Val222a225r223-	VSGA	I	-71.1	-23.4	45.3	-95.8	8.6	-	6.0	No
17.	Lys226	SGAK	IV	-95.8	8.6	-	-42.0	75.8	-	6.9	Yes

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	No. Residd	Length	Unit Rise	Residue Per turn	Pitch	Deviation from Ideal	Sequence
1.*	Leu8	Asn17	H	10	15.63	1.5	3.59	5.43	2.2	LGEKWKARLN
2.*	Ala21	Lys28	H	8	12.52	1	3.79	5.60	13.5	ALEFYSYK
3.*	Glu38	Lys45	H	8	12.47	1.4	3.71	5.60	11.9	EEARRALK
4.*	Gly58	Glu67	H	10	15.72	8 1.5 1 1.5 0	3.69	5.54	8.8	GSAKLRWLVE
5.	Gly86	Ala92	H	7	11.28	1.5	3.65	5.35	7.2	GWSYYAA
6.	Trp12	Ile123	G	3	-	4	-	-	-	WNI
7.	1	His13	G	3	-	-	-	-	-	VFH
8.	Val132 Pro154	4 Leu172	H	19	28.29	- 1.46	3.66	5.35	13.4	PEVEEARTLRVLSMVGDWL
9.	Ser189	Tyr202	H	14	20.03	1.4	3.60	5.23	9.7	STMMETLERLQRRY
10.	Thr229	Gly242	H	14	21.41	5 1.49	3.69	5.49	4.1	TIKSVSTTSQLLLG

Number of helix in chain 10; *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-methyltranferase 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, w**e 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 prelimianary 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 ZIKA 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] Triunfol M., *The Lancet Infectious Diseases* 2016 16:156 [PMID: 26723756].

[3] Mlakar J *et al. New England Journal of Medicine* 2016 **374**: 951-958 [PMID: 26862926].

[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].

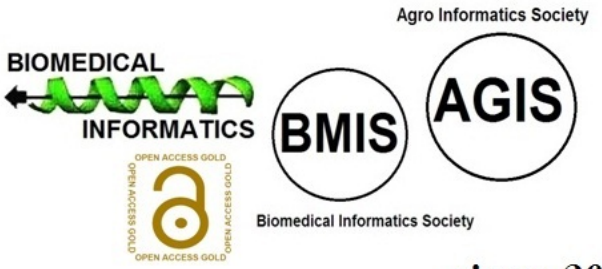
BIOINFORMATION
Discovery at the interface of physical and biological sciences

[9] Zhao Y *et al. Proceedings of the National Academy of Sciences* 2015 **112:**14834 [PMID: 26578813].
[10] Egloff MP *et al. The EMBO Journal*, 2002 **21:**2757 [PMID: 12032088].
[11] Quevillon E *et al. Nucleic Acids Research*, 2005 **33:**W116 [PMID: 15980438].
[12] Tarique M *et al. Frontiers in Microbiology* 2017 **8:**130 [PMID: 28232818].
[13] Castresana J, *Molecular Biology and Evolution* 2000 **17:**540 [PMID: 10742046].
[14] Jones DT, *Journal of Molecular Biology* 1999 **292:**195 [PMID: 10493868].
[15] Lewis PN *et al. Biochimica et Biophysica Acta (BBA) - Protein Structure* 1973 **303:**211 [PMID: 4351002].
[16] Hutchinson EG & Thornton JM, *Protein Sci* 1994 **3:** 2207 [PMID: 7756980].
[17] Venkatachalam CM, *Biopolymers* 1968 **6:**1425 [PMID: 5685102].
[18] May CE & R. Appels R, *Theor Appl Genet* 1987 **74:**617 [PMID: 24240218].
[19] Wilmot CM & Thornton JM, *Protein Eng* 1990 **3:**479 [PMID: 2371257].

Edited by P Kangueane
Citation: **Afaq *et al.*** Bioinformation 16(3): 229-235 (2020)

BIOINFORMATION
Discovery at the interface of physical and biological sciences

OPEN ACCESS GOLD



since 2005

BIOINFORMATION
Discovery at the interface of physical and biological sciences

indexed in