Available online freely at www.isisn.org

Bioscience Research

Print ISSN: 1811-9506 Online ISSN: 2218-3973

Journal by Innovative Scientific Information & Services Network

RESEARCH ARTICLE BIOSCIENCE RESEARCH, 2023 20(3): 667-676. **OPEN ACCESS**

The Sequence provides insight into a structure-based investigation of the critical *Japanese encephalitis* domain

Ziaullah M Sain[1], Asif Hussain Akber[2], Yousef Mutad Al Harbi[2], Khaled Ahmed Alahmari[3], Lahiq Ahmad Aseri[4], Awadh Ali Ayed Alqarni[5], Osama Faqihi[6], Nabeela Al-Abdullah[7]and Qamre Alam[8*]

[1]Department of Microbiology, Faculty of Medicine, Rabigh, King Abdulaziz University, Jeddah. 21589. **Saudi Arabia**
[2]Central Military Laboratory and Blood Bank Department - Virology Division, Prince Sultan Military Medical City, Riyadh 12233, **Saudi Arabia**
[3]Central Military Laboratory and Blood Bank Department - ED Lab Division, Prince Sultan Military Medical City, Riyadh 12233, **Saudi Arabia**
[4]Central Military Laboratory and Blood Bank Department - Hematology Lab Division, Prince Sultan Military Medical City, Riyadh 12233, **Saudi Arabia**
[6]Wadi Aldwasir Military Hospital, Laboratory Department, Blood Bank, 0117840323, Wadi-AlDwasir, **Saudi Arabia**
[6]Respiratory Therapy Department, King Fahad Military Medical Complex Dhahran, **Saudi Arabia**
[7]King Abdulaziz University, Faculty of Nursing, Public Health Department, Infection Control Unit, King Abdulaziz University Hospital Jeddah, **Saudi Arabia**
[8*]Molecular Genomics and Precision Medicine Department, ExpressMed Diagnostic and Research, Block, 359, Zinj, **Kingdom of Bahrain**

***Correspondence:** alamqa2022@gmail.com, qamrealam@expressmedlabs.com **Received: 08-06-2023, Revised: 24-06-2023, Accepted: 30-06-2023 e-Published: 07-07-2023**

Mosquitoes transmit the flavivirus known as the *Japanese encephalitis* virus (JEV), which is related to dengue, yellow fever, and other viruses. In this manuscript, we have done an insilico examination of the JEV (*Japanese encephalitis*) protein (Helicase ATP binding domain). We predicted a crucial part that belongs to the DEAH-Box RNA helicase. The available tools Prosite, Pfam, and Inter ProScan help to identify the helicase ATP binding domain (HABD). We also predict alignment, amino acid composition, charged amino acid, atomic level study, and molecular weight, including theoretical Pi. We also study the modelled structure of the helicase ATP binding domain (helices, beta turns) and their details, including secondary structure. This family takes into account an enormous number of RNA helicases. The Ramachandran plot was used for structure validation, including the topology observation of the helicase domain protein. The domain contains strands, α-helix, and circles. The part has a beta-turn, and their deposits are separate between the under 7å. The numeral of collections per turn for alpha helices, the helix contributes to the extent of the deviancy of the helix geometry as of an ideal helix. We also investigate the protein, DNA, and RNA binding sites. In the current research, we appraise the structure-based information and its composition (amino acid and atomic), helices, beta turns, and binding sites' details.

Keywords: JEV, Pfam, InterProScan, α-helix, Prosite, Secondary structure, Ramachandran plot, binding sites.

INTRODUCTION

The *Flavivirus* genus, which is a member of the *Flaviviridae* family, contains several viruses that are carried by arthropods. The West Nile virus (WNV), dengue virus (DEN), yellow fever virus (YFV), zika virus (ZIKV), and tick-borne encephalitis virus are among these viruses. The zoonotic cycle of JEV transmission between pigs, birds, and mosquitoes is maintained throughout the south & southeast areas of Asia (Solman et al. 2003, Tsai et al. 2000). The JEV virus, which attaches to the nervous system and has a high mortality rate, is spread through mosquito bites. JEV has a plus-sense single-stranded RNA genome. It consists of the structural proteins membrane, envelope, and capsid. Additionally, there are seven non-structural proteins, designated as NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5, which are translated as precursor poly-proteins (Sumiyoshi et al. 1987, Linares et al. 1989, Chambers et al. 1990). In this manuscript, we focused on the NS3 domain (protease-helicase) of the non-structural protein of JEV (Weinert et al. 2015, Crystal et al. 2019).

Established on sequence assessment and conserved features, helicases are divided into three super families. They are motor enzymes that catalyze the unwinding and

remodeling of double-stranded nucleic acids using the energy from NTP hydrolysis (Luking et al. 1998). The sequence of motif II allows for the further classification of helicase into the DEAD, DExD, and DExx subfamilies (Ahmadian et al. 1997). All helicases use the motif I/Walker A motif, which is a phosphate-binding P-loop, as well as the motif II/Walker B motif, which is an aspartic acid loop that binds Mg^{2+} (Koonin et al. 1991).

The helicase, NTPase (nucleoside 5′-triphosphatase), & 5′-terminal RNA triphosphatase functions of the protein (NS3) are all carried out by its catalytic domain. Functionally, the NS3 proteins of JEV, DEN, and YFV have been confirmed to contain helicase and ATPase enzymatic activity (Cui et al. 1998, Kuo et al. 1996, Warrener et al. 1993). Prior to the Walker A according to research, a conserved motif (Q) is required for DEAD-box helicases to operate as ATPases [Tanner et al.2003]. It has been shown that the JEV and HCV proteins' motif II (DExH/D) is essential for the functions of an ATPase & an RNA helicase (Utama et al. 2000). Rep from Escherichia coli (*E. coli*) of the superfamily 1 and the NS3 helicases from DEN, YFV, and HCV and UvrB from Bacillus caldotenax, which are members of the superfamily 2, are among the helicases whose crystal structures have been reported (Subramanya et al. 1996, Korolev et al.1997, Cho et al. 1998, Kim et al. 1998). The RNA binding, RNA helicase, and NTPase enzymes are involved in the viral genome's transcription and replication (Theis et al. 1999). The current study, we assess the depth structure of the domain (NS3) that may be played in the molecular processes.

MATERIALS AND METHODS

This manuscript includes a different methodology section.

Organization of the detail domain

The 3432 amino acid JEV (Japanese encephalitis) protein detailed domain organization. We concentrated on the critical (HABD) helicase ATP binding domain (188 aa), one of the essential domains. The sequence of this domain (the Helicase ATP binding domain) was obtained from the NCBI genome database for this investigation. Pfam, InterProScan & Scan Prosite were some implements we make the most for the manual domain organization (Tarique et al. 2017, Afaq et al. 2020, Bateman et al. 2002).

Amino acid sequence and its composition

The amino acid sequence (188 amino acids) of JEV was submitted to ExPASy Server on the link (https://web.expasy.org/protparam/), generated an aligned amino acid pattern with the Number of amino acids, molecular weight, and theoretical PI (Gasteiger et al. 2005, Boeckmann et al. 2003). The total amino acid composition of the protein (CSV format) is put on the excel sheet, and draw a sound representation of several residues with percentages covered throughout the protein [Apweileret al.2004]. Similarly, we have done deep analysis to identify the total number of atoms (2936). The sequence analysis of helicase ATP binding protein was done using expect protein (http://www.predictprotein.org), online available bioinformatics facility for sequence analysis & protein structure & function prediction. This software provides us with the total composition of the protein, including several aligned proteins and several matched pdb structures (Bernhofer et al. 2021, Burkhard et al. 2004).

Modelling & secondary structure prediction

188 amino acids of the protein were modelled structurally (secondary) via the SWISS-MODEL (https://swissmodel.expasy.org) using an appropriate 2z83.1.A (template) [Tarique et al 2013]. One may also see how the protein functionalities are appreciated by examining their structures. The template was used to create the structure models. For the secondary structure prediction using psipred, the produced PDB file from modelling was used (Armold et al. 2006, Jones et al. 2022, Moffat et al. 2022, Kandathil et al. 2022). The PROCHECK program used the generated file (PDB) to validate the simulated structure using a Ramachandran plot. The concurrently discovered motifs include the beta-turn, game-turn, helix, strand, & beta-hairpin. We used PDB sum (Sheik et al. 2002, Laskowskiet et al. 1997) to understand the produced models' topology.

Helices and beta turns

The well-planned colour illustrations "Helical turn & net" show the residues in each helix. The residues are color-coded according to the types of amino acids: hydrophobic (green), polar (blue), & charged (red). Concerning recovered helical value, the turns and nets accept 3.6 residues per turn (May et al. 1987, Wilmot et al. 1990). The four surrounding residues (i, i+1, i+2, & i+3) constitute a helicase domain. The key two residues must not be helical, and their C-alpha atom distances must be less than 7 (Ramachandran et al. 1968, Hutchinson et al. 1994). Using the polyproline area of the plot's phi, psi, and accumulation i+1 edges, two subclasses of kind VIa turns were identified. A requirement for VIa1, VIa2, & VIb kind turns is that residue I will apply cis-proline. Turns grouped as type IV do not meet the criteria mentioned earlier (Lewiset et al. 1973, Richardson et al. 1981).

Binding site prediction

The NCBI database is used to extract the protein's primary sequence, which is then sent to the server at http://www.predictprotein.org to generate several regions of binding sites, including protein, DNA, and RNA binding sites. The different coloured boxes in this viewer represent expected traits that correlate to specific locations in the primary sequence. Proteins, DNA, and RNA are critical in

determining all fundamental biological processes, including mechanisms (Yachdav et al. 2014).

RESULTS

Organization of domain

The details of domain organizations have been done using the available bioinformatics tools/ server. As we know that JEV (*Japanese encephalitis*) contains (1-3432 amino acid) many domains such as nonstructural protein NS-1 (797-1146 amino acid), NS2A (1158-1339 amino acid), NS2B (1374-1504 amino acid), Helicase ATP binding domain (1672-1859 amino acid), NS4A (2129-2268 amino acid) and NS4B (2273-2518 amino acid) (Figure1A). We carefully focused on the binding domain for the study in this article, Helicase ATP Binding Domain (HABD). We further used InterProScan to filter the helicase ATP binding domain. We found the HABD with 187 stretches of amino acid in length (Figure 1B). The HABD belongs to DEAH-Box RNA helicase. The amino acid sequence of HABD is shown in different colors, PTR and DEAH (Figure 1C). We also identified the Helicase ATP Binding Domain with various tools, InterPro, Prosite, and Pfam (Figure 1D).

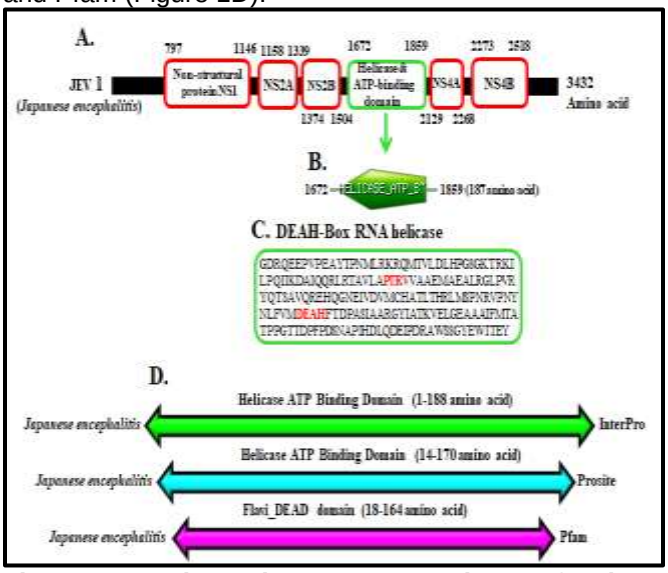

Figure 1: Schematic representation of domain organization JEV domain. (A) The important conserved domains are written inside the boxes and highlighted in the helicase domain. (B) Helicase ATP binding domain (HABD) contains 187 amino acids. (C) Essential amino acid is shown in different colors. (D) Identification of Helicase ATP Binding Domain with various tools.

DNA viruses that replicate mainly in the nucleus frequently practice a cellular helicase. In contrast, most RNA viruses that replicate largely in the cytoplasm possess a self-encoding helicase.

Sequences and its composition

A sequence of *Japanese encephalitis* virus essential helicase ATP binding domain with 188 amino acids. The alignment of the amino acid sequence was done with the help of the Protparam server. We found the molecular weight (21003.84) and theoretical Pi (5.67) (Figure 2A). Figure 2B represents the total amino acid composition and number of residues/ parentages throughout the protein. Amino acid residues (Ala): 11.2%, Arg: 7.4%, Asn: 3.2%, Asp: 5.9%, Cys: 0.5%, Gln: 4.8%, Glu: 6.9%, Gly: 4.8%, His: 3.2%, Ile: 5.9%, Leo: 6.9%, Lys: 2.7%, Met: 3.7%, Phe: 2.1%, Pro: 8.5%, Ser: 3.7%, Thr: 8%, Trp: 1.1%, Tyr: 3.2% & Val: 6.4% to cover throughout the protein. We further calculated this protein's negatively charged residues and positively charged residues. We discovered that negatively charged amino acids (Asp + Glu) have 24 and positively charged amino acids (Arg + Lys) have 19 in number. (Figure 2C). Similarly, we moved toward the atomic composition, and we found that atoms Nitrogen, Sulfur, Hydrogen, Carbon & Oxygen. The graphical representation reveals that C: 31%, H: 50%, N: 9%, O: 10%, and S: 0% throughout the protein (Figure 2D).

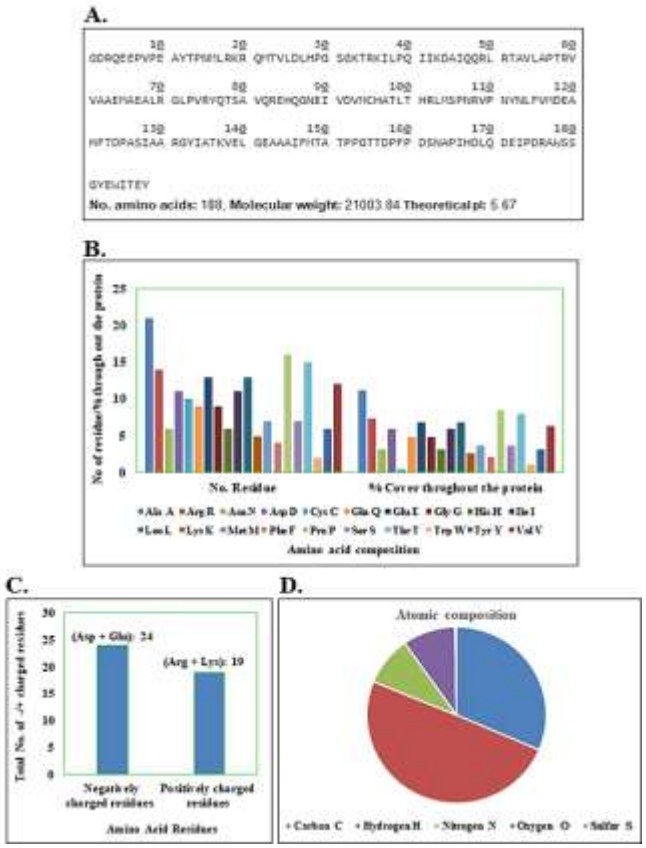

Figure 2: Aligned of JEV amino acid. (A) The amino acid sequences Helicase ATP binding domain were submitted to the protparam server, and the data were obtained. (B) The total amino acid composition and number of residue/ parentages throughout the protein. (C) The graph represents the positively charged and negatively charged amino acid residues. (D) The diagram illustrates the atomic composition of protein.

Modelling and Secondary structure

The Swiss model server successfully provided the anticipated model of helicase ATP binding domain protein and its 3-dimensional coordinate file in PDB format. The secondary structure was constructed, and the results obtained as of the server contain predicted secondary structure (secondary) by a confidence score. JEV-Helicase domain, template (2z83.1.A), and overlay pictures are all shown in Figure 3A. Strand, helix, and coil secondary structure perdition are revealed in a different hand. Figure 3B's strand is represented by the color yellow, the helix by the color purple, and the coil by the color grey. Similarly, we discovered that the protein contained tiny amounts of nonpolar, polar, hydrophobic, and aromatic molecules and cysteine (Figure 3C).

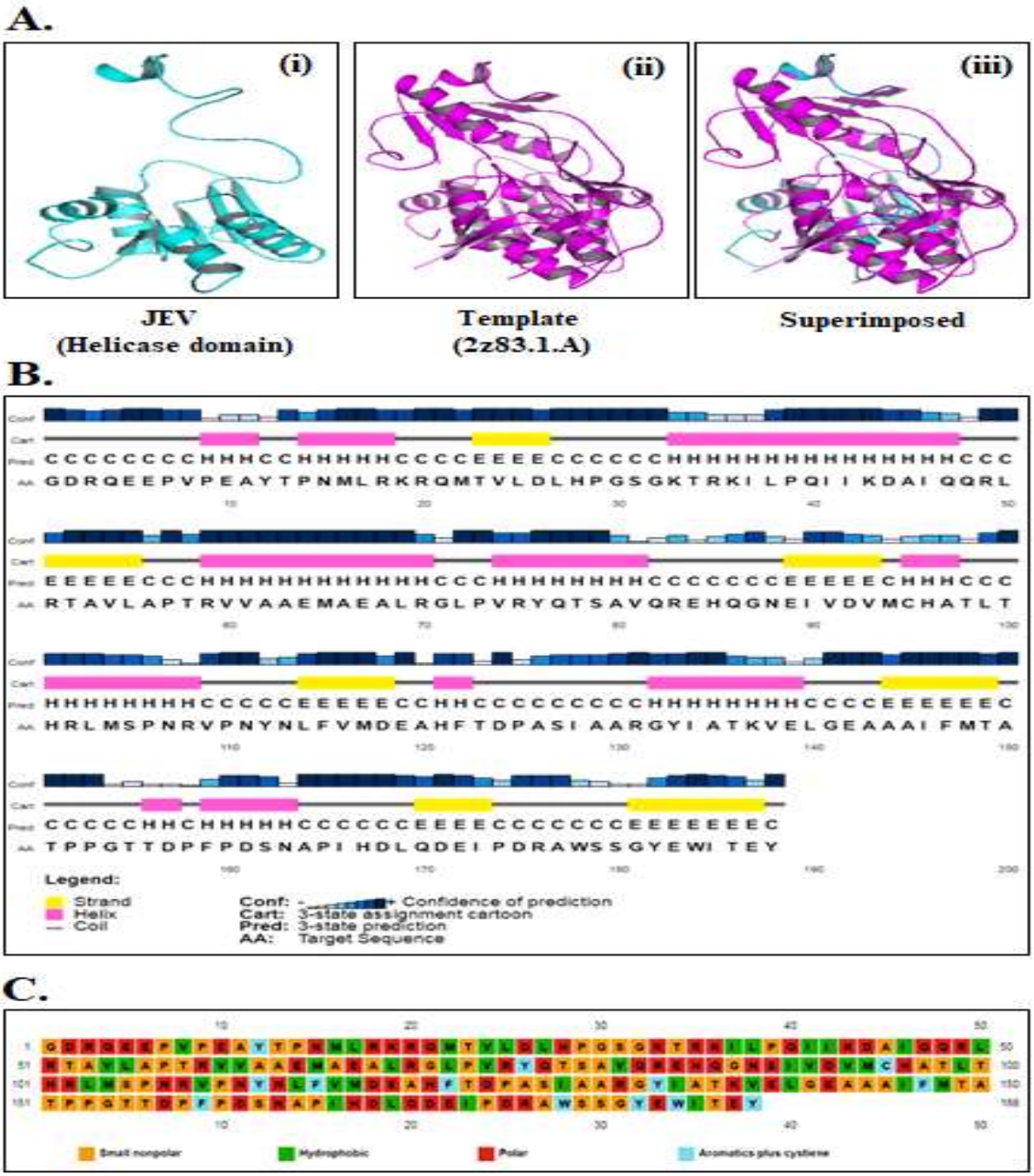

Figure 3: Structure modelling (A) (i) JEV-Helicase domain (ii) Template (2z83.1.A) (iii) Superimposed image. (B) The secondary structure prediction (Strand, helix, and coil) was generated. The yellow amino acid color represents the strand, purple for the helix, and gray for the coil. (C) Similarly, small nonpolar, polar, hydrophobic, and aromatics plus cysteine.

Ramachandran plot and topology

To assess the predicted model helicase protein with an ATP-binding domain. The best-expected model's Ramachandran plots were obtained from PROCHECK servers and showed the model's consistency. A fair eminence model for the helicase domain protein can be inferred from the PROCHECK Ramachandran plot, which showed residues in the most regions (favoured) & few in different regions (allowed) (Figure 4A). Using the PDBsum server, the topology of the protein containing the domain (Helicase ATP binding domain) was observed. The sequence of structure (secondary) elements in the helicase domain protein and their relative 3-D positions and approximate alignments are substantially simplified (Figure 4B).

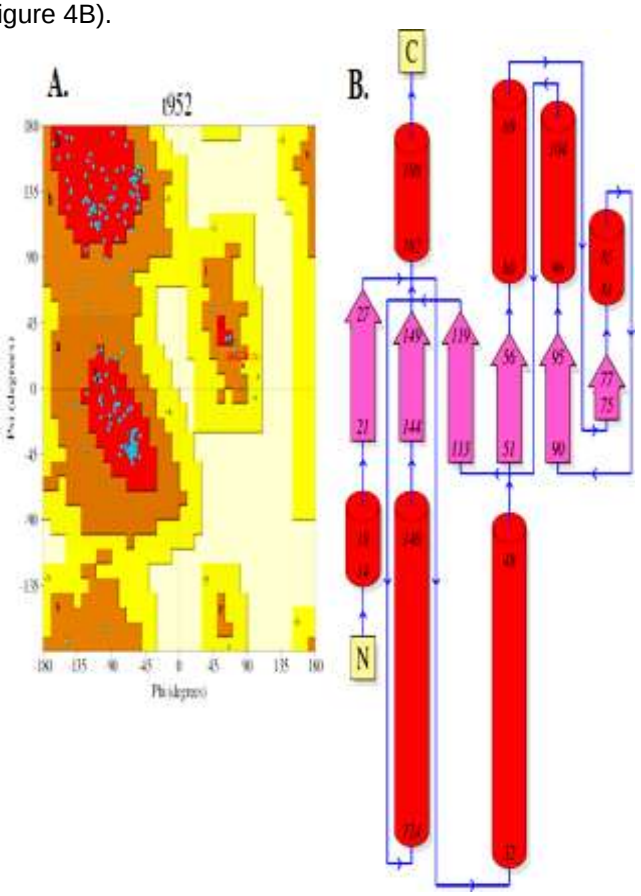

Figure 4: Ramachandran plot (A) Ramachandran plot of protein. (B) The observed topology of the protein (Helicase ATP binding domain) was done using the PDBsum server.

Empathy of helices

The data listed in the table for each helix includes the helix number, which is assigned consecutively starting by way of one on the N-end of the protein, the buildup numbers concerning the beginning & end of the helices, & the helix type, either alpha helix (H) or helix (G). The number of deposits in the helix is followed by details about its geometry, including its length and unit rise (measured in Angstroms), the number of buildups per turn (ideally 3.6) for alpha helices, the helix contribution Angstroms, and the degree to which it deviates from a perfect helix. This last esteem ought to be 0 for an ideal helix. The geometrical parameters are not determined for helices with less than four buildups. The amino acid (aa) sequence of the helix is shown in the table's final segment. Table 1 shows the helices of the DHX8 protein (Helicase ATP binding domain). The helical value of 3.6 residues per turn was approved. (Figure 5).

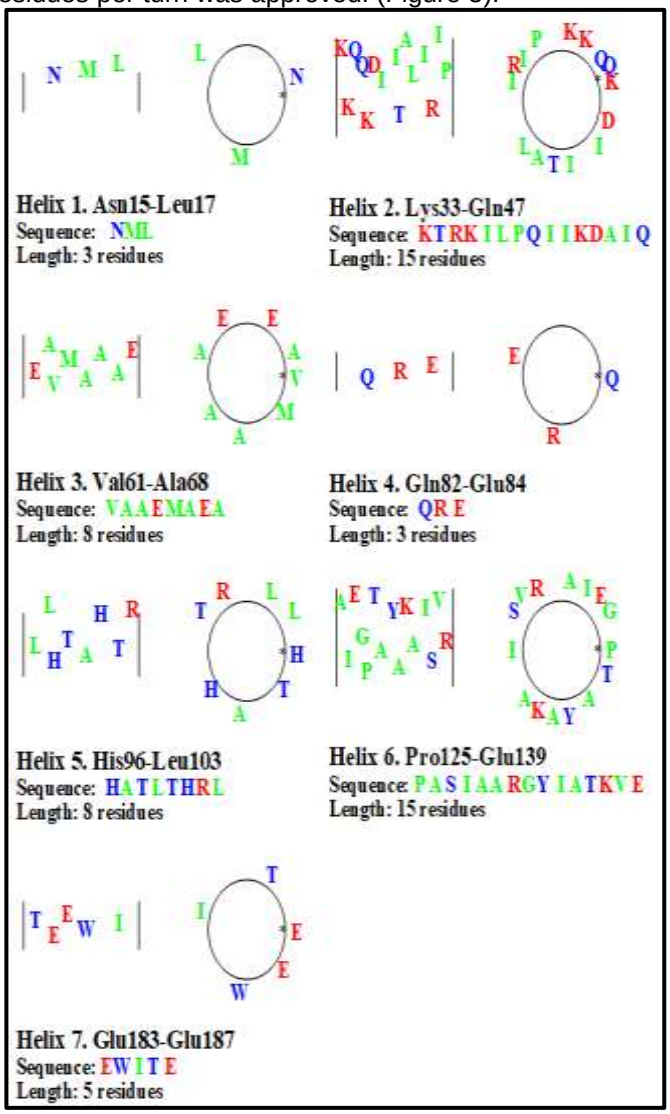

Figure 5: The helical diagrammatic depicts the residues (amino acids) preparation in every helix. Green (hydrophobic), red (charged), and blue (polar) colours indicate residues (amino acids). The assessment of residues per turn by helices was approved.

Table 1: Helices in the helicase domain from *Japanese encephalitis*

S.No.	Start	End	Type	No. residue	Length	Unit rise	Residue s/turn	Pitch	Deviation from ideal	Sequence
1.	Asn15	Leu17	G	3	-	-	-	-	-	NML
2.	Lys33	Gln47	H	15	21.87	1.42	3.84	5.48	14.3	KTRKILPQIIKDAIQ
3.	Val61	Ala68	H	8	12.17	1.46	3.67	5.35	13.2	VAAEMAEA
4.	Gln82	Glu84	G	3	-	-	-	-	-	QRE
*5.	His96	Leu103	H	8	12.93	1.57	3.62	5.68	13.7	HATLTHRL
*6.	Pro125	Glu139	H	15	23.13	1.51	3.65	5.50	6.6	PASIAARGYIATKVE
*7.	Glu183	Glu187	H	5	8.03	1.47	3.76	5.53	24.6	EWITE

Number of helices in chain A: 7

*Asterisked motifs relate to those exemplified in the motif plots at the top of the page.

Empathy of beta turn

Proteins can have beta twists, an unconventional supporting structure that alters how the polypeptide chain behaves over time. In proteins and polypeptides, they recur frequently. There are four dangerous amino buildups in each. The helicase ATP binding domain of a protein's beta-turn has been extensively studied. The turning plots show a Ramachandran plot with the circle with brown (residues i+1) & square with green (i+2) plotted. If the focus two deposits are not helical, as determined by either the Kabsch & Sander criteria or the creator-described criteria, & the space between the Calpha molecules of buildup I & buildup i+3 is less than 7 (Figure 6), the two deposits are not helical. Every beta-turn's information has surfaced:

The turn type

The buildup quantities of deposits I and i+3, in turn

The one-letter stretch amino acid codes for those deposits Phi, psi, and chi1 are specified for each focal buildup (i+1 and i+2). The final column shows how accumulations I and i+3's Calpha iotas differ and if hydrogen haven applies to these deposits. Tables 2 and 3 contain information regarding beta turns in the protein.

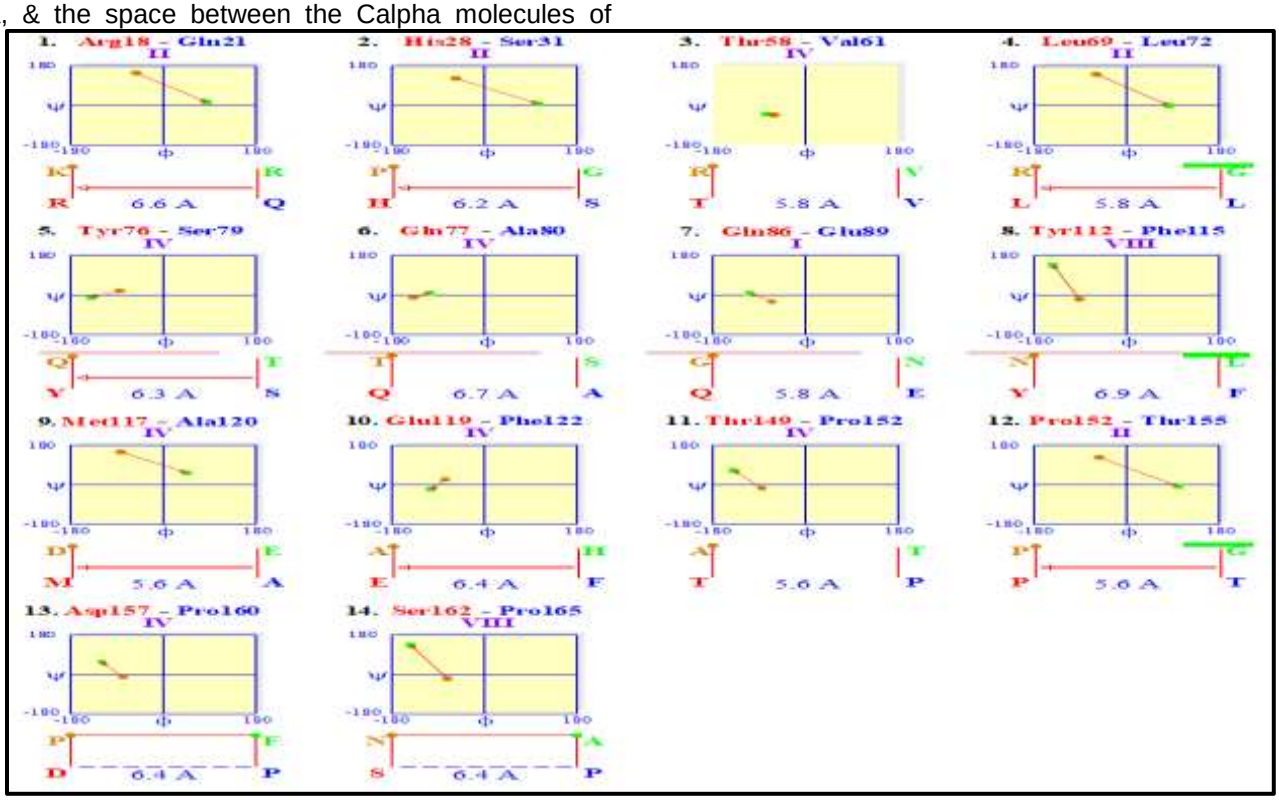

Figure 6: The plots for turns display a Ramachandran plothru, which has a circle with brown (residues i+1) and a square with green (i+2) schemed on it. The representation map of the turn through the four residues (amino acids) with the distance between them labelled as i-i+3. In the residue (amino acid), I donate a hydrogen bond to residue (i+3), as shown by the red arrow. The Ramachandran plot displays the number of residues (amino acids) and the turn type.

Table 2: Strands (Beta) in the helicase domain (ATP bining) from *Japanese encephalitis*.

S.No.	Start	End	Sheet	No. of residue	Edge	Sequence
1	Met22	Asp26	A	5	Yes	MTVLD
2	Thr52	Ala56	A	5	No	TAVLA
3	Arg75	Tyr76	A	2	Yes	RY
4	Val91	Cys95	A	5	No	VDVMC
5	Leu114	Asp118	A	5	No	LFVMD
6	Ala144	Thr149	A	6	No	AAIFMT

Number of strands in chain A: *6*

Table 3: Turns (Beta) in the helicase domain from *Japanese encephalitis*.

S.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	Arg18-Gln21	RKRQ	II	-52.9	147.2	-72.4	84.5	15.6	-145.9	6.6	No
2	His28-Ser31	HPGS	II	-55.9	122.8	-23.7	104.3	7.7	-	6.2	No
3	Thr58-Val61	TRVV	IV	-59.7	-43.7	-64.1	-76.4	-40.0	88.3	5.8	Yes
4	Leu69-Leu72	LRGL	II	-59.2	141.6	-75.7	82.4	-0.9	-	5.8	No
5	Tyr76-Ser79	YQTS	IV	-82.9	20.3	-69.0	-139.2	-11.1	37.7	6.3	No
6	Gln77-Ala80	QTSA	IV	-139.2	-11.1	37.7	-106.8	9.2	-58.3	6.7	Yes
7	Gln86-Glu89	QGNE	I	-68.2	-30.1	-	-111.8	6.7	62.7	5.8	Yes
8	Tyr112-Phe115	YNLF	VIII	-91.9	-18.3	-54.0	-142.9	136.3	177.3	6.9	Yes
9	Met117-Ala120	MDEA	IV	-82.8	149.5	-157.9	45.9	54.0	-61.7	5.6	No
10	Glu119-Phe122	EAHF	IV	-76.1	24.8	-	-103.7	-20.0	56.3	6.4	No
***11**	Thr149-Pro152	TATP	IV	-84.9	-15.4	-	-139.6	64.8	45.2	5.6	Yes
***12**	Pro152-Thr155	PPGT	II	-53.7	124.7	-29.3	100.8	-6.5	-	5.6	No
***13**	Asp157-Pro160	DPFP	IV	-79.9	-10.8	4.6	-121.6	56.0	-58.7	6.4	Yes
***14**	Ser162-Pro165	SNAP	VIII	-72.2	-18.5	-63.2	-143.3	134.4	-	6.4	Yes

Number of beta turns in chain A: 14

* Asterisked motifs relate to those clarified in the motif plots by the top of the page.

* Residue colouring corresponds to the colouring given in the motif plots and in the RasMol displays.

Binding site extrapolation

To generate various regions of binding sites, such as protein, DNA, and RNA binding sites, the main sequence (protein) is taken from the NCBI database and sent to the server (http://www.predictprotein.org). The protein binding site schematic is shown in Figure 7A: There are eight distinct places with varying quantities of amino acids where the binding sites are present. This crucial location is also appropriate for interacting proteins or necessary ligands. The DNA binding site can be found in Figure 7B in nine different areas. The mechanism at the DNA level may benefit from this binding—the RNA binding site in Figure 7C. The RNA helicase subclass of proteins includes this one. In terms of biological support, it will be helpful too. With the various coloured boxes, this viewer displays expected features matching locations in the primary sequence. Proteins, DNA, and RNA crucially determine all essential biological processes, including mechanisms.

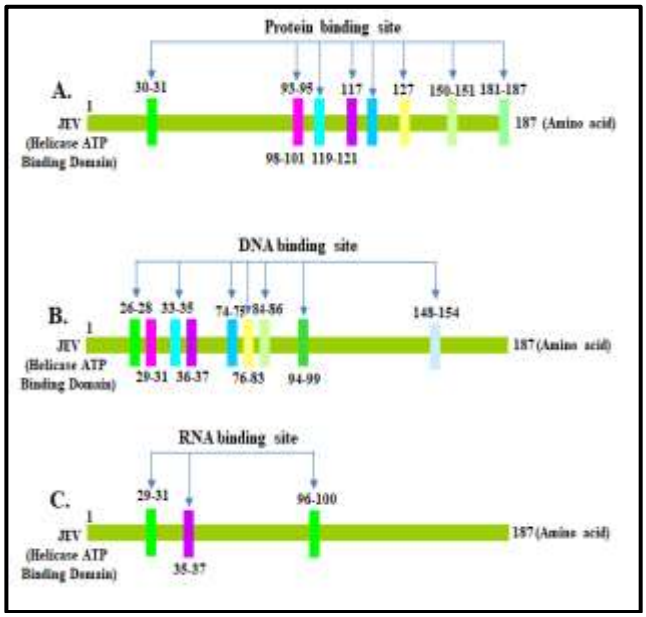

Figure 7: JEV (Helicase ATP binding domain) binding sites. (A) Important protein binding sites (B) Binding sites (DNA) (C) Binding sites (RNA).

CONCLUSION

We have done a thorough bioinformatics examination of the JEV **(***Japanese encephalitis***)** protein (Helicase ATP binding domain) in this work. In light of the consequences of this investigation, we have built up a novel learning-based computational methodology for identifying Helicase ATP binding domain and ideal evaluate an approach to predicted helicase ATP binding domain with different available tools; we also indicate alignment, amino acid composition; charged amino acid, atomic level study and its molecular weight including theoretical Pi. We also study the modelled structure of the helicase ATP binding domain (helices, beta turns) and their details. The conserved motif DEAH is the group's utmost distinctive feature. This family takes into account an enormous number of RNA helicases. The domain contains strands, α-helix, and circles. The number of assortments per turn for alpha helices, the helix contributes to the extent of the abnormality of the helix geometry from an ideal helix. The present study evaluates an approach to predicted hydrophobicity, amino acid charge, antigenic region, structure (helices, beta turns), & their details. The plots for each turn demonstrate a Ramachandran plot with deposits plotted on it. The buildup numbers & turn type are shown over the Ramachandran plot. The geometrical parameters are not determined for helices. This study will help us understand the helix's amino acid succession and the details of the structure.

CONFLICT OF INTEREST

The authors declared that the present study was performed without any conflict of interest.

ACKNOWLEDGEMENT

The authors express heartfelt gratitude towards, Express Med Diagnostic and Research, Block, 359, Zinj, Kingdom of Bahrain, for providing all the obligatory facilities in the course of this work.

AUTHOR CONTRIBUTIONS

ZMS and QA: designed and first drafted the manuscript. AHA YMA, KAA: data analysis, and revised the manuscript. LAA, AWA, and OF: data collection, and editing of the manuscript. All authors read and approved the final version. MR designed experiments and reviewed the manuscript. All authors read and approved the final version

REFERENCES

Afaq S, Atiya A, Malik A, Alwabli AS, Alzahrani DA, Al-Solami HM, Alzahrani O, Alam Q, Kamal MA, Abulfaraj AA et al: Analysis of methyltransferase (MTase) domain from Zika virus (ZIKV). Bioinformation 2020, 16(3):229-235.

Ahmadian M.R., Stege P., Scheffzek K., Wittinghofer A. Confirmation of the arginine-finger hypothesis for the GAP-stimulated GTP-hydrolysis reaction of Ras. Nat. Struct. Biol., 4 (1997), pp. 686-689

Apweiler, R., Bairoch, A., Wu, C. H., et al. (2004) UniProt: the Universal Protein knowledgebase. Nucleic Acids Res. 432, D115–D119.

ARMOLD K: The SWISSMODEL workspace: a web-based environment for Protein structure homology modilling. Bioinformatics 2006, 22:195-201.

Bateman A, Birney E, Cerruti L, Durbin R, Etwiller L, Eddy SR, Griffiths-Jones S, Howe KL, Marshall M, Sonnhammer EL: The Pfam protein families database. Nucleic acids research 2002, 30(1):276-280.

Bernhofer M, Dallago C, Karl T, Satagopam V, Heinzinger M, Littmann M, Olenyi T, Qiu J, Schütze K, Yachdav G, Ashkenazy H, Ben-Tal N, Bromberg Y, Goldberg T, Kajan L, O'Donoghue S, Sander C, Schafferhans A, Schlessinger A, Vriend G, Mirdita M, Gawron P, Gu W, Jarosz Y, Trefois C, Steinegger M, Schneider R, Rost B. Predict Protein-Predicting Protein Structure and Function for 29 Yea rs. Nucleic Acids Res. 2021 Jul 2;49(W1):W535-W540. doi: 10.1093/nar/gkab354.PMID: 3399920.

Boeckmann, B., Bairoch, A., Apweiler, R., et al. (2003) The Swiss-Prot protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res. 31, 354–370.

Burkhard Rost, Guy Yachdav, Jinfeng Liu. The PredictProtein server. *Nucleic Acids Research*, Volume 32, Issue suppl_2, 1 July 2004, Pages W321–W326, https://doi.org/10.1093/nar/gkh377.

Chambers, T. J., Weir, R. C., Grakoui, A., McCourt, D. W., Bazan, J. F., Fletterick, R. J., and Rice, C. M. (1990). Evidence that N-terminal domain of nonstructural protein NS3 from yellow fever virus is a serine proteases responsible for site-specific cleavages in the viral polyprotein. Proc. Natl. Acad. Sci. USA 87, 8898–8902.

Cho H.S., Ha N.C., Kang L.W., Chung K.M., Back S.H., Ja ng S.K., Oh B.H. Crystal structure of RNA helicase from genotype 1b hepatitis C virus. A feasible mechanism of unwinding duplex RNA. J. Biol. Chem., 273 (1998), pp. 15045-15052

Crystal Structure of Catalytic Domain of Japanese Encephalitis Virus NS3 Helicase/Nucleoside Triphosphatase at a Resolution 1.8 Å. Available

online: http://www.rcsb.org/structure/2Z83 (accessed on 23 June 2019).

Cui T., Sugrue R.J., Xu Q., Lee A.K., Chan Y.C., Fu J. Recombinant dengue virus type 1 NS3 protein exhibits specific viral RNA binding and NTPase activity regulated by the NS5 protein. Virology, 246 (1998), pp. 409-417

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.; Protein. Identification and Analysis Tools on the ExPASy Server; (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005). pp. 571-607.

Hutchinson EG, Thornton JM: A revised set of potentials for β-turn formation in proteins. Protein Science 1994, 3(12):2207-2216.

Jones DT & Thornton JM (2022) The impact of AlphaFold2 one year on. Nat Methods, 19(1):15-20. doi: 10.1038/s41592-021-01365-3.

Kandathil SM, Greener JG, Lau AM, Jones DT (2022) Ultrafast end-to-end protein structure prediction enables high-throughput exploration of uncharacterised proteins. PNAS, 119(4):e2113348119. doi: 10.1073/pnas.2113348119.

Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Tho mson J.A., Murcko M.A., and Lin C., Caron P.R. Hepatitis C virus NS3 helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. Structure, 6 (1998), pp. 89-100.

Koonin E.V. Similarities in RNA helicases. Nature, 352 (1991), p. 290

Korolev S., Hsieh J., Gauss G.H., Lohman T.M., Waksma n G. Major domain swiveling revealed by the crystal structures of complexes of E. coli Rep helicase bound to single-stranded DNA and ADP. Cell, 90 (1997), pp. 635-647

Kuo M.D., Chin C., Hsu S.L., Shiao J.Y., Wang T.M., Lin J .H. Characterization of the NTPase activity of Japanese encephalitis virus NS3 protein. J. Gen. Virol., 77 (1996), pp. 2077-2084

Laskowski RA, Hutchinson EG, Michie AD, Wallace AC, Jones ML, Thornton JM: PDBsum: a Web-based database of summaries and analyses of all PDB structures. Trends in biochemical sciences 1997, 22(12):488-490.

Lewis PN, Momany FA, Scheraga HA: Chain reversals in proteins. Biochimica et Biophysica Acta (BBA)-Protein Structure 1973, 303(2):211-229.

Linares, R. A., Cahour, A., Despres, P., Girard, M., and Bouloy, J. (1989). Processing of yellow fever virus polyprotein: Role of cellular proteases in the maturation of the structural proteins. J. Virol. 63, 4199–4209

Luking A., Stahl U., Schmidt U. The protein family of RNA helicases Crit. Rev. Biochem. Mol. Biol., 33 (1998), pp. 259-296

May C, Appels R: Variability and genetics of spacer DNA sequences between the ribosomal-RNA genes of hexaploid wheat (Triticum aestivum). Theoretical and Applied Genetics 1987, 74(5):617-624.

Moffat L, Kandathil SM, Jones DT (2022) Design in the dark: Learning deep generative models for de novo protein design. bioRxiv, doi: https://doi.org/10.1101/2022.01.27.478087.

Ramachandran G, Venkatachalam C: Stereochemical criteria for polypeptides and proteins. IV. Standard dimensions for the cis-peptide unit and conformation of cis-polypeptides. Biopolymers: Original Research on Biomolecules 1968, 6(9):1255-1262.

Richardson JS: The anatomy and taxonomy of protein structure. In: Advances in protein chemistry. vol. 34: Elsevier; 1981: 167-339.

Sheik S, Sundararajan P, Hussain A, Sekar K: Ramachandran plot on the web. Bioinformatics 2002, 18(11):1548-1549.

Solomon T., Ni H., Beasley D.W., Ekkelenkamp M., Cardo sa M.J., Barrett A.D.Origin and evolution of Japanese encephalitis virus in southeast AsiaJ. Virol., 77 (2003), pp. 3091-3098.

Subramanya H.S., Bird L.E., Brannigan J.A., Wigley D.B. Crystal structure of a DExx box DNA helicase. Nature, 384 (1996), pp. 379-383

Sumiyoshi H., Mori C., Fuke I., Morita K., Kuhara S., Kond ou J., Kikuchi Y., Nagamatu H., Igarashi A. Complete nucleotide sequence of the Japanese encephalitis virus genome RNA. Virology, 161 (1987), pp. 497-510

Tanner N.K., Cordin O., Banroques J., Doere M., Linder P. The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. Mol. Cell, 11 (2003), pp. 127-138

Tarique M, Ahmad M, Ansari A, Tuteja R: Plasmodium falciparum DOZI, an RNA helicase interacts with eIF4E. Gene 2013, 522(1):46-59.

Tarique M, Ahmad M, Chauhan M, Tuteja R: Genome wide in silico analysis of the mismatch repair components of Plasmodium falciparum and their comparison with human host. Frontiers in microbiology 2017, 8:130.

Theis K., Chen P.J., Skorvaga M., Van Houten B., Kisker C. Crystal structure of UvrB, a DNA helicase adapted for nucleotide excision repair. Embo J., 18 (1999), pp. 6899-6907

Tsai T.F. New initiatives for the control of Japanese encephalitis by vaccination: minutes of a WHO/CVI meeting, Bangkok, Thailand, 13–15 October 1998. Vaccine, 18 (Suppl 2) (2000), pp. 1-25

Utama A., Shimizu H., Hasebe F., Morita K., Igarashi A., S hoji I., Matsuura Y., Hatsu M., Takamizawa K., Hagiw ara A., Miyamura T. Role of the DExH motif of the Japanese encephalitis virus and hepatitis C virus NS3 proteins in the ATPase and RNA helicase activities. Virology, 273 (2000), pp. 316-324

Warrener P., Tamura J.K., Collett M.S. RNA-stimulated NTPase activity associated with yellow fever virus NS3 protein expressed in bacteria. J. Virol., 67 (1993), pp. 989-996

Weinert, T.; Olieric, V.; Waltersperger, S.; Panepucci, E.; Chen, L.; Zhang, H.; Zhou, D.; Rose, J.; Ebihara, A.; Kuramitsu, S.; et al. Fast native-SAD phasing for routine macromolecular structure determination. Nat. Methods 2015, 12, 131–133.

Wilmot C, Thornton J: β-Turns and their distortions: a proposed new nomenclature. Protein Engineering, Design and Selection 1990, 3(6):479-493.

Yachdav G, Kloppmann E, Kajan L, Hecht M, Goldberg T, Hamp T, Hönigschmid P, Schafferhans A, Roos M, Bernhofer M, Richter L, Ashkenazy H, Punta M, Schlessinger A, Bromberg Y, Schneider. Nucleic Acids Res 2014, 42, 337-343.