Original Research Article

Homology Modeling of Human DNA Repair Protein RAD51 Homolog 3 (RAD51C) in Breast Cancer

Ruzianisra Mohamed[1,2*], Siti Syairah Mohd Mutalip[2], Suhaidah Mohd Joffry[2], Nor Asfarisya Diana Zamri[3] and Nur Batrisyia Asmizan[4]

[1]Bioinformatics Unit, Faculty of Pharmacy, Universiti Teknologi MARA, Selangor Branch, Puncak Alam Campus, 42300 Bandar Puncak Alam, Selangor, Malaysia.
[2]Department of Pharmacology & Life Sciences, Faculty of Pharmacy, UiTM Puncak Alam Campus, 42300 Bandar Puncak Alam, Selangor, Malaysia.
[3]Fakulti Sains, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia.
[4]Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, Malaysia.

ABSTRACT

Breast cancer is known as one of the most predominant cancers that affect both females and males worldwide. The most crucial risk factor in breast cancer is the mutations in the RAD51C gene that have been considered in most hereditary breast cancers. RAD51C, the RAD51 paralogs, is also a deoxyribonucleic acid (DNA) repair protein related to breast and ovarian cancers. DNA double-strand breaks (DSBs) account for the significant detrimental form of DNA damage. RAD51C mutants also have been recognized in breast/ovarian cancer patients. However, the role of the RAD51C protein in hereditary breast cancer and its three-dimensional (3D) structures remains unclear. Thus, this study was conducted to identify the 3D structure of RAD51C protein from its amino acid sequences. The homology modeling for the 3D structure of the RAD51C protein was carried out by using three automated webservers: I-TASSER, SWISS-MODEL, and Phyre2. PyMOL was applied to visualize the 3D structure of RAD51C protein. Next, the MolProbity, ProSA, and SAVES v6.0 programs have been employed to check the stereo-chemical quality of RAD51C protein. The RAD51C-IT models were found to be the best models for the RAD1C protein after being evaluated and validated, and the models were constructed using full-length RAD51C protein sequences. Thus, these protein models can be utilized as a virtual screening tool in discovering potential inhibitors of RAD51C protein.

Keywords: RAD51C protein, homology modeling, breast cancer

Corresponding author:
Ruzianisra Mohamed
Bioinformatics Unit, Faculty of Pharmacy,
UiTM Cawangan Selangor, 42300 Bandar Puncak Alam,
Selangor, Malaysia.
Email: ruzianisra@uitm.edu.my

Received: 22 July 2024; accepted: 24 Oct 2024
Available online: 15 Nov 2024
http://doi.org/10.24191/IJPNaCS.v7i2.11

Copyright© 2024 UiTM Press. This is an open access article licensed under CC BY-SA
https://creativecommons.org/licenses/by-sa/4.0/

1.0 Introduction

DNA double-strand breaks (DSB_S) are the most common DNA lesions that occur in the cell areas and are generally associated with the homologous recombination (HR) pathway as the DNA damage response (1). The HR pathway is a crucial deoxyribonucleic acid (DNA) damage-repairing mechanism, responsible for maintaining genome stability and has been identified to be dysregulated in breast and ovarian cancer (2,3). The DNA damage response through the HR pathway is a cascading process that involves various protein complexes. Tumor suppressor proteins such as BRCA1, BARD1, BAP1, and RAD51 have been reported to be associated with the HR repairing process (2–5).

The introduction of ionizing radiation (IR) and DNA-damaging agents triggered the formation of DSB_S on the DNA strands and simultaneously activated the DNA damage response through the HR pathway. Early in the process, the BAP1 protein was localized to the chromatin near the DSB_S and recruited the BRCA1-BARD1 protein complex (2). This protein complex then localized to the damaged DNA site and aided the resection of the broken DNA strand to generate a single-strand template (6). The activated single-strand template was coated with replication protein A (RPA). RPA facilitates the growth of the RAD51 filament to envelop the ssDNA (7,8) by recruiting the BRCA2-PALB2 tumor suppressor protein complex, eventually substituted with the RAD51 paralogs protein (9). The RAD51 proteins then initiate the DNA repair through the HR pathway (6).

The association of RAD51 paralogs such as RAD51B, RAD51C, RAD51D, XRCC2, and XRCC3 has been identified to form two distinctive protein complexes which are BCDX2 and CX3. These protein complexes are responsible for the HR repairment mechanistic pathway (10,11). The RAD51 paralog C emerged as the most important protein in HR DNA damage repair as it is required to be localized to the damaged DNA site. This protein initiates the HR pathway and is an input signal for the arrest of the cell cycle progression (10,12).

The importance of the RAD51C in the DNA damage response was further justified through various mutations identified in its gene that lead to the increased risk in the prevalence of breast and ovarian cancers (10,13). The knockdown of RAD51C gene expression in breast and ovarian cancer cells increased their sensitivity to several drugs such as camptothecin, cisplatin, and olaparib which are responsible for inducing the formation of DSB_S (10). A study by Wu *et al.* (2022) (14) found a strong association between RAD51 and aggressive cancer biology, cancer cell proliferation, and poor survival in breast cancer. According to Yang *et al.* (2020) (15), RAD51C finds its place on commonly utilized cancer panels due to documented links between harmful genetic variations within these genes and the occurrence of tubo-ovarian carcinoma (TOC). Without RAD51C, the HR pathway cannot be initiated to repair the DSB_s, leading to cell death.

This result makes it obvious that silencing the expression of the RAD51C gene provides a potential therapeutic strategy in combating breast and ovarian cancers. High-throughput virtual screening against chemical databases paves the way for the discovery of novel inhibitors of the RAD51C protein. Unfortunately, the crystal structure of the RAD51C protein is unavailable in the Protein Database Bank (PDB). The absence of the protein RAD51C crystal structure hinders our efforts to visualize the 3D structure of the protein and identify its active sites.

Recently, homology modeling is a method that accurately predicts the 3D structure of a protein based on its amino acid sequences (16,17). The homology modeling process was carried out with the Modeler v9.24 program developed by Sali and Blundell in 1993 (18). This method

greatly contributed to narrowing down the large gap between the deposited proteins' amino acid sequences and the experimentally determined crystal structures in the Protein Data Bank database (19,20). Homology modeling tools such as I-TASSER, SWISS-MODEL, and Phyre2 are widely used in computational biology due to their unique approaches, strengths, and reliability in predicting protein structures when experimental methods (like X-ray crystallography or nuclear magnetic resonance, NMR) are unavailable. These specific tools are often chosen over others, along with their comparative strengths. The application of homology modeling in modern drug discoveries has been utilized to provide initial insights into the physical structure of a protein before the discovery of its crystal structure (21,22).

Homology modeling was developed based on several theories which are 1: evolutionary proteins share a similar protein structure; 2: the conformation of a protein structure is highly conserved compared to its amino acid sequence, in which, even a small change in the sequence will lead to a different protein structure; and 3: the functional site or domain of a protein with the same function is highly conserved in term of its structural folding (23). Thus, homology models play a critical role in drug discovery, as these models enable researchers to understand the 3D structure of the target protein, which is essential for various drug discovery applications, such as virtual screening, structure-based drug design, and the development of novel inhibitors.

Recently, we have been struck by the outbreak of COVID-19 caused by the novel coronavirus (SARS-CoV-2), which led us into a global recession and caused panic in the public healthcare system (24). Researchers have taken various initiatives to find potentially effective drugs against the deadly virus in response to this outbreak. Computer-aided drug discovery (CADD) methods emerged as the frontline in discovering potential drugs due to their fast and efficient processes. The adoption of computer-aided drug design (CADD) offers a promising alternative to rational drug design, serving to cut down both time and financial expenditures (25).

Studies by Iheagwam and Rotimi (2020) (26) and Onawole *et al*. (2021) (21) are examples of scientific discoveries that take advantage of CADD in generating the homology model of SARS-CoV-2 spike protein and identifying potential inhibitors against the virus. Using the same motivation, in this study, we employed homology modeling approaches to predict the 3D protein structure of RAD51C protein based on its amino acid sequence. Three protein structure prediction programs such as SWISS-Model, Phyre2, and I-TASSER were used to build the homology models of RAD51C. The qualities of the models were validated, and the best model was selected as the potential RAD51C protein model.

2.0 Materials and Methods

2.1 Selection of Protein Sequence

The amino acid sequence of the human RAD51C protein (Accession ID: O43502) was retrieved from UniProt (https://www.uniprot.org/) and saved in FASTA format. UniProt is a comprehensive database that includes the well-vetted protein collection in UniProtKB/Swiss-Prot, with each protein entry providing a detailed summary encompassing both experimentally confirmed and computationally predicted information (27). The phylogenetic analysis and the exploration of possible templates for the RAD51C protein was conducted through the Basic Local Alignment Search Tool Protein (BLASTP) (https://blast.ncbi.nlm.nih.gov/Blast.cgi) based on the Protein Data Bank (PDB) database.

2.2 Prediction of Protein Sequence and Homology Modeling

The possible templates for the RAD51C protein were ranked based on the coverage and percentage identity against the target sequence. Higher coverage and percentage identity were considered the best quality for a homolog template. The homology models of RAD51C protein were generated by using three different protein structure prediction webservers which are SWISS-MODEL (https://swissmodel.expasy.org/) (19), Phyre2 (http://www.sbg.bio.ic.ac.uk/Phyre2/html/page.cgi?id=index) (28), and I-TASSER (https://zhanggroup.org/I-TASSER/) (29) using default settings. Possible template sequences were explored using HHblits for SWISS-MODEL and Phyre2 while LOMETS was used in I-TASSER modeling programs.

The SWISS-MODEL is a user-friendly web-based modeling environment incorporating databases and the software needed for homology modeling. The workspace helps the user create and assess protein homology models at varying degrees of complexity. For modeling scenarios where substantially, identical structural templates are available, a highly automated modeling approach with minimal user participation is offered. Tools for choosing a template, creating models, and assessing the quality of structures can be accessed directly from the workspace or through the menu on the website (19).

Given a certain degree of grouping homology, "fitting" the arrangement (target) of a protein into a known structure (layout) may result in the production of an existing cryptic protein structure using Phyre2, an underutilized GUI for homology modeling. Its user-friendly software includes built-in profile watchers, Ramachandran plots, JSP models, and basic model improvement.

An online tool called the I-TASSER server uses structural data to automatically predict protein structure and annotate its function. The goal is to use advanced algorithms that are constantly improving to produce accurate predictions of protein structure and function. This tool generates functional annotations and structural predictions using a hierarchical process that begins with the amino acid sequence. I-TASSER also has included advanced methods for biological function prediction, precise neighbor structure estimation, and atomic-level structure refinement.

2.3 Assessment of Protein Models

The structure of the protein models was visualized using the PyMOL (https://pymol.org/2/) (30), validated by using MolProbity (http://molprobity.biochem.duke.edu) and Protein Structure Analysis (ProSA, https://prosa.services.came.sbg.ac.at/prosa.php)) servers to confirm the protein's tertiary structure (31,32). The SAVES v6.0 server, (https://saves.mbi.ucla.edu/) which includes the ERRAT, overall G-score, and Ramachandran Plot Analysis, was also used to evaluate the quality of each protein model (33,34).

3.0 Results and Discussion

3.1 The Amino Acid Sequences of RAD51C Protein

Homo sapiens RAD51C protein sequences of 376 amino acids were selected for the protein model generation. Closely related templates of RAD51C were analyzed by using the BLASTP program based on this protein sequence through the Protein Data Bank (PDB) database. A total of 30 amino acid sequences which are homologs to the RAD51C protein sequence were identified by the program. These template sequences were ranked based on total coverage, E-value, and percentage identity against the target sequence. A consensus score was given to each template sequence to identify the best template for the RAD51C protein. The amino acid sequences of RAD51C protein are shown in Figure 1.

```
          1.....................................................60
O43502      MRGKTFRFEMQRDLVSFPLSPAVRVKLVSAGFQTAEELLEVKPSELSKEVGISKAEALET
         61....................................................120
O43502      LQIIRRECLTNKPRYAGTSESHKKCTALELLEQEHTQGFIITFCSALDDILGGGVPLMKT
        121....................................................180
O43502      TEICGAPGVGKTQLCMQLAVDVQIPECFGGVAGEAVFIDTEGSFMVDRVVDLATACIQHL
        181....................................................240
O43502      QLIAEKHKGEEHRKALEDFTLDNILSHIYYFRCRDYTELLAQVYLLPDFLSEHSKVRLVI
        241....................................................300
O43502      VDGIAFPFRHDLDDLSLRTRLLNGLAQQMISLANNHRLAVILTNQMTTKIDRNQALLVPA
        301....................................................360
043502      LGESWGHAATIRLIFHWDRKQRLATLYKSPSQKECTVLFQIKPQGFRDTVVTSACSLQTE
        361.............376
043502      GSLSTRKRSRDPEEEL*
```

Figure 1: The amino acid sequences of RAD51C protein.

3.2 Structure Prediction of SWISS-MODEL, Phyre2, and I-TASSER

The RAD51C 3D protein models were generated using SWISS-MODEL, Phyre2, and I-TASSER protein modeling programs. The RAD51C protein models were generated based on the alignment of the best template sequence against the RAD51C protein sequence in each protein modeling program. The 3D structure of the protein models from the SWISS-MODEL (RAD51C-SM), Phyre2 (RAD51C-P2), and I-TASSER (RAD51C-IT) are shown in Figures 2, 3, and 4, respectively. The superposition of the generated protein models is shown in Figure 5. The generated protein model comprises 333 amino acids for RAD51C-SM, 300 amino acids for RAD51C-P2, and 376 amino acids for RAD51C-IT. Based on these results, the RAD51C-IT model is the only homology model generated based on the full-length RAD51C protein sequences.

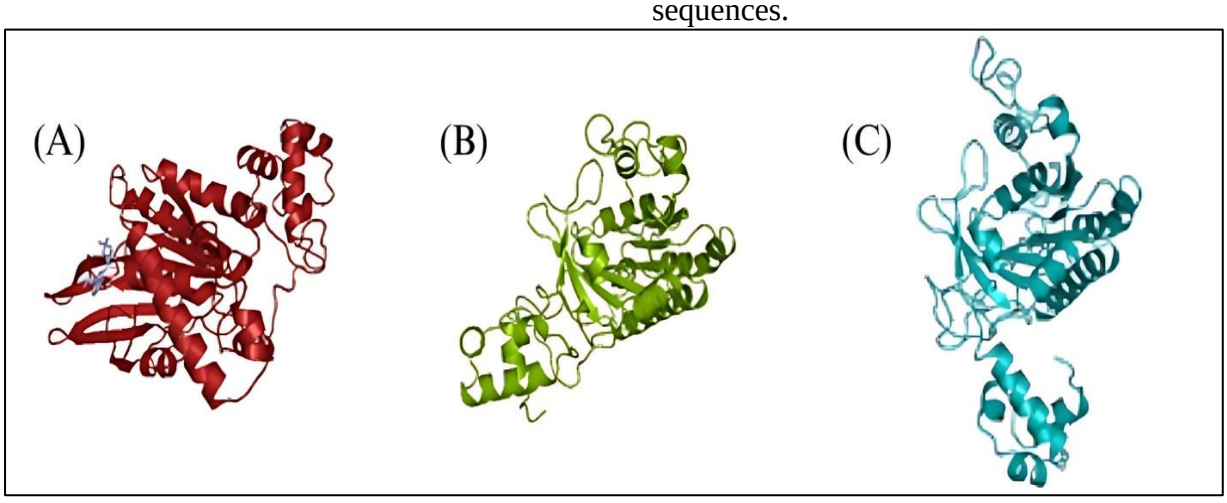

Figure 2: The 3D structure of the RAD51C-SM homology models (A-C) generated by the SWISS-MODEL webserver.

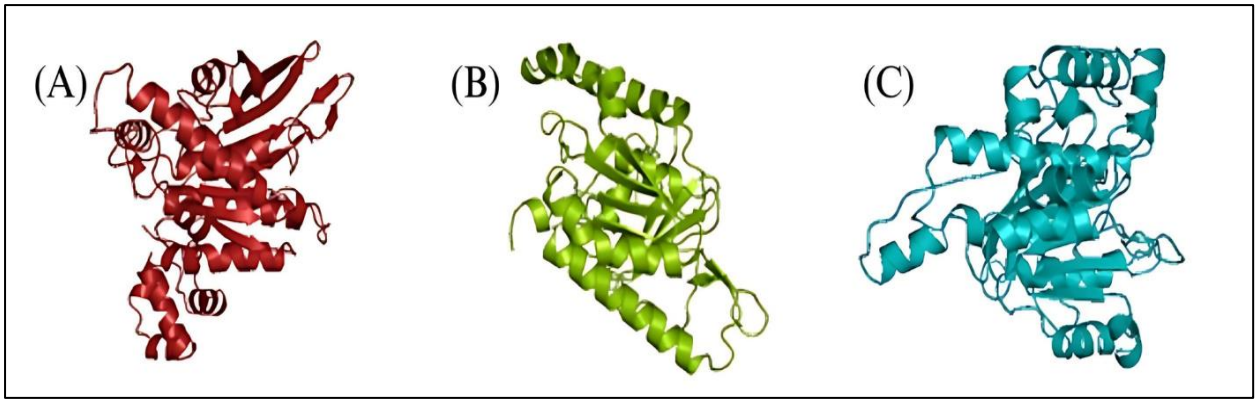

Figure 3: The 3D structure of the RAD51C-P2 homology models (A-C) generated by the Phyre2 webserver.

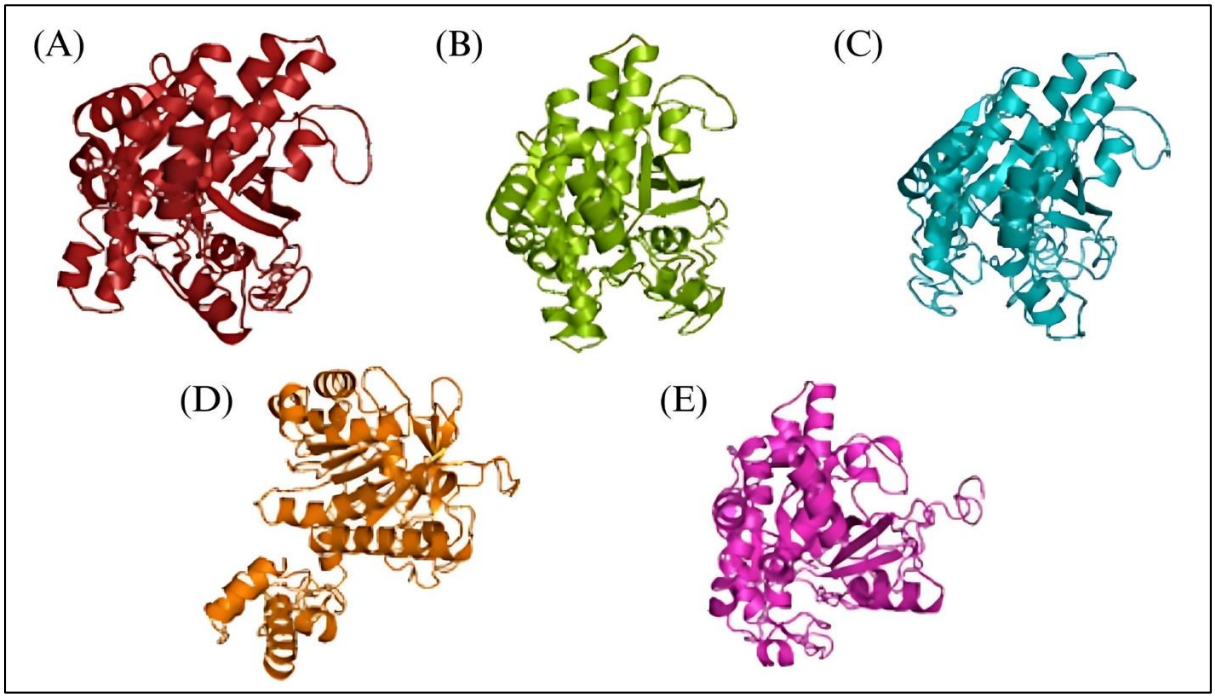

Figure 4: The 3D structure of RAD51C-IT homology models (A-E) generated by the I-TASSER webserver.

The results summary of the analyses from SWISS-MODEL and Phyre2 is tabulated in Table 1 and Table 2, respectively. The structural qualities of the protein models generated from these programs were evaluated and validated using the MolProbity and the PROSA programs. The distribution of amino acids in the Ramachandran plot was assessed using the PROCHECK program. RAD51C-P2 model emerged with the highest residues distributed in the most favored region of the Ramachandran plot (90.8%), followed by RAD51C-SM (90.2%) and RAD51C-IT (72.9%). In addition, the RAD51C-SM model also presented the lowest residues distributed in the disallowed region of the Ramachandran plot.

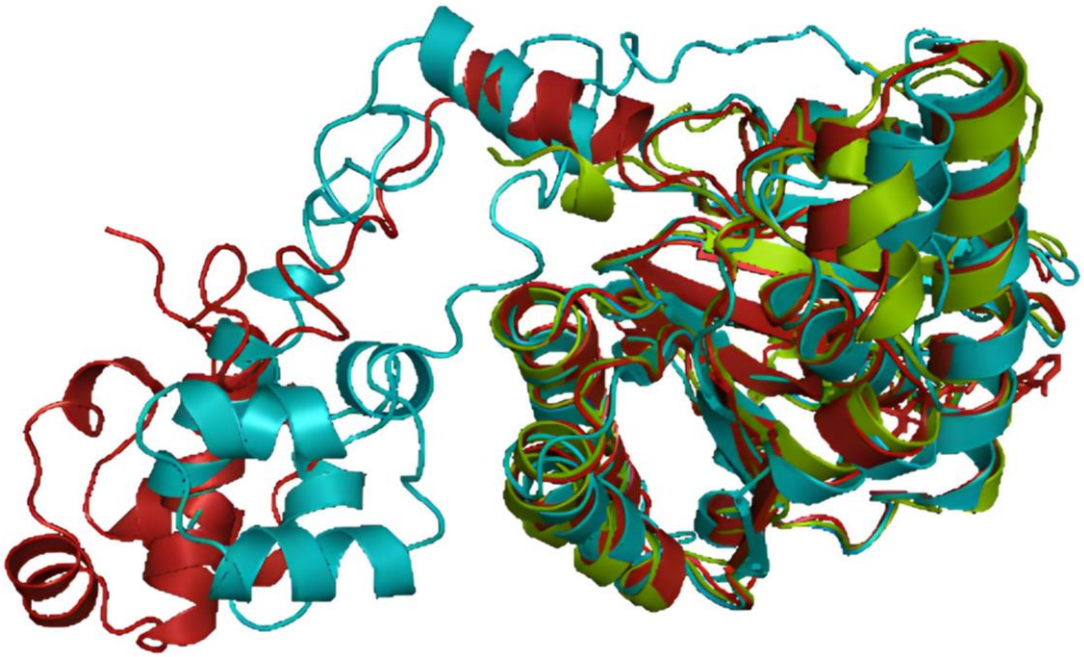

Figure 5: Superposition of RAD51C homology models. RAD51C-SM model is shown in firebrick, RAD51C-P2 is shown in split-pea and RAD51C-IT is shown in teal.

Table 1: Protein modeling results of RAD51C by the SWISS-MODEL.

Template	Sequence identity	Sequence similarity	Range	Coverage	GMQE	QMEAN
8GBJ.1.B	100.00	0.61	9-348	1.00	0.86	0.83
2ZUC.1.B	33.44	0.36	17-350	0.80	0.58	0.67
2ZUB.1.B	33.44	0.36	17-349	0.80	0.57	0.67

Table 2: Protein modeling results of RAD51C by Phyre2.

Model	Template	Aligned residues	Sequence identity	Coverage
1	c8fazC_	314	100%	84%
2	c8gjaC_	266	61%	70%
3	c1pznA_	291	36%	89%
4	c2dflA_	288	35%	89%
5	c5jzcG_	297	32%	88%
6	c3ldaA_	288	32%	90%
7	c1szpC_	271	33%	87%
8	c1t4g4_	284	31%	88%

3.3 Evaluation and Analysis of the Z-Score Graph

The overall quality of a homology model compared to the native protein structures was analyzed based on the value of the Z-score against its number of residues. The Z-score of the generated protein models is shown in Figure 6. The Z-score value below zero indicates the high quality of the predicted model (35). Protein model RAD51C-SM presented a Z-score value of -6.69 concerning its number of residues (333 residues), RAD51C-P2 (-6.92) concerning its number of residues (300 residues), and RAD51C-IT (-8.40) to its number of residues (376 residues). Based on the results, all generated models have been positioned inside the range of characteristics for native protein structures solved through X-ray crystallography and NMR. This result indicates that the generated protein models are of a good quality homology model.

3.4 Evaluation and Analysis of Local Model Quality Graph

The local model quality of the generated protein models is shown in Figure 7. The knowledge-based energies of the models' amino acids were plotted against their sequence positioning. Amino acids that reside in the positive energy region correlate with incorrect protein structure geometry. On the other hand, amino acids residing in the negative energy region were predicted to have a high-quality structure. Based on the results, all generated models presented good overall local model quality with some parts of the models predicted as problematic or erroneous. However, RAD51C-IT presented lower energy values compared to other protein models.

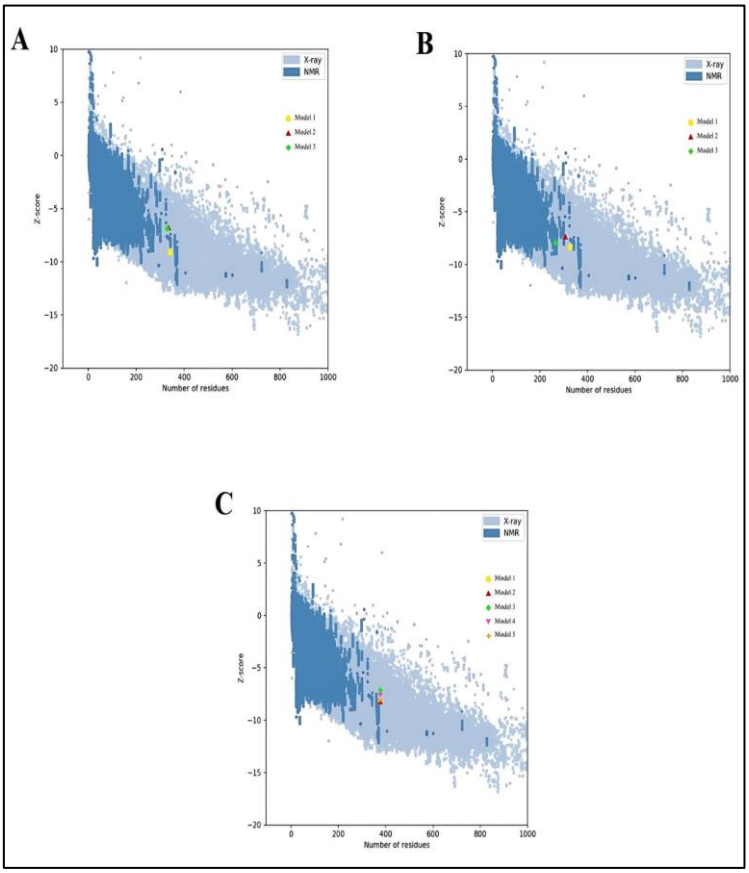

Figure 6: The Z-score of (A) RAD51C-SM, (B) RAD51C-P2, and (C) RAD51C-IT protein models.

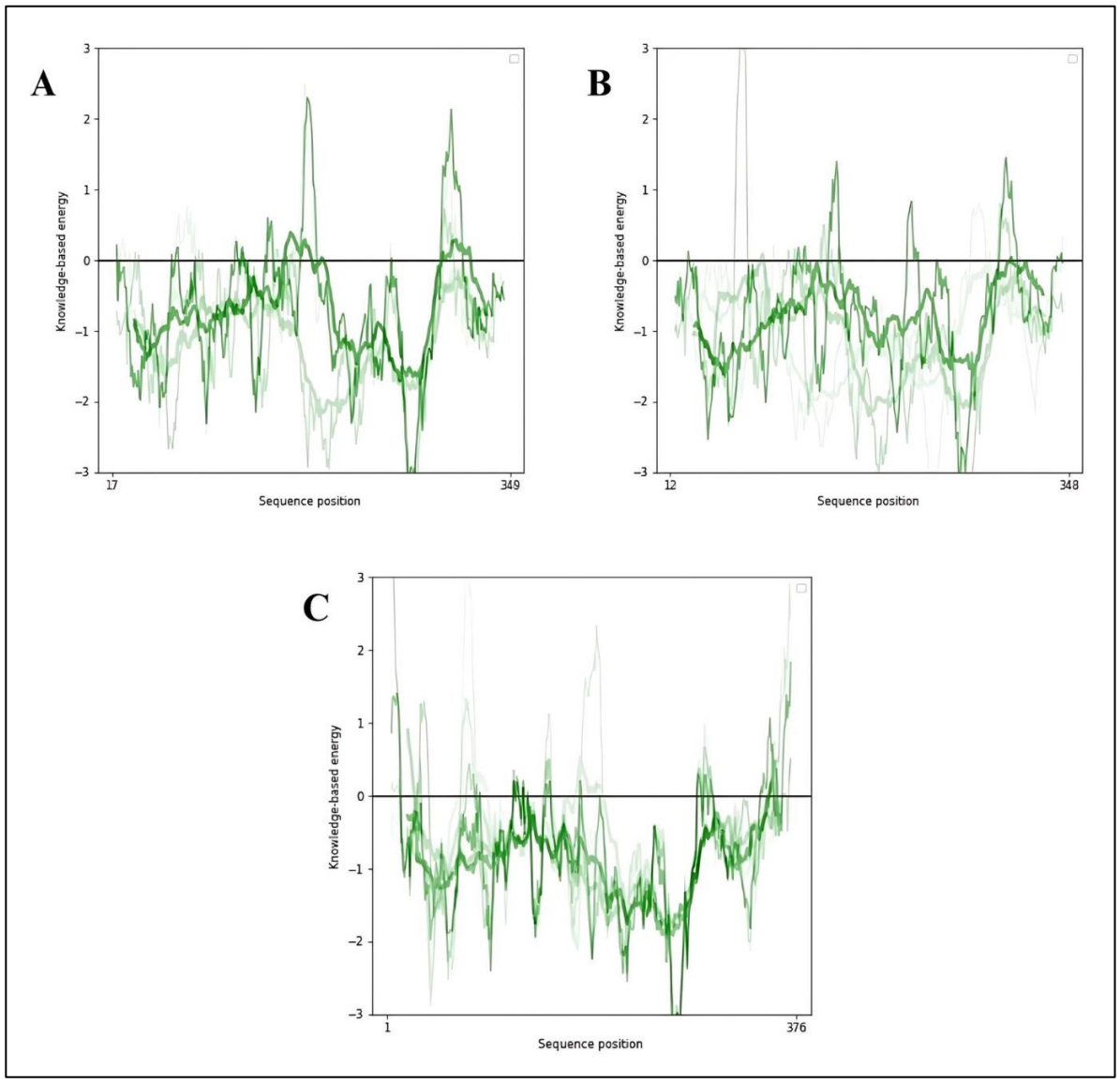

Figure 7: The local model qualities of (A) RAD51C-SM, (B) RAD51C-P2, and (C) RAD51C-IT protein models.

3.5 Evaluation and Analysis by Ramachandran Plot

Ramachandran plots were generated using the SAVES v6.0 software to validate predicted protein structures using parameters such as favored regions, allowed areas, and generously allowed regions of amino acid residues (34). The Ramachandran favored region is displayed in red, the additionally allowed region in yellow, the generously allowed region in faint yellow, and the disallowed region in white. Validation of the generated protein models using the Ramachandran plot analysis has shown that the RAD51C-SM model emerged with the highest residues distributed in the most favored region. However, this value does not pass the standard of a high-quality protein model which is 98.0% of the residues must reside in the most favored region while less than 0.20% of the residues in the disallowed region. Thus, refinement of the protein model can be further conducted through

energy minimization using molecular dynamics simulation. The overall result of the Ramachandran plots for SWISS-MODEL, Phyre2, and I-TASSER is shown in Figure 8. A comparison of the protein models against high-resolution structures conducted using the ProSA webserver showed that all generated protein models reside within the range of characteristics of native protein structures. These protein models were predicted to have lower structural errors as their size in amino acid numbers is within range of experimentally determined crystal structures. In addition, the local model quality of the generated protein models also indicates good quality models as most of their residues have been positioned in the negative region of the graph. However, it was observed that the RAD51C-IT model possessed lower energy values compared to other protein models.

It must be noted that only the RAD51C-IT protein model was generated based on the full-length RAD51C sequence (376 amino acids). Any missing residues from the homology models could result in misinformation in understanding the biological function of RAD51C on a molecular level. Based on the analyses of the structural qualities of the generated protein models, the RAD51C-IT model emerged as the most plausible RAD51C protein model compared to other models. Even though the distribution of RAD51C-IT residues in the most favored and disallowed regions of the Ramachandran plot is unfavorable, energy minimization of its structure can be conducted to rectify the errors. Thus, in this study, RAD51C-IT was predicted as the best homology model for RAD51C protein. RAD51C protein model quality assessment by ProSA and percentage of residues (%) are shown in Table 3 and amino acid positions of RAD51C protein and SAVES v6.0 for each predicted model are shown in Table 4.

Based on the results of the RAD51C protein model quality assessment by ProSA and the percentage of residues (Table 3), model 2 from Phyre2 has the highest percentage of residues in the favored region (90.8%) indicating the most stable confirmation of the predicted protein, followed by model 1 from SWISS-MODEL with 90.2% of residues, and model 1 from Phyre2 (89.0%).

4.0 Conclusion

Despite the importance of RAD51C protein in the homologous recombination (HR) pathway, its 3D protein structure remains elusive, and the active sites of RAD51C are unknown. Thus, in this study, three homology models have been successfully generated using three different protein modeling webservers. Evaluation and validation processes conducted on these models have revealed the RAD51C-IT model as the best model for the RAD1C protein. However, further refinement of the structure of the RAD51C-IT models can be conducted to rectify errors and improve the quality of the model. This protein model can be utilized as a virtual screening tool in discovering potential inhibitors of RAD51C protein.

Authorship contribution statement

RM: designed the study, planned the experiments, verified the analytical methods, and wrote the manuscript. **SSMM**: verified the analytical methods and wrote the manuscript. **SMJ**: verified the analytical methods and wrote the manuscript. **NADZ**: carried out the experiments and wrote the manuscript. **NBA**: carried out the experiments and wrote the manuscript.

Acknowledgment

This research was supported by the UiTM internal grant through the Geran Penyelidikan MyRA 2023 (600-RMC 5/3/GPM (039/2023)).

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript.

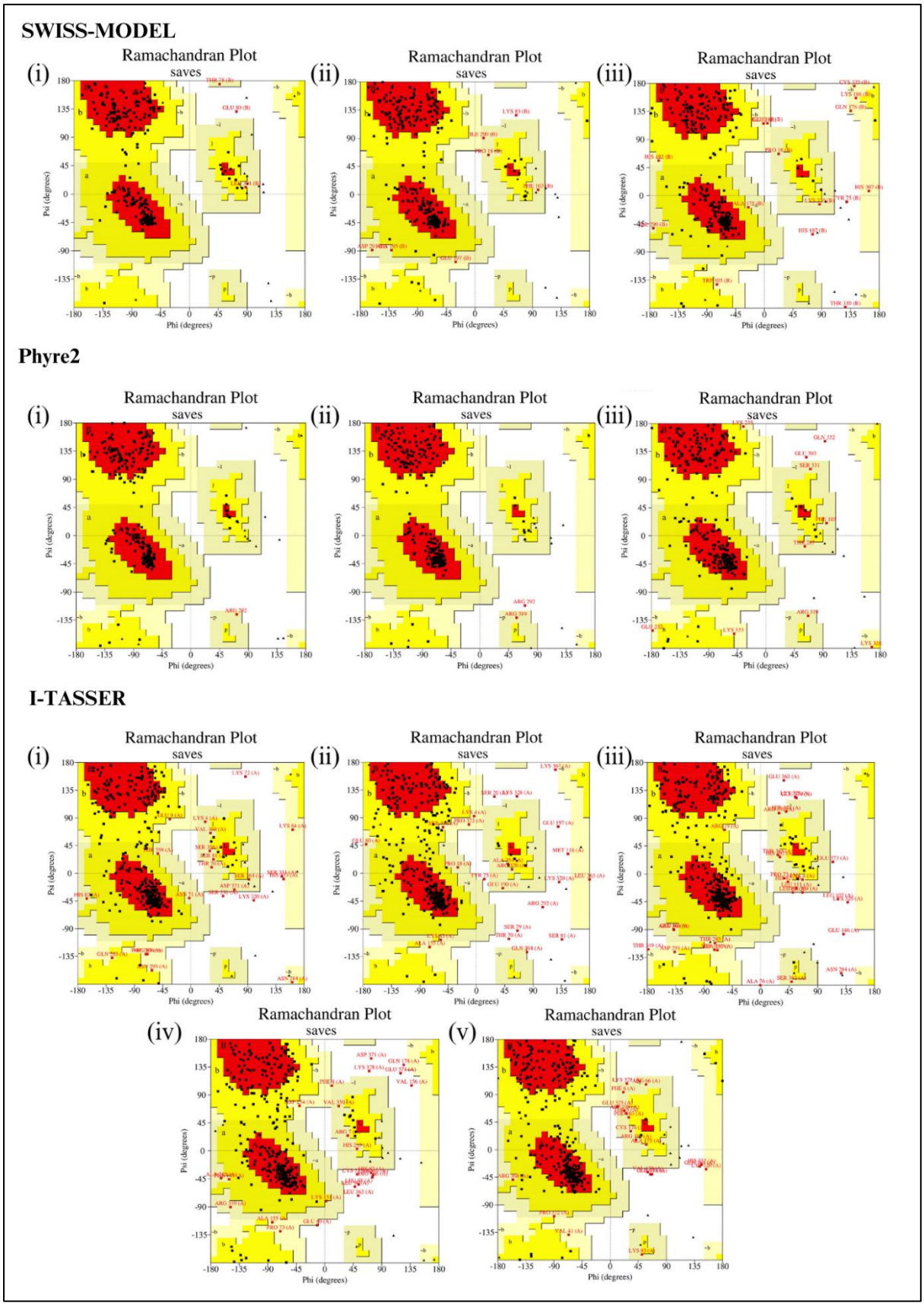

Figure 8: Ramachandran plots for the 3D protein models generated by (i) SWISS-MODEL, (ii) Phyre2, and (iii) I-TASSER webservers. Each subfigure shows the distribution of residues

in different regions: red regions indicate the most favored regions, yellow represents additional allowed regions, light yellow represents generously allowed regions, and white regions indicate disallowed regions. The quality of each protein model is assessed based on the percentage of residues present in each region.

Table 3: RAD51C protein model quality assessment by ProSA and percentage (%) of residues.

Webserver	Model	ProSA Z-score	Percentage of Residues (%)			
			Favored region	Additional allowed region	Generously allowed region	Disallowed region
SWISS-MODEL	1	-9.12	90.2	8.8	0.7	0.3
	2	-6.81	84.7	13.3	1.7	0.3
	3	-6.77	83.6	11.7	3.7	1.0
Phyre2	1	-8.38	89.0	10.7	0.3	0.0
	2	-7.96	90.8	8.4	0.4	0.4
	3	-7.40	82.1	14.2	2.9	0.7
I-TASSER	1	-8.16	72.9	20.6	5.0	1.5
	2	-8.22	69.6	24.2	2.7	3.5
	3	-7.12	71.1	20.9	5.0	2.9
	4	-7.66	68.7	24.5	3.2	3.5
	5	-8.04	72.6	21.8	4.1	1.5

Table 4: Amino acid positions of RAD51C protein and SAVES v6.0 for each predicted model.

Webserver	Model	Amino acids	SAVES v6.0	
			ERRAT (%)	Overall G-factors
SWISS-MODEL	1	340	86.67	-0.10
	2	334	91.92	-0.17
	3	333	85.40	-0.20
Phyre2	1	324	81.61	0.38
	2	266	79.92	0.26
	3	305	70.03	0.23
I-TASSER	1	376	87.23	-0.63
	2	376	79.29	-0.67
	3	376	85.91	-0.71
	4	376	84.51	-0.71
	5	376	89.95	-0.59

References

1. Kawale AS, Sung P. Mechanism and significance of chromosome damage repair by homologous recombination. Essays Biochem. 2020;64(5):779–790.
2. Louie BH, Kurzrock R. BAP1: Not Just a BRCA1-Associated Protein. Cancer Treat Rev. 2020;90(102091).
3. Bonilla B, Hengel SR, Grundy MK, Bernstein KA. RAD51 Gene Family Structure and Function. Annu Rev Genet. 2020;23(54):25–46.
4. Ismail IH, Davidson R, Gagné JP, Xu ZZ, Poirier GG, Hendzel MJ. Germline mutations in BAP1 impair its function in DNA double-strand break repair. Cancer Res. 2014;74(16):4282–4294.
5. Yu H, Pak H, Hammond-Martel I, Ghram M, Rodrigue A, Daou S, *et al*. Tumor suppressor and deubiquitinase BAP1 promotes DNA double-strand break repair. PNAS. 2014;111(1):285–290.
6. Zhao W, Steinfeld JB, Liang F, Chen X,

Maranon DG, Ma CJ, *et al*. Promotion of RAD51-mediated homologous DNA pairing by BRCA1-BARD1. Nature. 550(7676):360–365.

7. Mehta A, Haber JE. Sources of DNA double-strand breaks and models of recombinational DNA repair. Cold Spring Harb Perspect Biol. 2014;6(9):1–17.

8. Skoko JJ, Cao J, Gaboriau D, Attar M, Asan A, Hong L, *et al*. Redox regulation of RAD51 Cys319 and homologous recombination by peroxiredoxin 1. Redox Biol. 2022;56(102443):1–14.

9. Suwaki N, Klare K, Tarsounas M. RAD51 paralogs: Roles in DNA damage signaling, recombinational repair and tumorigenesis. Semin Cell Dev Biol. 2011;22(8):898–905.

10. Liao SG, Liu L, Wang YJ. Effect of RAD51C expression on the chemosensitivity of Eμ-Myc p19Arf-/- cells and its clinical significance in breast cancer. Oncol Lett. 2018;15(5):6107–6114.

11. Huang R, Zhou PK. DNA damage repair: historical perspectives, mechanistic pathways and clinical translation for targeted cancer therapy. Signal Transduct Target Ther. 2021;6(1):254.

12. Wang Z. Regulation of cell cycle progression by growth factor-induced cell signaling. Cells. 2021;10(3327).

13. Suszynska M, Ratajska M, Kozlowski P. BRIP1, RAD51C, and RAD51D mutations are associated with high susceptibility to ovarian cancer: Mutation prevalence and precise risk estimates based on a pooled analysis of ~30,000 cases. J Ovarian Res. 2020;13(50):1–11.

14. Wu R, Patel A, Tokumaru Y, Asaoka M, Oshi M, Yan L, et al. High RAD51 gene expression is associated with aggressive biology and with poor survival in breast cancer. Breast Cancer Res Treat. 2022;193(1):49–63.

15. Yang X, Song H, Leslie G, Engel C, Hahnen E, Auber B, *et al*. Ovarian and Breast Cancer Risks Associated with Pathogenic Variants in RAD51C and RAD51D. J Natl Cancer Inst. 2020;112(12):1242–1250.

16. Hameduh T, Haddad Y, Adam V, Heger Z. Homology modeling in the time of collective and artificial intelligence. Comput Struct Biotechnol J. 2020;18:3494–3506.

17. Husain S, Mohamed R, Abd Halim KB, Mohd Mutalip SS. Homology modeling of human BAP1 and analysis of its binding properties through molecular docking and molecular dynamics simulation. J Biomol Struct Dyn. 2023;41(15):7158–7173.

18. Eswar N, John B, Mirkovic N, Fiser A, Ilyin VA, Pieper U, *et al*. Tools for comparative protein structure modeling and analysis. Nucleic Acids Res. 2003;31(13):3375–3380.

19. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, *et al*. SWISS-MODEL: Homology modelling of protein structures and complexes. Nucleic Acids Res. 2018;46(W1):W296–303.

20. Michalska K, Joachimiak A. Structural genomics and the Protein Data Bank. J Biol Chem. 2021;296:100747.

21. Onawole AT, Sulaiman KO, Kolapo TU, Akinde FO, Adegoke RO. COVID-19: CADD to the rescue. Virus Res. 2020;285:198022.

22. Adelusi TI, Oyedele AQK, Boyenle ID, Ogunlana AT, Adeyemi RO, Ukachi CD, et al. Molecular modeling in drug discovery. Informatics Med Unlocked. 2022;29:100880.

23. Hasani HJ, Barakat K. Homology Modeling: an Overview of Fundamentals and Tools. Int Rev Model Simulations. 2017;10(2).

24. Atzrodt CL, Maknojia I, McCarthy RDP, Oldfield TM, Po J, Ta KTL, *et al*. A Guide to COVID-19: a global pandemic caused by the novel coronavirus SARS-CoV-2. FEBS J. 2020;287:3633–3650.

25. Nascimento IJ dos S, de Aquino TM, da Silva-Júnior EF. The New Era of Drug Discovery: The Power of Computer-aided Drug Design (CADD). Lett Drug Des Discov. 2022;19:951–955.

26. Iheagwam FN, Rotimi SO. Computer-aided analysis of multiple SARS-CoV-2 therapeutic targets: Identification of potent molecules from African medicinal plants. Scientifica (Cairo). 2020;1–25.

27. The UniProt Consortium. UniProt: The universal protein knowledgebase in 2023. Nucleic Acids Res. 2023;51:523–531.

28. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modelling, prediction and analysis. Nat Protoc. 2015;10(6):845–858.

29. Yang J, Zhang Y. I-TASSER server: New development for protein structure and function predictions. Nucleic Acids Res. 2015;43:W174–181.

30. Schrödinger L, DeLano W. PyMOL [Internet]. 2020. Available from: http://www.pymol.org/pymol

31. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, *et al*. MolProbity: All-atom structure validation for macromolecular crystallography. Acta Crystallogr. 2010;66:12–21.

32. Wiederstein M, Sippl MJ. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 2007;35(Suppl.2):407–410.

33. Colovos C, Yeates Todd O. Verification of protein structures: Patterns of nonbonded atomic interactions. Protein Sci. 1993;2:1511–1519.

34. Hollingsworth SA, Karplus PA. A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins. Biomol Concepts. 2010;1(3–4):271–283.

35. Zhang LI, Skolnick J. What should the Z-score of native protein structures be? Protein Sci. 1998;7:1201–1207.