Protein Profiling and 3d Structure Prediction Studies on Mutated HCC-Associated Catenin Beta-1 Gene

Priyadharshini M. and D. Leelavathi

Research Journal of Agricultural Sciences
An International Journal

P- ISSN: 0976-1675
E- ISSN: 2249-4538

Volume: 13
Issue: 04

Res. Jr. of Agril. Sci. (2022) 13: 1196–1200

CARAS

Res. Jr. of Agril. Sci. (Jul-Aug 2022)
13(4): 1196–1200

ISSN: 0976-1675 (P)
ISSN: 2249-4538 (E)

www.rjas.org

Full Length Research Article

Protein Profiling and 3D Structure Prediction Studies on Mutated HCC-Associated Catenin Beta-1 Gene

Priyadharshini M.*[1] and D. Leelavathi[2]

Received: 30 May 2022 | Revised accepted: 29 Jul 2022 | Published online: 03 Aug 2022

A B S T R A C T

Hepatocellular carcinoma (HCC) is considered to be the third leading cause of mortality due to cancer globally. Mutations in *CTNNB1* (Catenin Beta-1) genes are considered to be responsible for the development of HCC with inconstant frequencies based on the etiology. The main objective of this *in silico* research work is to predict the 3D structure of the mutated protein sequence of the CTNNB1 (target) and to explore the positive electrostatic surfaces of the modelled protein structure. In this study, we use proteomics servers for sequence retrieval. We then explore the 3D structure of CTNNB1 using an automated homology modelling server which is then validated. Further, we identify the potential electrostatic regions of CTNNB1 and analyze them using advanced *in silico* tools. All the information will be predicted in 3D form using molecular visualization tools. Our results clearly elucidate the presence of positive electrostatic surfaces on the predicted 3D structure of the hepatocellular carcinoma gene-coded mutated protein, CTNNB1. The identified results play an indispensable role in the molecular interaction of drugs and proteins. Based on the results, we can validate the efficiency of the drugs and predict drugs using structure-based drug designing technique.

***Key words*:** CTNNB1, Mutated HCC gene, Protein modelling, Positive electrostatic surfaces

Hepatocellular carcinoma (HCC), with 748,000 cases and 695,000 deaths in 2008, is the most common liver malignancy as well as the third cause of death from cancer worldwide [1]. The regions of higher incidence are Eastern and South-Eastern Asia as well as Middle and Western Africa, with age-standardized rates (ASR) ranging from 12.2 to 24.0 per 100,000 persons. In Europe, a remarkable North-South geographic gradient has been observed, with ASRs ranging from 3.8 per 100,000 men in Northern to 9.8 per 100,000 men in Southern Europe [1]. The highest incidence of HCC has been reported in men living in Southern Italy with a peak of 34.8 cases per 100,000 in the Campania region [2].

The majority of HCC cases are attributable to hepatitis B (HBV, 54%) and C (HCV, 31%) viruses, although a substantial geographic variation exists in the world. In Africa and Asia, where HBV is endemic, more than 60% of liver cancers are related to HBV infection, 20% are due to HCV and the remaining are caused by alcohol abuse and dietary exposure to aflatoxins. Conversely, in the United States, Europe, Egypt and Japan, more than 60% of HCCs are associated with HCV, 20% with HBV and the rest are mainly attributable to excessive alcohol consumption and liver metabolic diseases [3].

* **Priyadharshini M.**
✉ priyadharshinimk96@gmail.com

[1-2] P. G. and Research Department of Zoology, Ethiraj College for Women, Chennai - 600 008, Tamil Nadu, India

Both HBV and HCV, despite their distinct virologic features, are preferentially hepatotropic, not directly cytopathic, subvert the innate and adaptive immunity by several mechanisms and persist long into hepatocytes [4-5]. Chronic infections give rise to a complex, multistep process of hepatocarcinogenesis, lasting more than 30 years, that encompasses prolonged inflammation, hepatitis, regeneration, fibrosis, cirrhosis, dysplasia and finally HCC [6-8]. Cirrhosis is characterized by reduced hepatocyte proliferation, fibrosis, destruction of liver cells, and dysplastic nodules in around 15% of cirrhotic patients [9]. During this process, the molecular interaction between hepatitis viral products and the host cell machinery, with or without other environmental factors, contributes to the accumulation of a variety of genetic alterations which begin in preneoplastic lesions and increase in dysplastic hepatocytes and HCC [9]. Increased frequency of genetic damages may provide a progressive selective growth advantage of hepatocytes with a malignant phenotype, leading to phenotypically and genetically heterogeneous HCCs.

Mutations in CTNNB-1 genes play a significant role in the development of HCC [10]. Hence, CTNNB-1 genes were chosen as a target and the 3D structure of the mutated protein sequence of CTNNB1 was predicted to explore the positive electrostatic surfaces of the modelled protein structure, which shall be used in drug designing.

MATERIALS AND METHODS

Target selection

The protein sequence of the human CTNNB1 (Catenin beta-1 (UniProt ID: P35222) was retrieved based on the literature studies from the UniProt database in FASTA format (https://www.uniprot.org/uniprot/P35222.fasta).

3D structure prediction

The selected gene coded protein sequence of CTNNB1 was converted into the 3D structure using an automated homology protein modelling server called the Swiss Model server (https://swissmodel.expasy.org/).

Protein structure validation

After modelling, the 3D structure was validated using the ProCheck server(https://servicesn.mbi.ucla.edu/PROCHECK/) and the 3D structure was visualized using an advanced molecular visualization software called Discovery Studio Software.

Protein electrostatic patches prediction

The modelled 3D structure of the CTNNB1 (Catenin beta-1) was analyzed for protein surface electrostatic patches using the Patch Finder Plus tool (http://pfp.technion.ac.il/).

RESULTS AND DISCUSSION

>sp|P35222|CTNB1_HUMAN Catenin beta-1 OS=Homo sapiens OX=9606 GN=CTNNB1 PE=1 SV=1
MATQADLMELDMAMEPDRKAAVSHWQQQSYLDSGIHSGATTTAPSLSGKGNPEEEDVDTS
QVLYEWEQGFSQSFTQEQVADIDGQYAMTRAQRVRAAMFPETLDEGMQIPSTQFDAAHPT
NVQRLAEPSQMLKHAVVNLINYQDDAELATRAIPELTKLLNDEDQVVVNKAAVMVHQLSK
KEASRHAIMRSPQMVSAIVRTMQNTNDVETARCTAGTLHNLSHHREGLLAIFKSGGIPAL
VKMLGSPVDSVLFYAITTLHNLLLHQEGAKMAVRLAGGLQKMVALLNKTNVKFLAITTDC
LQILAYGNQESKLIILASGGPQALVNIMRTYTYEKLLWTTSRVLKVLSVCSSNKPAIVEA
GGMQALGLHLTDPSQRLVQNCLWTLRNLSDAATKQEGMEGLLGTLVQLLGSDDINVVTCA
AGILSNLTCNNYKNKMMVCQVGGIEALVRTVLRAGDREDITEPAICALRHLTSRHQEAEM
AQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLALCPANHAPLREQGAIPRLVQLL
VRAHQDTQRRTSMGGTQQQFVEGVRMEEIVEGCTGALHILARDVHNRIVIRGLNTIPLFV
QLLYSPIENIQRVAAGVLCELAQDKEAAEAIEAEGATAPLTELLHSRNEGVATYAAAVLF
RMSEDKPQDYKKRLSVELTSSLFRTEPMAWNETADLGLDIGAQGEPLGYRQDDPSYRSFH
SGGYGQDALGMDPMMEHEMGGHHPGADYPVDGLPDLGHAQDLMDGLPPGDSNQLAWFDTDL

Fig 1 CTNNB1_HUMAN: Uniprot database

The above picture represents the amino acid sequence of CTNNB1- Catenin beta-1 in FASTA format.

10	20	30	40	50	60
MATQADLMEL	DMAMEPDRKA	AVSHWQQQSY	LDSGIHSGAT	TTAPSLSGKG	NPEEEDVDTS
70	80	90	100	110	120
QVLYEWEQGF	SQSFTQEQVA	DIDGQYAMTR	AQRVRAAMFP	ETLDEGMQIP	STQFDAAHPT
130	140	150	160	170	180
NVQRLAEPSQ	MLKHAVVNLI	NYQDDAELAT	RAIPELTKLL	NDEDQVVVNK	AAVMVHQLSK
190	200	210	220	230	240
KEASRHAIMR	SPQMVSAIVR	TMQNTNDVET	ARCTAGTLHN	LSHHREGLLA	IFKSGGIPAL
250	260	270	280	290	300
VKMLGSPVDS	VLFYAITTLH	NLLLHQEGAK	MAVRLAGGLQ	KMVALLNKTN	VKFLAITTDC
310	320	330	340	350	360
LQILAYGNQE	SKLIILASGG	PQALVNIMRT	YTYEKLLWTT	SRVLKVLSVC	SSNKPAIVEA
370	380	390	400	410	420
GGMQALGLHL	TDPSQRLVQN	CLATLRNLSD	AATKQEGMEG	LLGTLVQLLG	SDDINVVTCA
430	440	450	460	470	480
AGILSNLTCN	NYKNKMMVCQ	VGGIEALVRT	VLRAGDREDI	TEPAICALRH	LTSRHQEAEM
490	500	510	520	530	540
AQNAVRLHYG	LPVVVKLLHP	PSHWPLIKAT	VGLIRNLALC	PANHAPLREQ	GAIPRLVQLL
550	560	570	580	590	600
VRAHQDTQRR	TSMGGTQQQF	VEGVRMEEIV	EGCTGALHIL	ARDVHNRIVI	RGLNTIPLFV
610	620	630	640	650	660
QLLYSPIENI	QRVAAGVLCE	LAQDKEAAEA	IEAEGATAPL	TELLHSRNEG	VATYAAAVLF
670	680	690	700	710	720
RMSEDKPQDY	KKRLSVELTS	SLFRTEPMAW	NETADLGLDI	GAQGEPLGYR	QDDPSYRSFH
730	740	750	760	770	780
SGGYGQDALG	MDPMMEHEMG	GHHPGADYPV	DGLPDLGHAQ	DLMDGLPPGD	SNQLAWFDTDL

Fig 2 Mutated sequence of CTNNB1

The above picture shows the amino acid sequence of CTNNB1 with the mutated position highlighted in yellow.

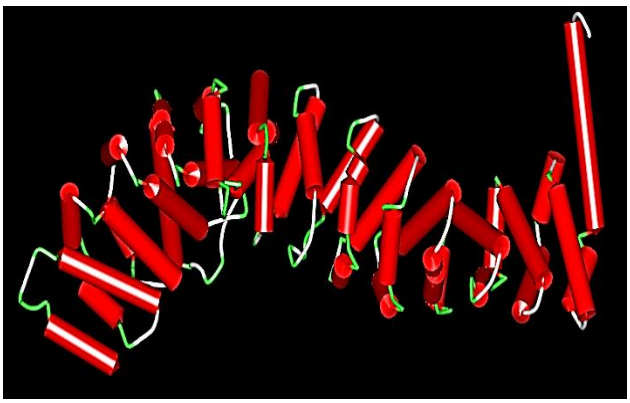

Fig 3 3D structure of CTNNB1 Catenin beta-1 protein

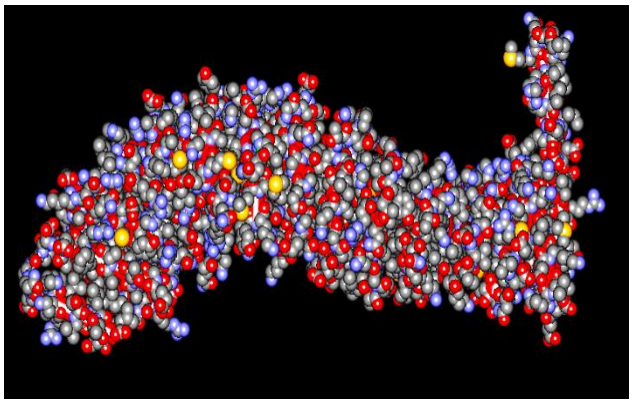

Fig 4 3D structure of the CTNNB1 Catenin beta-1 protein

The above picture represents the 3D structure of the Mutated CTNNB1 protein viewed in Discovery Studio software in the Schematic model with secondary model colour.

The above picture represents the 3D structure of the Mutated CTNNB1 protein viewed in Discovery Studio software in spacefill model with coloured atoms.

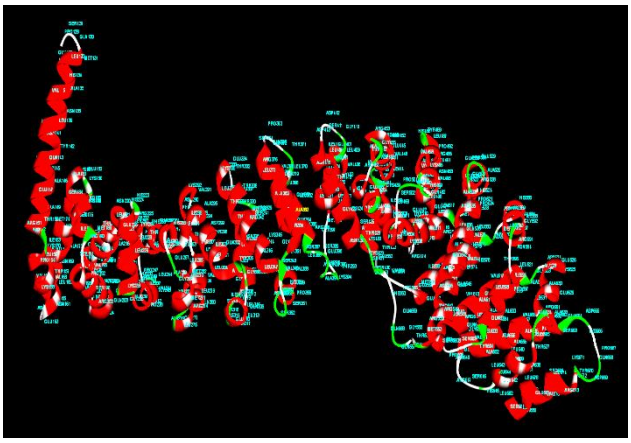

Fig 5 3D structure of the CTNNB1 Catenin beta-1 protein

The above picture represents the 3D structure of the Mutated CTNNB1 protein viewed in Discovery Studio software in a ribbon model with mutated amino acid ALA-383 being highlighted in yellow.

In the current *in silico* investigation, the target protein sequence of CTNNB1 (P35222) was retrieved from the UniProt database in FASTA format (Fig 1) to predict the 3D structure and identify the Electrostatic surfaces. Fig. 3, and 4,5 visualize the Mutated 3D structure of the modelled protein target CTNNB1 Catenin beta-1. The Swiss Model server was used to predict the 3D structure. Somatic activating mutations in CTNNB1 have been detected in several tumours, but the highest frequency is observed in hepatocellular carcinoma (HCC). These mutations frequently disrupt the normal protein turnover of β-catenin and result in its aberrant accumulation in cancer cells [11].

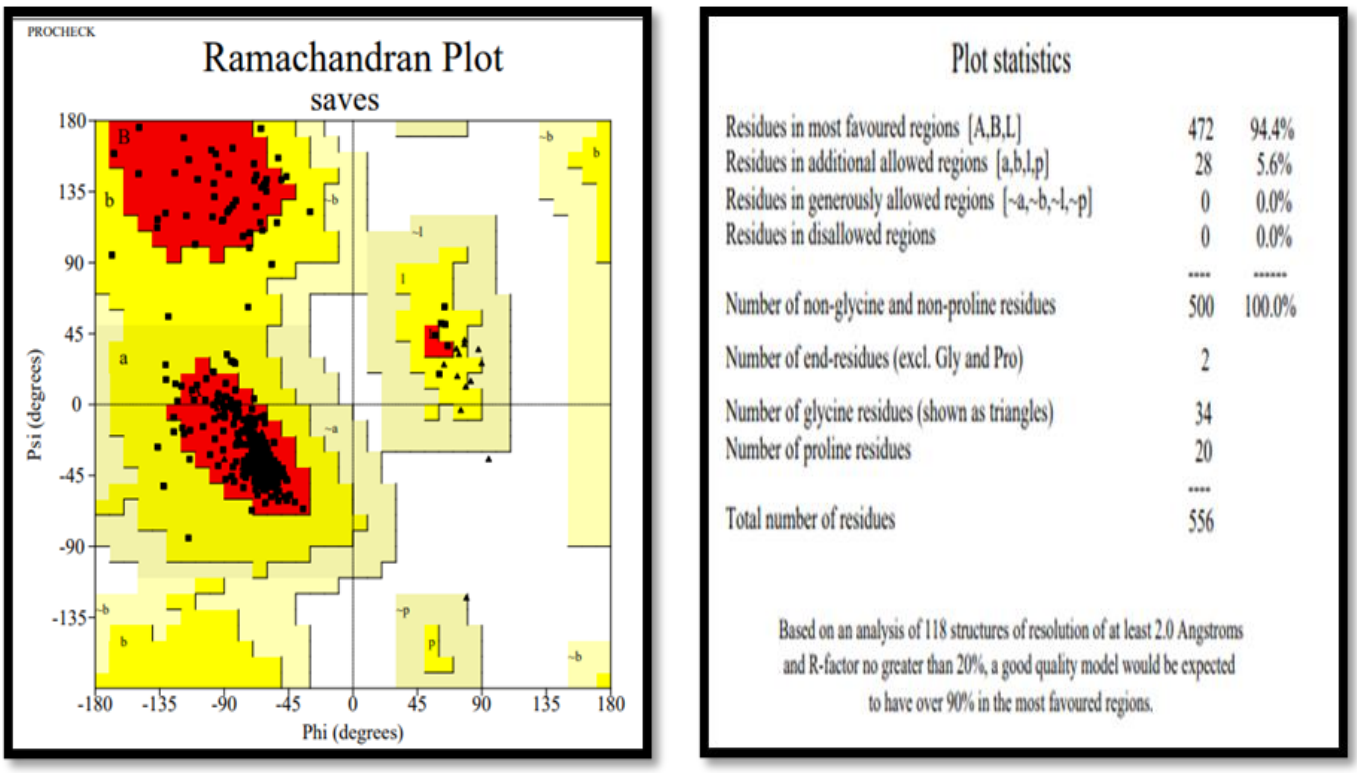

Fig 6 3D structure validation- ProCheck server

The above pictures show the assessment of the Ramachandran Plot for the modelled protein structure of CTNNB1.

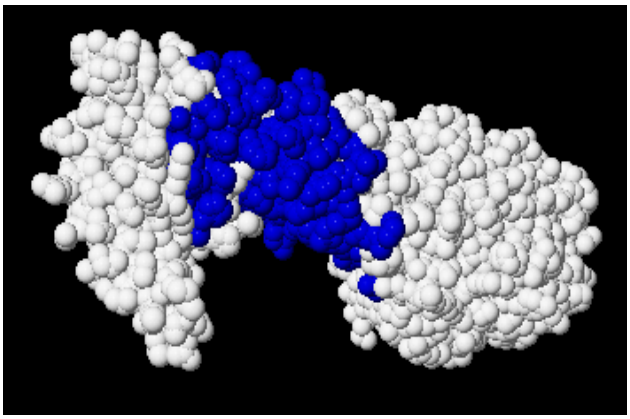

Fig 7 3D structure of the CTNNB1 Catenin beta-1 protein (Electrostatic surface model)

The above picture represents the 3D protein structure of CTNNB1 in the space fill model with the potential electrostatic surface (Blue Colour) viewed in blue colour using Protein Patch Finder plus tool.

In protein modelling, SWISS-MODEL was used to convert the amino acid sequence of CTNNB1 into a 3D structure [Fig 2-4]. SWISS-MODEL [12-15] was used to analyse the molecular and structural details of CTNNB1 elaborately for docking. SWISS-MODEL is a server for automated comparative modelling of three-dimensional (3D) protein structures. Waterhouse et al. computed models by the SWISS-MODEL server homology modelling pipeline which is based on ProMod3, an in-house comparative modelling engine based on Open Structure. The modelled 3D protein was thoroughly estimated using the ProCheck server [16] for assessment of the Ramachandran Plot. After modelling, the 3D structure of the mutated protein was validated using the ProCheck server. (Fig 6) shows the assessment of the Ramachandran Plot which confirms that there is no error (79.9%) in the modelled protein.

(Fig 7) clearly shows the protein structure of the CTNNB1 with the Electrostatic surface regions shown in blue colour and the overall 3D protein structure was visualized in the spacefill model using protein Patch finder plus.

Patch Finder Plus is an automatic server to extract and display the largest positive electrostatic patch on a protein surface. The PFplus [17] algorithm automatically assigns surface positive patches by looking for adjacent points on the protein surface that meet a given electrostatic potential cutoff. The algorithm is built on five major steps: (i) calculating the electrostatic potential of the protein on a 3D grid, using the Poisson–Boltzmann equation. (ii) Defining the grid points that fall closest to the protein surface while emitting all non-surface points. (iii) Extract all 3D patches of adjacent grid points that meet the defined cutoff. (iv) Choosing the largest positive patch for each protein chain. (v) Assigning the protein residues related to the patch.

The identified positive electrostatic surfaces are as follows: ARG274 TYR306 THR289 GLN309 VAL346 LYS288 TYR331 ARG376 ALA383 TRP338 ASN380 GLU334 TYR333 LEU252 THR339 ILE303 SER250 ASP299 ASN220 ASP249 LYS345 ASN308 TYR254 GLN379 ASN415 HIS223 MET271 PRO373 ARG342 ALA295 VAL208 GLY307 ASN287 GLN266 ARG212 VAL349 CYS419 HIS219 LYS292 VAL378 ASN290 CYS350 VAL291 SER374 THR332 CYS213 VAL248 PHE253 HIS176 ILE296 GLN302 LYS335 GLY216 ILE256 ALA284 HIS260 LYS270 PHE293 PRO247. Interestingly, the mutated amino acid Alanine 383 is present among the positive electrostatic regions. This identified amino acid region acts as a potential binding site for drugs.

CONCLUSION

In this research work, we observe that the positive patches present on the CTNNB1 protein surface indicate, protein functions such as nucleic acid and membrane binding. The overall results clearly illustrate that the mutated region is present on the positive patches of the electrostatic regions of CTNNB1 protein. Identification of drug targets (antigenic binding sites) is a fundamental challenge for pharmaceutical industries. Hence, our finding would be propitious in the field of computational biology; explicitly in structure-based drug designing.

Acknowledgement

The authors acknowledge the help extended by Dr. Balaji Munivelan, Ph. D., CEO and Senior Bio-informatician (bioinfobalaji@gmail.com) ABS Geno-informatics, Chennai, for his contribution towards Insilico drug docking studies.

LITERATURE CITED

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. 2010. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *International Journal of Cancer* 127(12): 2893-2917.
2. Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, Boyle P. 2007. *Cancer Incidence in Five Continents.* Volume IX. IARC Press, International Agency for Research on Cancer, 2007.
3. El-Serag HB. 2012. Epidemiology of viral hepatitis and hepatocellular carcinoma. *Gastroenterology* 142(6): 1264-1273.
4. Protzer U, Maini MK, Knolle PA. 2012. Living in the liver: hepatic infections. *Nature Reviews Immunology* 12(3): 201-213.
5. Rehermann B, Nascimbeni M. 2005. Immunology of hepatitis B virus and hepatitis C virus infection. *Nature Reviews Immunology* 5(3): 215-229.
6. Roberts LR, Gores GJ. 2005. Hepatocellular carcinoma: molecular pathways and new therapeutic targets. *In*: *Seminars in Liver Disease* 25(2): 212-225. Copyright© 2005 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA.
7. Arzumanyan A, Reis HM, Feitelson MA. 2013. Pathogenic mechanisms in HBV-and HCV-associated hepatocellular carcinoma. *Nature Reviews Cancer* 13(2): 123-135.
8. Tsai WL, Chung RT. 2010. Viral hepatocarcinogenesis. *Oncogene* 29(16): 2309-2324.
9. Thorgeirsson SS, Grisham JW. 2002. Molecular pathogenesis of human hepatocellular carcinoma. *Nature Genetics* 31(4): 339-346.
10. Tornesello ML, Buonaguro L, Tatangelo F, Botti G, Izzo F, Buonaguro FM. 2013. Mutations in TP53, CTNNB1 and PIK3CA genes in hepatocellular carcinoma associated with hepatitis B and hepatitis C virus infections. *Genomics* 102(2): 74-83.
11. von Kries JP, Winbeck G, Asbrand C, Schwarz-Romond T, Sochnikova N, Dell'Oro A, Behrens J, Birchmeier W. 2000. Hot spots in β-catenin for interactions with LEF-1, conductin and APC. *Nature Structural Biology* 7(9): 800-807.

12. Bienert S, Waterhouse A, de Beer TA, Tauriello G, Studer G, Bordoli L, Schwede T. 2017. The SWISS-MODEL Repository—new features and functionality. *Nucleic Acids Research* 45(D1): 313-319.
13. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R. 2018. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research* 46(W1): W296-W303.
14. Guex N, Peitsch MC, Schwede T. 2009. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* 30(S1): S162-S173.
15. Studer G, Tauriello G, Bienert S, Biasini M, Johner N, Schwede T. 2021. ProMod3—A versatile homology modelling toolbox. *PLoS Computational Biology* 17(1): p.e1008667.
16. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* 26(2): 283-291.
17. Shazman S, Celniker G, Haber O, Glaser F, Mandel-Gutfreund Y. 2007. Patch Finder Plus (PFplus): a web server for extracting and displaying positive electrostatic patches on protein surfaces. *Nucleic Acids Research* 35(Suppl_2): W526-W530.