Open Acess
Jurnal Biota Vol. 10 No. 2 (2024) https://doi.org/10.19109/Biota.v10i2. 21681

Research Article

Functional 3D Structure Analysis of Quasispecies Variants of Hepatitis B Virus Surface and Core Protein in Advanced Liver Disease and Chronic HBV Infection Patients in Indonesia: In Silico

Samuel Johanes Aldrian Mahulette[1], Adhisa Fathirisari Putri[1], Azki Afidati Putri Anfa[1], Yoshihiko Yano[2], Jajar Setiawan[3], Wahyu Aristyaning Putri[1*]

[1] Department of Tropical Biology, Faculty of Biology, Universitas Gadjah Mada, Yogyakarta, Indonesia

[2] Department of Gastroenterology, Kobe University Hospital, Kobe, Japan

[3] Department of Physiology, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia

*Email: wahyuaristyaningputri@ugm.ac.id

Received: 23/01/2024	Revised: 29/05/2024	Accepted: 30/05/2024

ABSTRACT

Hepatitis B Virus (HBV) is an endemic virus and belongs to Hepadnaviridae family. This virus can result in variations of quasispecies due to its high rate of mutation. A quasispecies variant is a small population and develops as a result of mutation and can become a wild-type population. This research aims to study and carry out 3D modeling on 12 in-house full sequence HBV genome isolates from Indonesia and obtain predictive visualization data to become a reference for further research leading to the production of anti-virals and natural treatments for HBV. 12 in-house full HBV genome sequences obtained from previous research were used to carry out 3D modeling and structural analysis of the surface protein, core protein, and polymerase protein. Analysis was carried out in silico using programs available online. Phylogenetic analysis was carried out using MEGA11, translation of nucleotides into protein sequences using the ExPAsy Translate portal, physiochemical analysis using ProtParam portal, and functional domain testing using the MOTIF tool from GenomeNet. Then 3D modelling using Phyre2 and SWISS-MODEL. The major mutation of the S protein occurs in L21S and mutations in the C protein mainly occur in P79Q and S87G. The model for S Protein from homology structure prediction is not reliable thus it still needs more templates from experimental techniques. While C Protein structure prediction can provide information for further research in alternative natural antiviral treatment.

Keywords: *3D model; Endemic; Genome; HBV; Protein.*

Introduction

Hepatitis B Virus (HBV) is a leading cause of liver damage and liver tissue cancer growth. Over 300 million people worldwide suffer from Chronic Hepatitis B (CHB), resulting in 4 million deaths each year [1-3]. Asia is home to over 70% of total HBV cases globally, with estimates suggesting this percentage [4-6]. In a joint research effort in 2016, conducted by [7], meta-analysis and modeling were used to estimate the prevalence of HBsAg in 120 countries. This research revealed that Indonesia ranked fourth in HBsAg-positive infections.

Indonesia, a Southeast Asian country, experiences high HBV endemicity due to its unique archipelagic geography [5], [6]. This geographical isolation limits

the availability of comprehensive databases for analyzing HBV variants' pathogenicity and clinical characteristics. Nevertheless, in many cases of advanced liver disease (ALD) and chronic HBV infection (CHB) among Indonesian patients, new subgenotypes and genotype variants have been discovered [8]. Additionally, researchers have identified 12 whole genome sequences of HBV quasispecies in ALD and CHB patients in Java, all belonging to the dominant B3 genotype in the Southeast Asian region, particularly Indonesia [3].

Patients infected with HBV will be identified through laboratory diagnosis, which involves checking the patient's serological status using markers or a combination of several antigen markers, such as HB surface antigen (HBsAg), HB core antigen (HBcAg), HBeAg, and/or antibody markers like HB surface antibody (anti-HBs/HBsAb), HB core antibody (anti-HBc), HB e antibody (anti-HBe), and anti-HBc IgM [1], [4]. The severity of HBV infection can also be determined by conducting the HBsAg, HBeAg/anti-HBe, and HBV DNA tests, followed by assessing blood parameters such as aspartate aminotransferase (AST) and alanine transaminase (ALT). Additionally, transient elastography (Fibroscan) or needle liver biopsy can be used as non-invasive and invasive methods for detecting liver cirrhosis [1].

HBV infection can lead to fibrosis in hepatocyte cells, and in chronic conditions, it can progress to liver cirrhosis (LC). In the long term, HBV infection can also increase the risk of hepatocellular carcinoma (HCC) or liver cancer [1], [3], [4], [9]. Liver cirrhosis is a condition affecting liver tissue, characterized by the accumulation of regenerative nodules surrounded by fibrous fibers, resulting from chronic liver tissue damage. This condition is preceded by collagen or fibrosis encapsulation of damaged tissue [10].

In previous research, quasispecies were observed in the S and X regions of the HBV ORFs [3]. Quasispecies variants are created due to viral replication mutations, leading to diverse virus populations. These quasispecies populations are categorized into minor populations (1-5%), intermediate populations (5-20%), and major populations (>20%). When a mutant variant constitutes more than 80% of the population, it is considered to have replaced the wild-type strain [3], [11]. HBV is a virus that has diverse DNA. The HBV genome consists of the X gene (X), precore/core gene (preC/C), presurface antigen/surface antigen gene (preS/S), and polymerase gene (P) that control the replication and transcription [12], [13]. The HBV-X protein (HBx), is involved in the pathological process of HBV and plays a major role in HBV replication by either promoting viral replication or changing host gene expression linked to HCC [12]. HBx mutations have been implicated in the pathophysiology of HBV and are essential for the development of HCC [12], [14]. The core promoter region (preC/C) contributes to the pathogenicity, morphogenesis, and essential for replication of the virus [15]. The S gene (preS/S) encodes a series of surface antigen polypeptides that are embedded inside the viral envelope, whereas the P gene encodes the reverse transcriptase of the virus [12]. S gene mutations have the potential to impact HBsAg secretion, immunogenicity, and antigenicity, and complicate illness diagnosis [16]. Models of protein structure can be examined in silico. Utilizing in silico analysis, it was possible to deduce the function of proteins, identify potential binding sites and partners for those interactions, create or enhance new enzymes or antibodies, and explain the effects of current mutations [17]. SWISS-MODEL and Phyre2 can used for protein modelling [18]. In this study, we used 12 in-house full genome sequences of HBV, including HBV genotype B3, to compare predicted protein structures. These sequences were obtained from [3], with

eight samples originating from LC and/or HCC patients and four from CHB patients. Researchers suspect that the tertiary structure of the protein surface will exhibit significant differences. However, the differences from the wild-type genotype B3 may need to be more pronounced in the core protein and polymerase protein.

Materials and methods

1. Retrieving in-house sequences

In the previous research, thirty hepatitis B surface antigen (HBsAg)-positive were collected at the General Hospital of Surabaya and Hajj Hospital (Surabaya, Indonesia), but only twelve samples were successfully analyzed for the full genome. Twelve in-house isolates of genotype B3 were obtained from previous research [3]. In this study, HBV genotype B3 isolates were also used as the wild type. HBV DNA isolation was carried out using the Qiagen DNA blood prep kit. The complete HBV genome was isolated using multiple primer sets. Purification was conducted using the gel purification method [19]. Direct sequencing was performed at PT. Genetika Science. The HBV genotype B3 sequence from GenBank and the sequences of the 12 HBV isolates were used for phylogenetic analysis. This analysis utilized the Clustal W tool from GenomeNet. Subsequently, a phylogenetic tree was constructed using Molecular Evolutionary Genetics Analysis (MEGA) Software version 11.0.8 with 1000 bootstrap reconstructions [3].

2. Protein modelling

The wild-type isolates and 12 in-house isolates were translated using the ExPASy (Expert Protein Analysis System) Website Portal provided by the Biozentrum at the University of Basel, Switzerland (https://www.expasy.org/). This bioinformatics portal and its tools are hosted on an online server managed by the Swiss Institute of Bioinformatics (SIB). The translation was performed by selecting the translate portal.

Following translation, the ProtParam portal was utilized to conduct a physicochemical analysis of the proteins translated from the 12 in-house and single wild-type isolates. The physicochemical analysis covered molecular weight, amino acid composition, instability index, half-life, aliphatic index, theoretical pI, and extinction coefficient [20]. Subsequently, the protein sequences of the 12 in-house isolates and the wild-type isolate were assessed for functional domains using the MOTIF tool from GenomeNet (https://www.genome.jp/tools/motif/). MOTIF sequences consist of amino acid compositions that may hold biological significance [21].

SWISS-MODEL server, used for modeling and viewing ANOLEA and QMEAN6 values. ANOLEA measures model packaging quality by estimating the average empirical atomic force magnitude, while QMEAN6 provides a global and local assessment of the Qualitative Model Energy Analysis [22]. Homology modeling and evaluation of 3D structural models were conducted for 12 in-house isolates and wild-type isolates using the Phyre2 website server hosted by the Imperial College of London. Target isolates were submitted to the Phyre2 server in intensive modeling mode, and the modeling process relied on data retrieved from the Protein Data Bank (PDB). Alignment scores were generated for the top 10 structures, which were then used to create 3D models. Model quality was assessed based on their similarity to the template model [23].

Results and Discussion

Results

1. Mutation mapping

Sequences of S and C proteins from eight samples from [3] conducted using SWISS-MODEL software are shown in Table 1 and Table 2. Mutations in the S and C proteins occur sporadically but are more common in ALD than in CHB patients. The major mutation of the S protein occurs in L21S and mutations in the C protein mainly

occur in P79Q and S87G (Table 1). The results of this study are slightly different from research by Abavisani [24], where the hot mutations in the S protein from HBV-CHB patients were S53L (37,7%), A184V/G (39,3%), and S210K/N/R/S (39,3%), while L21S was only around 29,1%. In other research from Ahmed et al [25], mutations in ALD patients were more common in T1631C (65,6%), and no mutations in L21S, P79Q, and S87G.

Table 1. Protein sequence alignment and mutation mapping of the S Protein. 12 In-house S Protein Query were align with HBV Genotype B3 (AB713527.1)

Genotype B3	A	L	Q	N	G	L	Q	S	C
A1 ALD	T	.	.	.	.	.	.	.	.
B1 ALD	.	.	.	.	E	.	.	.	.
C1 ALD	S	S	.	.	.	.	.	.	.
D1 ALD	.	.	.	.	.	.	.	.	Y
A2 ALD	.	S	R	.	E	.	.	.	.
C2 ALD	T	S	.	S	.	.	.	N	.
D2 ALD	.	.	.	.	.	.	.	.	.
E2 ALD	.	.	.	.	.	.	.	.	.
F1 CHB	.	.	.	.	.	.	P	.	.
G1 CHB	.	.	.	.	.	.	.	.	.
B2 CHB	.	.	.	.	.	.	.	.	Y
F2 CHB	.	S	.	.	E	P	.	.	.

Table 2. Protein sequence alignment and mutation mapping of the C Protein. 12 In-house S Protein Query were align with HBV Genotype B3 (AB713527.1)

Genotype B3	P	V	I	L	Y	S	I	S	E	P	E	S
A1 ALD	.	.	.	.	.	T	.	G	.	Q	.	.
B1 ALD	T	.	.	.	.	.	.	.	.	Q	D	G
C1 ALD	.	A	.	.	.	.	.	.	.	.	.	G
D1 ALD	.	.	.	.	.	.	.	.	.	.	.	.
A2 ALD	T	.	.	.	.	.	.	.	.	Q	D	G
C2 ALD	.	.	V	.	.	T	V	.	Q	.	.	N
D2 ALD	.	.	.	.	.	.	.	.	.	.	.	.
E2 ALD	.	.	.	.	.	.	.	.	.	.	.	.
F1 CHB	T	.	.	.	.	.	.	.	.	Q	.	G
G1 CHB	.	.	.	.	.	.	.	.	.	.	.	.
B2 CHB	.	.	.	.	.	.	.	.	Q	Q	.	.
F2 CHB	.	.	.	M	F	.	.	.	.	.	.	G

2. Protein structure prediction SWISS-MODEL

Protein structure prediction with SWISS-MODEL on the S protein produces two models, each containing two structures. Model 1 is built based on the Woodchuck Hepatitis Virus homodimer, and Model 2 is built based on the heterodimer Tumor necrosis factor receptor superfamily member five and only covers 33 aa (118-151) in the MHR and the a-determinant part of the S protein. Model C protein has a homo-tetramer oligo state. The outside of Model C protein consists of the least hydrophobic layer, with a hydrophobic layer inside (Figure 1). The Qmean Z-score results on the S and C protein structures with SWISS-MODEL are listed in Table 1. Qmean Z-Score -7.74 on Model 1 predicts a low-quality model (below -4.0). Qmean Z-Score -3.93 on Model 2 predicts a good quality model (above -4.0). Qmean Z-Score on C Protein predicts a good quality model (above -4.0).

Table 3. Qmean Z-Score for protein structure from SWISS-MODEL

Structure	Qmean Z-Score				
	Qmean	Cβ	All atom	Solvation	Torsion
S Protein (Model 1)	-7,74	-7,67	-2,97	-3,03	-5,63
S Protein (Model 2)	-3,93	-0,78	-1,24	-2,58	-3,68
C Protein	-1,06	-2,41	0,77	0,56	-0,87

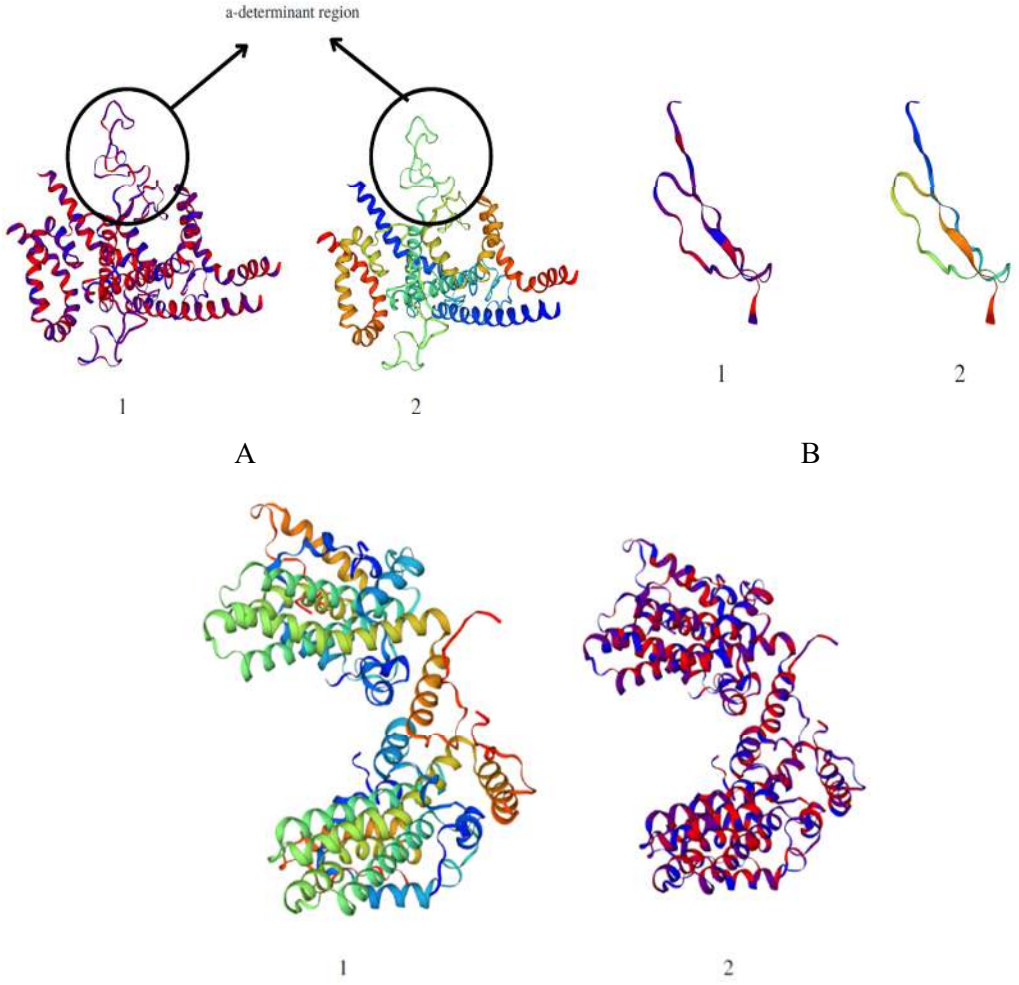

Figure 1. S Protein structure prediction from SWISS-MODEL (A) Model 1 (B) Model 2. Structure (1) color annotation by the hydrophobicity of the protein; the red color on the annotation represents the most hydrophobic region. Blue color depicts the least hydrophobic region. Structure (2) color annotation rainbow respectively from N-terminus to C-terminus. C Protein structure prediction from SWISS-MODEL (C). Structure (1) color annotation by the hydrophobicity of the protein; red color on the annotation represents the most hydrophobic region. Blue color depicts the least hydrophobic region. Structure (2) color annotation rainbow respectively from N-terminus to C-terminus.

3. *Protein structure prediction Phyre2*

Protein structure prediction was done using Phyre2 software, as in Figure 2. S Protein Structure from Phyre2 has a bad quality assessment. Most of the structures were modeled by ab initio. S Protein modeled by Phyre2 does not have any pocket proteins. C Protein Structure from Phyre2 has a good quality assessment (Figure 3). Most of the structures were modeled from updated template. Based on conservation and pocket detection analysis, aa K96 to F110 are the most conserved regions with four protein pocket sites.

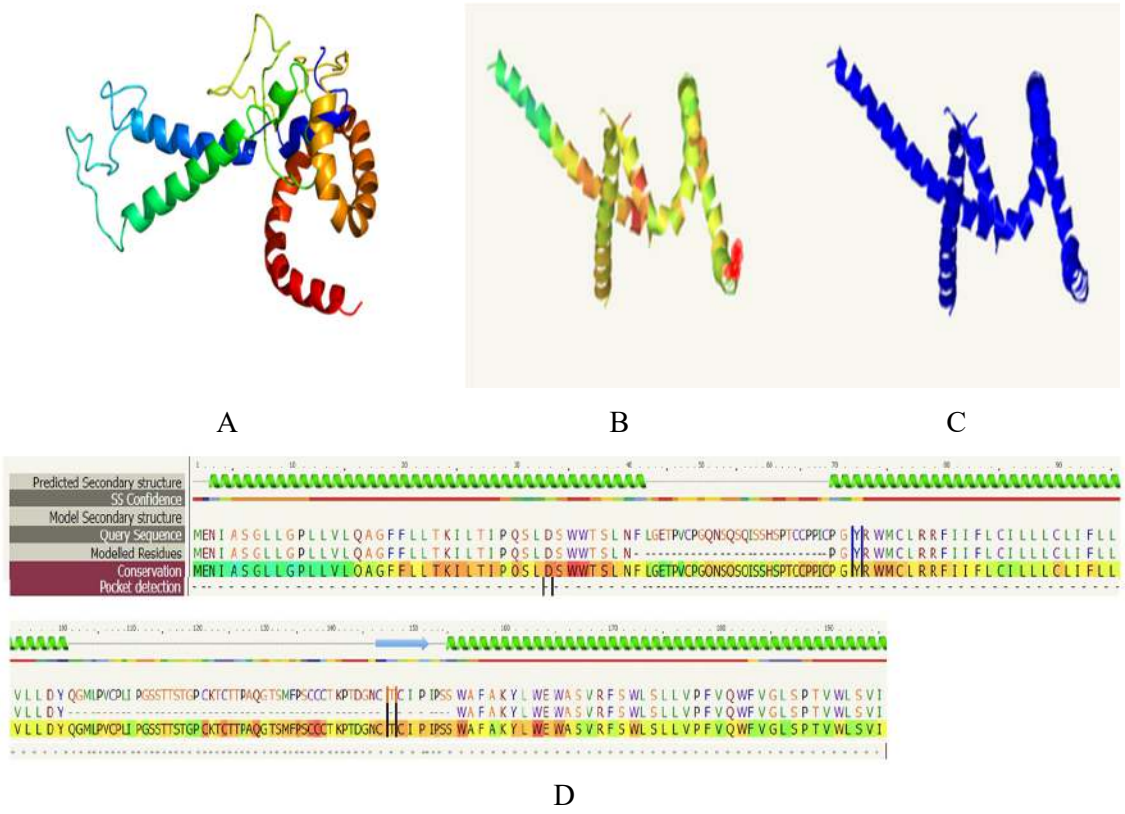

Figure 2. (A)S Protein structure prediction from Phyre2 (B) Conservation site analysis of the predicted model (C) Pocket detection site analysis of the predicted model. (D) Complete query of the S Protein predicted secondary structure along with the confidence level and analysis result of conservation and pocket detection analysis.

4. *Protein ligand binding SWISS-MODEL*

Protein-ligand binding analysis was carried out with SWISS-MODEL on S protein and C protein with a nonbonded interaction graph, as in Figure 4 There is a slight potential to block S Protein based on ligand hotspot analysis from Model 1. Hotspot aa W36 has the highest nonbonded interaction of 9.14% on Model 1 (Figure 4 A). Model 2 has the highest nonbonded interaction of 22.13% on hotspot T125 (Figure 4 B). Nonbonded interaction in Model 2 potentially provides more information regarding ligand blocking of S Protein as an alternative treatment. The C protein hotspot aa W102 has the highest average nonbonded interaction of 6.03% (Figure 4 C). Followed by L37.B of 5.81%. Hotspot aa W102 is also a protein pocket site and the most conserved site based on SWISS-MODEL analysis. There is a potential to bind C Protein using alternative compounds.

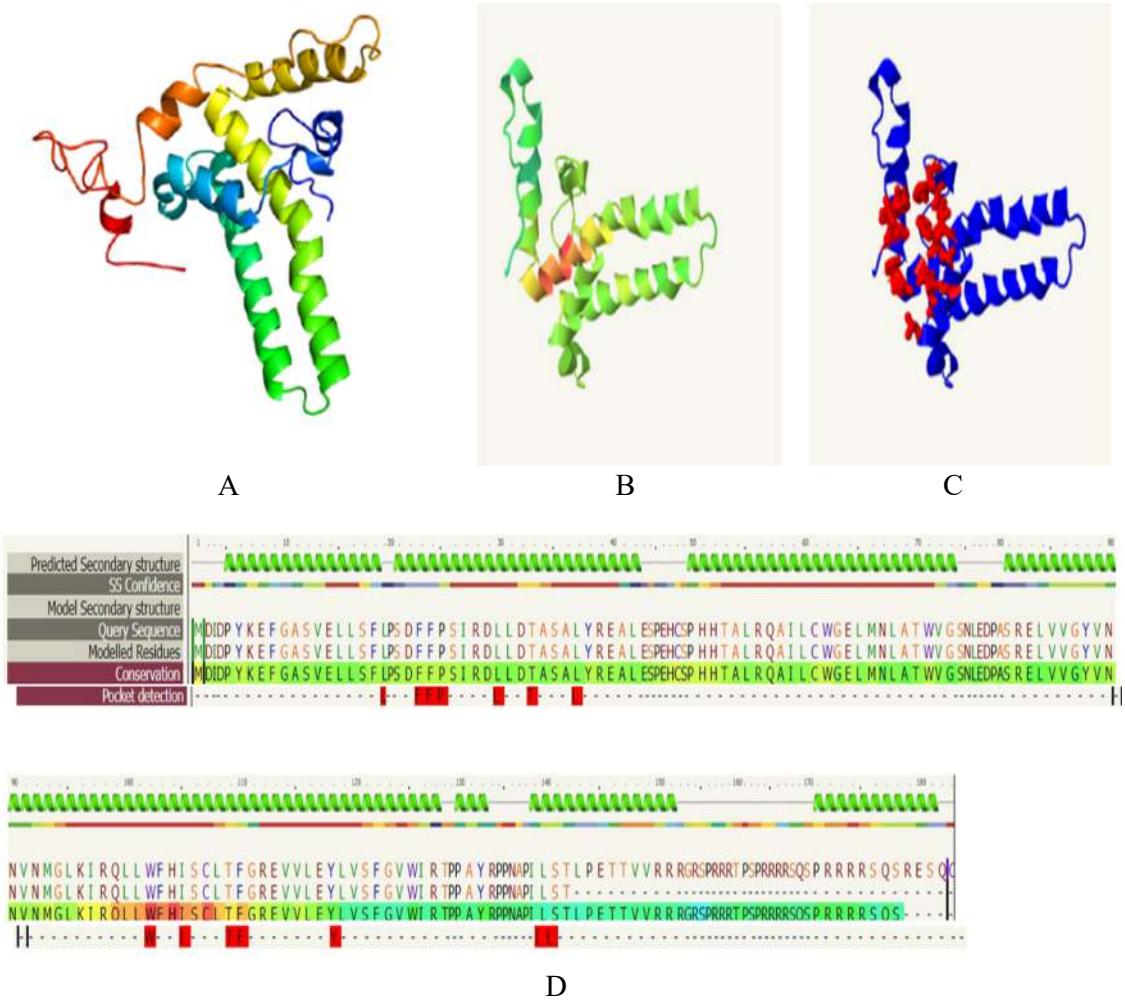

Figure 3. (A) C Protein structure prediction from Phyre2 (B) Conservation site analysis of the predicted model (C) Pocket detection site analysis of the predicted model. (D) Complete query of the C Protein predicted secondary structure along with the confidence level and analysis result of conservation and pocket detection analysis.

Discussion

HBV currently consists of at least 9 genotypes (A to I), with 96% of chronic HBV infections generally caused by genotypes C (26%), D (22%), E (18%), A (17%), and B (14%) [8]. Generally, Indonesia is dominated by genotype B and subgenotype B3, unlike other Asian countries, which are dominated by subgenotypes B1 and B2 [15]. The absence of in-silico research on the B3 subgenotype in Indonesia limits the comparability of our findings with other studies.

Mutations in the B3 subgenotype were analyzed using SWISS-MODEL software. L21S is the main mutation in the S protein (Table1), and P79Q and S87G are in the C protein (Table 2). Mutations in the S protein are associated with several liver disorders[20]. Another study showed that mutations in the HBV genotype may cause liver cirrhosis and could act as molecular markers for the diagnosis of the clinical symptoms of chronic HBV disease [15], [26].

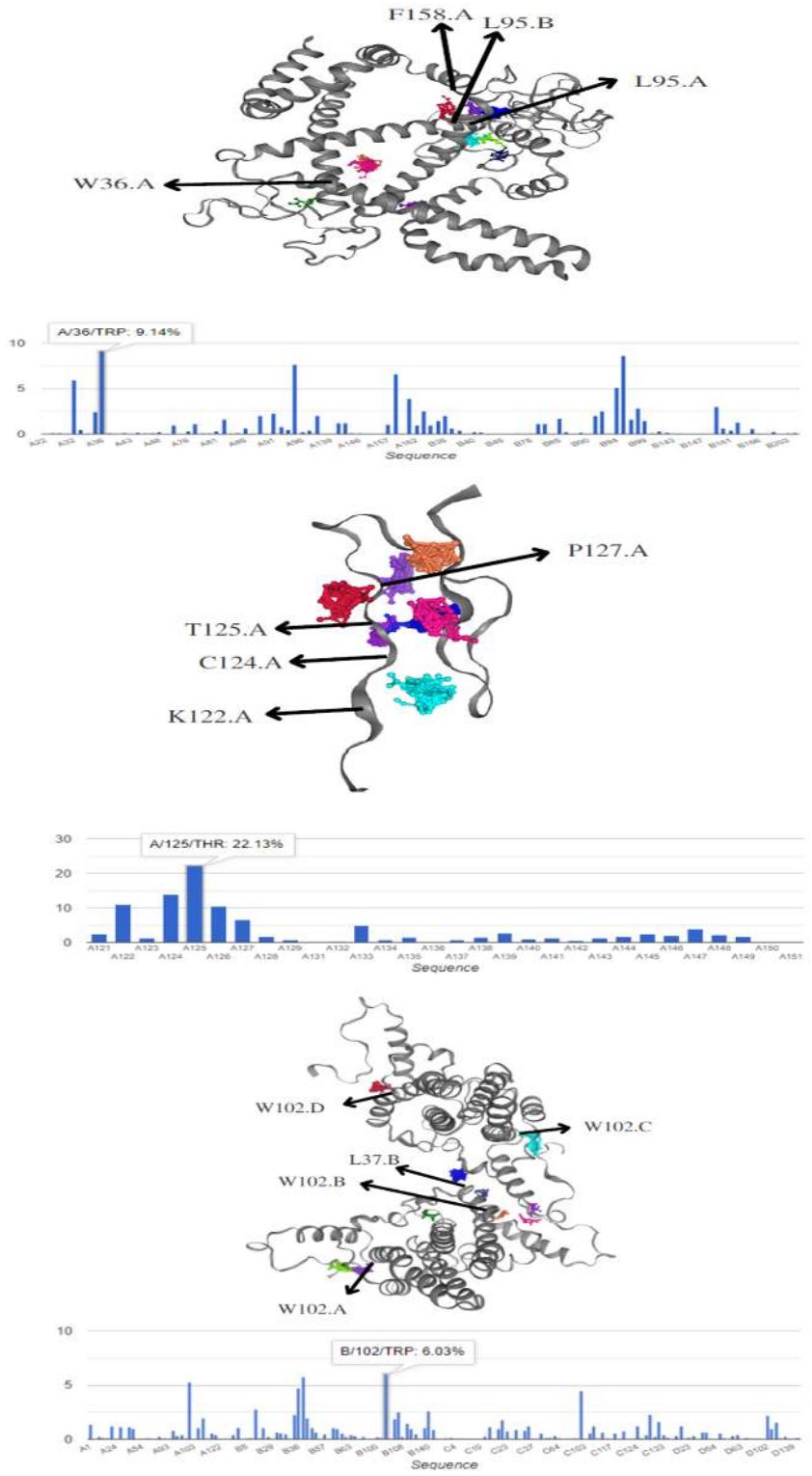

Figure 4. S Protein ligand binding analysis (A) Model 1 S Protein from SWISS-MODEL with nonbonded interaction graph (B) Model 2 S Protein from SWISS-MODEL with nonbonded interaction graph. (C) C Protein ligand binding analysis from C Protein modelled by SWISS-MODEL with nonbonded interaction graph.

SWISS-MODEL and Phyre2 software are used to predict protein structures based on current sequences. SWISS-MODEL provides quality estimates at several stages of the modeling process to help the user identify optimal templates and is also utilized for the fully automated template selection procedure. Once models have been built, their quality is assessed by the QMEAN scoring function (Z-score) [27]. S Protein has two different models (model 1 and model 2), and only model 2 shows good quality (Figure 1, Table 3). Model 1 was analyzed with homodimer Woodchuck hepatitis virus (WHV) and had a Z-score value of -7.14. The woodchuck model system is a vital instrument of natural viral infection. HBV and WHV infections and virions are identical. Because the genomes of HBV and WHV can resemble one other by up to 65%, comparing the two viruses is crucial for the development of antivirals [28]. Model 2 was analyzed with heterodimer tumor necrosis factor receptor and had a Z-score value of -3.93. This shows that model 2 is more effective in developing antivirals compared to model 1. Tumor necrosis factor-alpha-induced protein 1 (TNF-αIP1) was shown to be more highly expressed in HBV [29]. In addition, tumor necrosis factor-alpha (TNF-alpha) can detect HBV infection and inhibit viral DNA replication in mouse animal models [30]. The region highlighted as "a determinant" in model 1 (Figure 1A) is part of the surface gene. It's prone to mutations that can lead to various issues like immune evasion and vaccine resistance. Mutations in the S gene can cause amino acid substitutions, particularly in the HBsAg "a determinant" area. These substitutions can reduce sensitivity in diagnostic tests and lead to failures in response to both the Hepatitis B vaccine (HepB) and Hepatitis B Immunoglobulin (HBIG). These mutations, known as vaccine escape mutations, were reported by [31]. [32] found a higher incidence of these mutations in children who received plasma-derived vaccines (0.3%) compared to those who received recombinant vaccines (0.06%). The C protein has a homo-tetramer oligo state with a Z-score value of -1,06. Mutations in the core protein are known to cause severe liver disease disorders such as liver fibrosis, cirrhosis, and hepatocellular carcinoma [33], [34]. In this study, the C protein consists of the least hydrophobic layer. According to [2], the interaction of mutants L60, L95, and K96 between L protein and C protein occurs in the hydrophobic area preventing the development of mature viruses. Therefore, the C protein model can also be used as a reference for the development of antiviral treatment.

Conservation and pocket detection are functional parameters in Phyre2 protein structure analysis. The conservation model can provide information on the possibility of the presence or absence of a functional residue. Color indicators ranging from green to red indicate residue areas with high conservation value; the closer to red, the higher the conservation value (High). Meanwhile, color indicators from green to purple have a low conservation value; the closer to purple, the lower the conservation value (Low). Pocket detection is one of the functional protein parameters used to predict which amino acids can be used as active sites. The largest pocket is often found as an active site location. The largest pocket detected by the fpocket2 program [35], is shown in red wireframe mode. Pocket detection was not detected in the S Protein (Figure 2C) but was detected in the C protein with four active pocket sites (Figure 3C). The hydrophobic pocket of an external component can be bound by the HBV capsid. According to [36], the homolog of Triton X-100 is predicted to disrupt the HBV life cycle by either engaging in competition with the natural pocket factor or by impeding capsid dynamics into a single conformation. A novel target for medication to intervene in the HBV life cycle is the hydrophobic pocket. This allows the C protein to bind to the HBV capsid and inhibit the process of

transferring genetic material and HBV replication. The results suggest that the S Protein model requires additional templates to improve its quality, while the C protein model exhibits better quality and can serve as a reference for future research.

Research on the interactions between proteins and ligands is crucial to knowing the mechanism of biological regulation. This technique can identify potentially active compounds with the greatest for developing drugs and forecast the binding affinity of molecules inside certain receptor targets [37]. In this study, hotspot aa W36 has the highest nonbonded interaction of 9.14% on Model 1 of S protein (Figure 4A), whereas model 2 has the highest nonbonded interaction of 22.13% on hotspot T125 (Figure 4B). The C protein hotspot aa W102 has the highest average nonbonded interaction of 6.03% (Figure 4C), followed by L37.B of 5.81%. Hotspots were often linked to protein areas that bind low molecular weight molecules, a single ligand that has the same moiety as a substructure, and a single binding subpocket across various structures [38]. Protein C shows a lower affinity value between model 1 and model 2 of S protein. Lower values signify a stronger hydrogen bond between the drug and the protein receptor as well as a greater binding affinity [39], [40]. In addition, hotspot aa W102 of C protein is also a protein pocket site and the most conserved site based on SWISS-MODEL analysis. There is a potential to bind C Protein using alternative compounds.

Conclusions

Mutations in the B3 subgenotype consist of S protein mutations (L21S) and C protein mutations (P79Q and S87G) which cause chronic HBV in Indonesia. The model for S Protein from homology structure prediction can be said to be reliable thus it still needs more templates from experimental techniques. While C Protein structure prediction can provide information for further research in alternative natural antiviral treatment supported by a Z-score value above -4.0, has four active pocket site, and has the highest binding affinity capability.

References

[1] M. Guvenir and A. Arikan, "Hepatitis B virus: from diagnosis to treatment," *Pol. J. Microbiol.*, vol. 69, no. 4, pp. 391–399, 2020, doi: 10.33073%2Fpjm-2020-044.

[2] F. Pastor *et al.*, "Direct interaction between the hepatitis B virus core and envelope proteins analyzed in a cellular context," *Sci. Rep.*, vol. 9, no. 1, p. 16178, 2019, doi: 10.1038/s41598-019-52824-z.

[3] W. A. Putri *et al.*, "Association between quasispecies variants of hepatitis B virus, as detected by high-throughput sequencing, and progression of advanced liver disease in Indonesian patients," *Mol. Med. Rep.*, vol. 20, no. 1, pp. 16–24, 2019, doi: 10.3892/mmr.2019.10250.

[4] M. H. Nguyen, G. Wong, E. Gane, J.-H. Kao, and G. Dusheiko, "Hepatitis B virus: advances in prevention, diagnosis, and therapy," *Clin. Microbiol. Rev.*, vol. 33, no. 2, pp. 10–1128, 2020 doi: 10.1128/cmr.00046-19.

[5] M. D. Thedja, D. H. Muljono, N. Nurainy, C. H. C. Sukowati, J. Verhoef, and S. Marzuki, "Ethnogeographical structure of hepatitis B virus genotype distribution in Indonesia and discovery of a new subgenotype, B9," *Arch. Virol.*, vol. 156, pp. 855–868, 2011, doi: 10.1007/s00705-011-0926-y.

[6] D. H. Muljono, "Epidemiology of hepatitis B and C in Republic of Indonesia," *Euroasian J. Hepato-Gastroenterol.*, vol. 7, no. 1, p. 55, 2017, doi: 10.5005%2Fjp-journals-l0018-1212.

[7] D. Razavi-Shearer *et al.*, "Global prevalence, treatment, and prevention of hepatitis B virus infection in 2016:

a modelling study," *Lancet Gastroenterol. Hepatol.*, vol. 3, no. 6, pp. 383–403, 2018, doi: 10.1016/s2468-1253(18)30056-6.

[8] Y. Yano, T. Utsumi, M. I. Lusida, and Y. Hayashi, "Hepatitis B virus infection in Indonesia," *World J. Gastroenterol. WJG*, vol. 21, no. 38, p. 10714, 2015, doi: 10.3748/wjg.v21.i38.10714.

[9] F. J. van Hemert, H. L. Zaaijer, B. Berkhout, and V. V Lukashov, "Mosaic amino acid conservation in 3D-structures of surface protein and polymerase of hepatitis B virus," *Virology*, vol. 370, no. 2, pp. 362–372, 2008, doi: 10.1016/j.virol.2007.08.036.

[10] D. Schuppan and N. H. Afdhal, "Liver cirrhosis," *The Lancet*, vol. 371, no. 9615, pp. 838–851, 2008, doi: 10.1016/s0140-6736(08)60383-9.

[11] L. N. Yamani *et al.*, "Ultradeep sequencing for detection of quasispecies variants in the major hydrophilic region of hepatitis B virus in Indonesian patients," *J. Clin. Microbiol.*, vol. 53, no. 10, pp. 3165–3175, 2015, doi: 10.1128/jcm.00602-15.

[12] H. Kim, S.-A. Lee, and B.-J. Kim, "X region mutations of hepatitis B virus related to clinical severity," *World J. Gastroenterol.*, vol. 22, no. 24, p. 5467, 2016, doi: 10.3748/wjg.v22.i24.5467.

[13] M. Campos-Valdez, H. C. Monroy-Ramírez, J. Armendáriz-Borunda, and L. V Sánchez-Orozco, "Molecular mechanisms during hepatitis B infection and the effects of the virus variability," *Viruses*, vol. 13, no. 6, p. 1167, 2021, doi: 10.3390%2Fv13061167.

[14] Z.-H. Zhang, C.-C. Wu, X.-W. Chen, X. Li, J. Li, and M.-J. Lu, "Genetic variation of hepatitis B virus and its significance for pathogenesis," *World J. Gastroenterol.*, vol. 22, no. 1, p. 126, 2016, doi: 10.3748/wjg.v22.i1.126.

[15] R. Kumar, "Review on hepatitis B virus precore/core promoter mutations and their correlation with genotypes and liver disease severity," *World J. Hepatol.*, vol. 14, no. 4, p. 708, 2022, doi: 10.4254/wjh.v14.i4.708.

[16] H. Liu, S. Chen, X. Liu, and J. Lou, "Effect of S-region mutations on HBsAg in HBsAg-negative HBV-infected patients," *Virol. J.*, vol. 21, no. 1, p. 92, 2024, doi: 10.1186/s12985-024-02366-2.

[17] A. Kryshtafovych and K. Fidelis, "Protein structure prediction and model quality assessment," *Drug Discov. Today*, vol. 14, no. 7–8, pp. 386–393, 2009, doi: 10.1016%2Fj.drudis.2008.11.010.

[18] M. Basyuni *et al.*, "Protein modelling of triterpene synthase genes from mangrove plants using Phyre2 and Swiss-model," in *Journal of Physics: Conference Series*, IOP Publishing. p. 12095, 2018, doi: 10.1088/1742-6596/978/1/012095.

[19] Y. Qin *et al.*, "Improved method for rapid and efficient determination of genome replication and protein expression of clinical hepatitis B virus isolates," *J. Clin. Microbiol.*, vol. 49, no. 4, pp. 1226–1233, 2011.

[20] E. Gasteiger *et al.*, *Protein identification and analysis tools on the ExPASy server*. Springer, 2005, doi: 10.1385/1-59259-890-0:571.

[21] A. Bateman *et al.*, "The Pfam protein families database," *Nucleic Acids Res.*, vol. 32, no. suppl_1, pp. D138–D141, 2004, doi: 10.1093/nar/gkh121.

[22] M. Biasini *et al.*, "SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information," *Nucleic Acids Res.*, vol. 42, no. W1, pp. W252–W258, 2014, doi: 10.1093/nar/gku340.

[23] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, and M. J. E. Sternberg, "The Phyre2 web portal for protein

modeling, prediction and analysis," *Nat. Protoc.*, vol. 10, no. 6, pp. 845–858, 2015, doi: 10.1038/nprot.2015.053.

[24] M. Abavisani *et al.*, "Mutations in SARS-CoV-2 structural proteins: a global analysis," *Virol. J.*, vol. 19, no. 1, p. 220, Dec. 2022, doi: 10.1186/s12985-022-01951-7.

[25] Md. F. Ahmed, Md. H. Kabir, and A. Z. Md. T. Islam, "Impact of Feed Point Position on Patch Antenna's Return Loss and Bandwidth for UWB Applications," *J. Multidiscip. Appl. Nat. Sci.*, vol. 4, no. 1, pp. 30–38, Jun. 2023, doi: 10.47352/jmans.2774-3047.158.

[26] C.-H. Chen *et al.*, "Clinical significance of hepatitis B virus (HBV) genotypes and precore and core promoter mutations affecting HBV e antigen expression in Taiwan," *J. Clin. Microbiol.*, vol. 43, no. 12, pp. 6000–6006, 2005, doi: 10.1128/jcm.43.12.6000-6006.2005.

[27] A. Waterhouse *et al.*, "SWISS-MODEL: homology modelling of protein structures and complexes," *Nucleic Acids Res.*, vol. 46, no. W1, pp. W296–W303, 2018, doi: 10.1093/nar/gky427.

[28] A. A. Kukreja *et al.*, "Structurally similar woodchuck and human hepadnavirus core proteins have distinctly different temperature dependences of assembly," *J. Virol.*, vol. 88, no. 24, pp. 14105–14115, 2014.

[29] M. C. Lin *et al.*, "Tumor necrosis factor-α-induced protein 1 and immunity to hepatitis B virus," *World J. Gastroenterol.*, vol. 11, no. 48, p. 7564, 2005, doi: 10.1128/JVI.01840-14.

[30] H.-T. Tzeng *et al.*, "Tumor necrosis factor-alpha induced by hepatitis B virus core mediating the immune response for hepatitis B viral clearance in mice model," *PLoS One*, vol. 9, no. 7, p. e103008, 2014, doi: 1007-9327/11/7564.asp.

[31] K. Ko *et al.*, "Existence of hepatitis B virus surface protein mutations and other variants: demand for hepatitis B infection control in Cambodia," *BMC Infect. Dis.*, vol. 20, pp. 1–13, 2020, doi: 10.1186/s12879-020-05025-3.

[32] H.-Y. Hsu *et al.*, "No increase in prevalence of hepatitis B surface antigen mutant in a population of children and adolescents who were fully covered by universal infant immunization," *J. Infect. Dis.*, vol. 201, no. 8, pp. 1192–1200, 2010, doi: 10.1086/651378.

[33] A. Mohamadkhani *et al.*, "The role of mutations in core protein of hepatitis B virus in liver fibrosis," *Virol. J.*, vol. 6, pp. 1–9, 2009, doi: 10.1186/1743-422X-6-209.

[34] A. A. Al-Qahtani *et al.*, "The correlation between hepatitis B virus precore/core mutations and the progression of severe liver disease," *Front. Cell. Infect. Microbiol.*, vol. 8, p. 355, 2018, doi: 10.3389%2Ffcimb.2018.00355.

[35] V. Le Guilloux, P. Schmidtke, and P. Tuffery, "Fpocket: an open source platform for ligand pocket detection," *BMC Bioinformatics*, vol. 10, pp. 1–11, 2009, doi: 10.1186/1471-2105-10-168

[36] L. Lecoq *et al.*, "A pocket-factor–triggered conformational switch in the hepatitis B virus capsid," *Proc. Natl. Acad. Sci.*, vol. 118, no. 17, p. e2022464118, 2021, doi: 10.1073/pnas.2022464118.

[37] Y. Fu, J. Zhao, and Z. Chen, "Insights into the molecular mechanisms of protein-ligand interactions by molecular docking and molecular dynamics simulation: A case of oligopeptide binding protein," *Comput. Math. Methods Med.*, vol. 2018, 2018, doi: 10.1155/2018/3502514.

[38] A. E. Wakefield *et al.*, "Benchmark sets for binding hot spot identification

in fragment-based ligand discovery," *J. Chem. Inf. Model.*, vol. 60, no. 12, pp. 6612–6623, 2020, doi: 10.1021/acs.jmedchem.1c01803.

[39] M. Uzzaman, M. K. Chowdhury, and M. Belal Hossen, "Thermochemical, molecular docking and ADMET studies of aspirin metabolites," *Front Drug Chem Clin Res*, vol. 2, pp. 1–5, 2019, doi: 10.1016/j.imu.2021.100706.

[40] M. A. Thafar, M. Alshahrani, S. Albaradei, T. Gojobori, M. Essack, and X. Gao, "Affinity2Vec: drug-target binding affinity prediction through representation learning, graph mining, and machine learning," *Sci. Rep.*, vol. 12, no. 1, p. 4751, 2022, doi: 10.1038/s41598-022-08787-9.