

In Silico* Anaysis of Gene Sequence Variation and Molecular Characterization Segregation and Condensation Protein B (ScpB) *Streptococcus agalactiae

Fransisca Probo Setyoningrum^{1,2}, Ari Yuniastuti¹, Irwan Budiono¹ and Sri Ratna Rahayu¹

¹Department of Public Health, Faculty of Medicine, Universitas Negeri Semarang, Central Java, Indonesia

²Department of Medical Laboratory Technology, Faculty of Health Science, Universitas Pelita Harapan, Tangerang, Banten

Abstract. *Streptococcus agalactiae* is a causes sepsis in mothers and neonates. Cases of infection in the United States are estimated at 8000 cases of GBS infection early – onset occurs annually in cases per 1000 live births. Approximately 4-6% of infected babies die and 25-50% survive will experience sequelae such as mental retardation, blindness and deafness. Gold standard for diagnosing sepsis is the culture method, but this method has limitations in the examination, including requiring a long time, the culture results which are declared negative after being confirmed by the PCR method get positive results. The result of culture examination was not infected and when the baby was born get suffered from sepsis. The ScpB gene plays a role in immunogenic factors. Therefore, this study aims to molecularly characterizing variants in the ScpB gene of *Streptococcus agalactiae* back bacteria. Genetic variations in a population will influence traits and responses. This research aims to analyze genetic variations in the ScpB sequence *Streptococcus agalactiae* NCBI with PopSet 2572585596 using BsmAI restriction enzyme with RFLP method *in silico*. The results showed that there were genetic variations in the ScpB gene sequence and three allelic variations contained in 10 sequences. ScpB is also responsible for the response to antibiotics. Some antibiotics are resistance, even the antibiotic recommended by. FDA omandacycline, is also caused by. the biofilm which product from ScpB. In conclusion, there are 3 allelic variations result from *in silico* RFLP, these variations will certainly have varying effect including immune responses, biofilm production and resistance to antibiotics.
Keywords: *Streptococcus agalactiae*, ScpB gene, *in silico* RFLP, genetic variances, antibiotic resistance

1 Introduction

Streptococcus agalactiae is one of the bacteria that causes sepsis in pregnant women and neonatal [1]. Infection cases *Streptococcus agalactiae* in the United States it is estimated that there are 8000 cases of GBS infection early-onset occurs annually in cases per 1000 live births. Approximately 4-6% of infected babies die and 25-50% survive will experience sequelae such as mental retardation, blindness and deafness. In pregnant women, approximately 5 - 30 % have asymptomatic colonization, GBS in the genital and gastrointestinal tracts will receive the same colonization through vertical transmission either in utero or when it passes through the birth canal [2]. Therefore, appropriate and fast diagnostic tests are needed so that sepsis can be prevented and detected early [3]. This of course has an impact on improving the quality of life of pregnant women and neonates. The gold standard issued by the CDC to diagnose infection *Streptococcus agalactiae* is culture, but this method has several disadvantages, including the long examination time and often giving false negative results. This is proven by the negative results of the mother's culture examination but the neonate is born with sepsis caused by *Streptococcus agalactiae* [1, 4]. Therefore, research regarding the effectiveness of culture methods is often carried out and compared with other methods such as Polymerase Chain Reaction (PCR) which gives good results because it has a specificity of 95.9% [5]. Christie Atkinson Munch Peterson (CAMP) is a culture method used in the test, lacking sensitivity for identification *Streptococcus agalactiae* because the culture results were negative when

confirmed using the Phoenix system, PCR and 16s rRNA gene sequencing gave positive results for *Streptococcus agalactiae* [6]. This is caused by several possibilities, such as the patient consuming prophylaxis antibiotics so that the culture method is less effective than the molecular method [7].

The ScpB gene is a gene that is responsible for activation of the complement immune system and invasion of epithelial cells. Genetic variations are changes that occur in nucleotides, genes, chromosomes and genomes in a microorganism. Genetic variations in ScpB influence the ability to interrupt complement activation through cleavage of the neutrophil chemottractant C5a and invasion of human cells [8]. This variation is able to influence the level of virulence of *Streptococcus agalactiae* so there are times when pregnant women do not show signs and symptoms of infection, but when the neonate is born through the birth canal, it will quickly become infected and then develop sepsis. One way to detect genetic variations in the ScpB gene is to use molecular markers Restriction Fragment Length Polymorphism (RFLP).

RFLP uses retraction enzymes to identify DNA sequences. RFLP detection is based on the possibility of comparing the profiles of bands produced after restriction enzyme cutting of DNA to produce fragments of different lengths. RFLP is widely used because has a simpler and more sensitive way of using specific markers to analyze the similarity and variability of genes [9]. RFLP can be visualized in silico for outcome prediction genotyping before carrying out RFLP in the laboratory. Utility in silico is to predict success and provide hypotheses, provide new discoveries; this computer simulation analysis provides results that are close to real conditions. Literature studies are used to obtain information regarding the structure of nucleotide base sequences. The nucleotides were searched for PopSet genes from the NCBI genebank URL <http://www.ncbi.nlm.nih.gov> by searching for the PopSet number. In PopSet: 2572585596 there are ScpB genes from 10 isolates that have been submitted. This gene will be subjected to in silico RFLP to determine its genetic variations. Genes variation can be supported by analyzed molecular characterization using several tools bioinformatics. The proteins can be accessed from accession number of PopSet number. Characters that can be obtained include analysis amino acid composition, physico-chemical properties, primary and secondary structure. Asam amino composition and physico-chemical use Protparam, ProtScale and SOSUI then primary and secondary structure use PSIPRED, SOPMA and SWISS-MODEL.

2 Methods

2.1 Data Set

Literature studies were used to obtain information about the ScpB gene with PopSet identity number: 2572585596 which was downloaded in fasta format from NCBI *Streptococcus agalactiae* C5a peptidase (scpB) gene, partial cds. PopSet - NCBI (nih.gov). The PopSet has 10 gene sequences with accession numbers **Table 1**.

Table 1. List of Accession Numbers for *Streptococcus agalactiae* ScpB Gene Isolates

No	Accession Number Isolot Gen ScpB
1	OR174997.1
2	OR174996.1
3	OR174995.1
4	OR174994.1
5	OR174993.1
6	OR174992.1
7	OR174991.1
8	OR174990.1
9	OR174989.1
10	OR174988.1

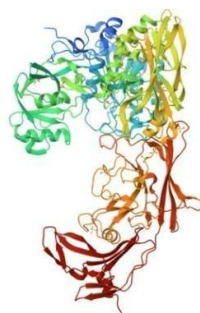


Fig. 1. 3D structure of the ScpB protein *Streptococcus agalactiae* visualized with RCBS Protein Data Base (PDB).

2.2 Data Collection

The next step is to look for restriction enzyme candidates that will be used <http://insilico.ehu.es/restriction/>. In this method the tools chosen are compare restriction patterns of many sequences then enter fasta which comes from NCBI. Further more Go to next step, the final step is choosing Get list of restriction enzymes so that enzymes are obtained the best restrictions that will be used in the next stage [10]. RFLP is carried out automatically with *in silico* with <http://www.benchling.com/>. Import all the downloaded isolates from PopSet, then click the scissor-shaped button, then click find enzyme Next, select the names of the restriction enzymes that were previously determined during screening, type them in the column. Then click menu run digest to restrict. Results can be seen in the menu virtual digest An electrophorogram will appear in the form of image [11]. Molecular characterization was also carried out to obtain the amino acid composition, physicochemical properties, primary structure and secondary structure using bioinformatics as a further analysis of the alleles formed. The aim is to see the character of the allelic variations that are formed. It requires fasta proteins that can be accessed from accession numbers that can be accessed from the NCBI PopSet. Fasta protein is the main data that will be used to analyze its amino acid composition, physicochemical properties, primary structure and secondary structure. Amino acid composition and physicochemical properties were obtained by accessing the ProtParam tool <http://web.expasy.org/protparam/>. Then to see the primary structure it is obtained by accessing <https://web.expasy.org/protscale/> strengthened by accessing <https://harrier.nagahama-i-bio.ac.jp/sosui/mobile/>. Secondary structure predictions were obtained PSIPRED Workbench (ucl.ac.uk) and to get a 3D image of the protein can be accessed via SWISS-MODEL (expasy.org).

2.3 Data Processing

The research results in the form of a description that explains the variations that will be obtained

3 Results And Discussion

3.1 Structure

The ScpB gene produces C5a peptidase which is responsible for activation of the complement immune system and invasion of epithelial cells. C5a peptidase is able to damage chemotactic signals by breaking down complement C5a to make it inactive [12]. Location of C5a peptidase.

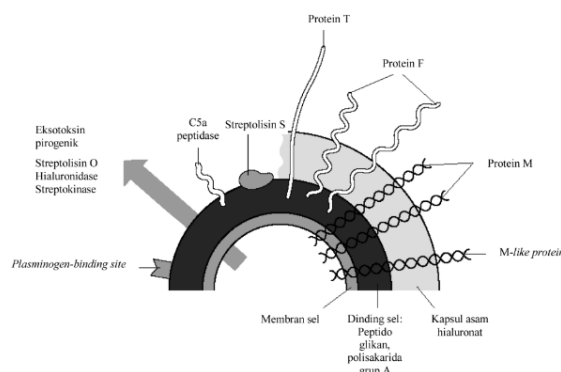


Fig. 2. Diagram of the structure of *Streptococcus agalactiae*.

Restriction enzyme candidate screening, obtained and selected the BsmAI enzyme on the side introduction of 5'-GTCTCN-3'. This restriction enzyme was chosen because there are cutting variations in the ScpB gene sequence. Restriction Endonuclease Enzyme is an enzyme that is able to cut internal parts of DNA that works specifically and is able to cut precisely at phosphodiester bonds. The ideal restriction enzyme is one that is capable of cutting at specific locations and is able to recognize 4 - 8 nucleotide sequences. Restriction enzyme cutting will produce cuts, namely cohesive ends and flat ends or often referred to as sticky end and blunt end. The optimum temperature required for the BsmAI enzyme to make cuts is 37°C. The BsmAI enzyme comes from bacteria *Bacillus stearothermophilus* [13].

RFLP is a technique commonly used for genotyping through cutting DNA sequences with the help of restriction enzymes [14]. The RFLP method is an analysis method using restriction enzymes that cut typical nucleotide sequences at certain different locations resulting in fragments of different lengths [15]. RFLP analysis is a codominant marker used to achieve various goals. Considering that restriction sites have certain DNA sequences, this means that variations in the presence of restriction sites reflect variations in DNA sequence. In other words, RFLP can function as an estimate of DNA variation. Variations are detected in the form of polymorphic (double) long sequence cuts where the assessment time of the variation sequence is possible from the fragment data itself, long sequence variations in a section can be assessed from nucleotide substitutions [16]. Studies in silico is an initial study before proceeding with other methods such as in vivo and in vitro to help predict and provide hypotheses about the activity of a compound or ligand because the process of both is sometimes difficult to explain in a simple way the mechanism of the ligand and target and requires more time and expensive costs cheap [17].

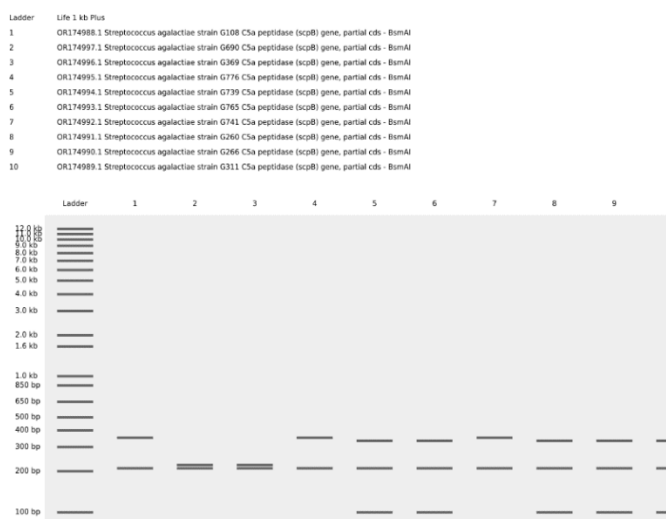


Fig 3. Electrophorogram of the results of restriction of the ScpB gene sequence using enzymes BSMAl *In silico*.

RFLP analysis on the 10 *ScpB* gene sequences *Streptococcus agalactiae* using the BsmAI restriction enzyme which can be seen in **(Figure 3)**. Produces three allele variations, the allele 1 produces two bands with sizes 352bp and 563bp in isolates 1, 4, 7, the second allele produces two bands with sizes 223bp and 211bp in isolates 2,3, the third allele produces three bands with sizes 335bp, 211bp and 17bp in isolates 5,6,8,9,10. Allele frequencies in the *ScpB* gene *Streptococcus agalactiae* with the BsmAI restriction enzyme can be seen in **Table 2**.

Table 2. Allele frequencies of the *ScpB* gene *Streptococcus agalactiae* RFLP results By *In silico*

Enzyme Restriction	Site Introduction	Size Fragment (bp)	Allele	Total Presence Fragment (N=10)	Percentage Presence Fragment (%)	Frequency Allele
BsmAI	GTCTCN	352bp	1	3	30%	0.30
		563bp				
		223bp	2	2	20%	0.20
		211bp				
		335bp	3	5	50%	0.50
		211bp 17bp				

The gene variations found in these 3 alleles indicate that the isolate *Streptococcus agalactiae* have variations of the *ScpB* gene. The *ScpB* gene is a gene that influences the ability of bacteria to inactivate neutrophil chemotactiles towards the center of infection, the ability to fight the host defense system and host extracellular matrix proteins that influence colonization, persistence, translocation of biofilm production, and invasion of the central nervous system. The ability to cause virulence is proven by two things that are being widely researched to this day:

1. Biofilm Formation

Biofilm is a collection of bacterial cells that form colonies covered by an adhesive composed of carbohydrates secreted by bacteria [18].

2. Resistance to antibiotics

The inability of antibiotics to eliminate bacteria so that it is concluded that they are resistant to certain antibiotics include **Table 3**.

Table 3. Resistant Antibiotics *Streptococcus agalactiae*

Antibiotics	Source
Penicillin, linezolid, vancomycin	Miranda PSD et al, 2018
Erythromycin, clindamycin, vancomycin	Hayes et al, 2020
Omadacycline, linezolid	Abrahamian F et al, 2019
Omadacycline	Li G et al, 2022

OMC (Omadacycline) antibiotics, antibiotics that have received approval from the FDA, are used for the treatment of skin and soft tissue infections, as well as pneumonia, and are also less effective in eliminating them due to their ability to *Streptococcus agalactiae* in biofilm formation [22]. The impact of these allele variations not only affects the properties of *Streptococcus agalactiae* but also alters the amino acid composition, physicochemical properties, primary structure, and secondary structure. The following are the results of molecular characterization of the three alleles. Allele 1 isolates numbers 1, 4, 7 (OR174997.1, OR174994.1, OR174991.1). Allele 2 isolates numbers 2, 3 (OR174996.1, OR174995.1). Allele 3 isolates numbers 5, 6, 8, 9, 10 (OR174993.1, OR174992.1, OR174990.1, OR174989.1, OR174988.1). From each allele, one representative isolate was randomly selected for further bioinformatic analysis. The molecular characterization of the selected proteins, namely allele 1 (OR174997.1), allele 2 (OR174995.1) and allele 3 (OR174990.1) will be observed:

1. Amino Acid Composition

The amino acid composition of each representative allele was analyzed using ProtParam as shown in **Table 4**.

Table 4. Amino Acid Composition (%) SpcB *Streptococcus agalactiae*

No	Amino Acid Composition	Result %		
		Allele 1 (OR174997.1) Protein ID: WMQ58569.1	Allele 2 (OR174995.1) Protein ID: WMQ58567.1	Allele 3 (OR174990.1) Protein ID: WMQ58562.1
1	Ala (A)	11.7%	11.2%	10.6%
2	Arg (R)	1.4%	2.7%	3.2%
3	Asn (N)	4.8%	5.3%	5.9%
4	Asp (D)	8.3%	9.6%	10.1%
5	Gln (Q)	4.8%	3.7%	3.2%
6	Glu (E)	2.1%	3.2%	3.7%
7	Gly (G)	6.2%	6.4%	6.4%
8	His (H)	0.0%	0.0%	0.5%
9	Ile (I)	5.5%	5.9%	5.9%
10	Leu (L)	8.3%	6.9%	6.9%
11	Lys (K)	11.0%	11.7%	10.1%
12	Met (M)	2.1%	1.6%	1.6%
13	Phe (F)	2.8%	3.7%	3.7%
14	Pro (P)	6.9%	5.9%	5.9%
15	Ser (S)	8.3%	6.9%	6.9%
16	Thr (T)	7.6%	6.9%	6.9%
17	Trp (W)	0.7%	0.5%	0.5%
18	Tyr (Y)	2.1%	3.2%	3.2%
19	Val (V)	5.5%	4.8%	4.8%

SpcB in *Streptococcus agalactiae* is composed of 19 amino acids. The three alleles show different amounts of amino acid composition. In the first allele the most dominant amino acid is Alanine (11.7%) and the lowest is Tryptophan (0.7%), in the second allele the most dominant amino acid is Lysine (11.7%) and the lowest is Tryptophan (0.5%), and in The most dominant allele for the three amino acids is Alanine (10.6%) then the lowest are Tryptophan (0.5%) and Histidine (0.5%).

2. Primary Structure and Physico-Chemical Properties of SpcB

Primary structure and physicochemical properties were analyzed using the ProtParam program on the ExPASy website <https://web.expasy.org/protparam/> and ProtScale on the ExPASy site <https://web.expasy.org/protscale/>. Primary structure and physicochemical properties are interconnected, this is to determine the efficiency and stability of proteins in biological systems. The results of the physicochemical properties of SpcB from the ProtParam program are shown in Table 5.

Table 5. Physicochemical profile of SpcB on the three allele variations

Parameter	Hasil		
	Allele 1 (OR174997.1) Protein ID: WMQ58569.1	Allele 2 (OR174995.1) Protein ID: WMQ58567.1	Allele 3 (OR174990.1) Protein ID: WMQ58562.1
Number of Amino Acid	145	188	188
Molecular Weight	15409.57kDa	20392.07kDa	20459.97kDa
Isoelectric Point	9.08	8.89	5.91
Value (pI)			
Atomic Composition	C=682, H=1108, N=182, O=216, S=3	C=904, H=1447, N=243, O=286, S=3	C=903, H=1435, N=245, O=290, S=3
Formulas	C628H1108O216S3	C904H1447N243O286S3	C903H1435N245O290S3
Number of Atoms	2191	2883	2876
Estimated Half Time	100 hours (mammalian reticulocytes, in vitro >20 hours (yeast, in vivo) >10 haours (<i>Escherichia coli</i> , in vivo)	100 hours (mammalian reticulocytes, in vitro >20 hours (yeast, in vivo) >10 haours (<i>Escherichia coli</i> , in vivo)	100 hours (mammalian reticulocytes, in vitro >20 hours (yeast, in vivo) >10 haours (<i>Escherichia coli</i> , in vivo)
Instability Index	43.31 (Protein unstable)	39.05 (Protein Stabil)	42.13 (Protein unstable)
Aliphatic Index	81.52	74.84	74.31
Grand Average of Hydrophaticity	-0.358	-0.546	-0.571
Total number of negative residues (Asp+Glu)	15	24	26
Total Number of Positive Residues (Arg+Lys)	18	27	25

The analysis results in **(Table 5)** show the number of amino acids and molecular weights vary. The resulting isoelectric point values are also different for alleles 1 and 2, indicating that the protein is basic, while for allele 3, it indicates that the protein is basic. Proteins are amphoteric compounds that can react with acids and bases. This is because the molecules have positive and negative charges. The isoelectric point (pI) is the pH at which an amino acid is neutral or has a charge of 0 (zwitterion form). When the isoelectric point is reached, the number of positive and negative charges is the same. If the pH is above the isoelectric point, the protein will have a negative charge. Conversely, if the pH is below the isoelectric point, the protein will have a positive charge [23,24].

The SpcB protein has a different stability index on alleles 1 and 3, while the protein is unstable on allele 2. A protein is considered stable if it has an index value <40 and a protein with an index value >40 is said to be unstable [25]. This index predicts protein stability based on its amino acid composition [26].

The protein aliphatic index for the three alleles shows that this protein is considered stable over a wide temperature range (thermostable). Aliphatic amino acids (A, I, L, and V) are responsible for the thermal stability of proteins. This index evaluates protein thermostability based on the percentage of each highly aliphatic amino acid that has thermally stable properties. The higher the index value, the more thermostable the protein is [27]. The aliphatic index plays a role in thermal stability and the amount. The GRAVY protein index for all three alleles shows negative values, where this protein is hydrophilic and will easily interact with water molecules [27]. Proteins with a negative GRAVY score are considered hydrophilic in nature with good solubility [25]. Further searches use the Protscale program on the ExPASy site <https://web.expasy.org/protscale/> to determine the level of hydrophobicity of the SpcB protein in the three alleles. **Figure 4.**

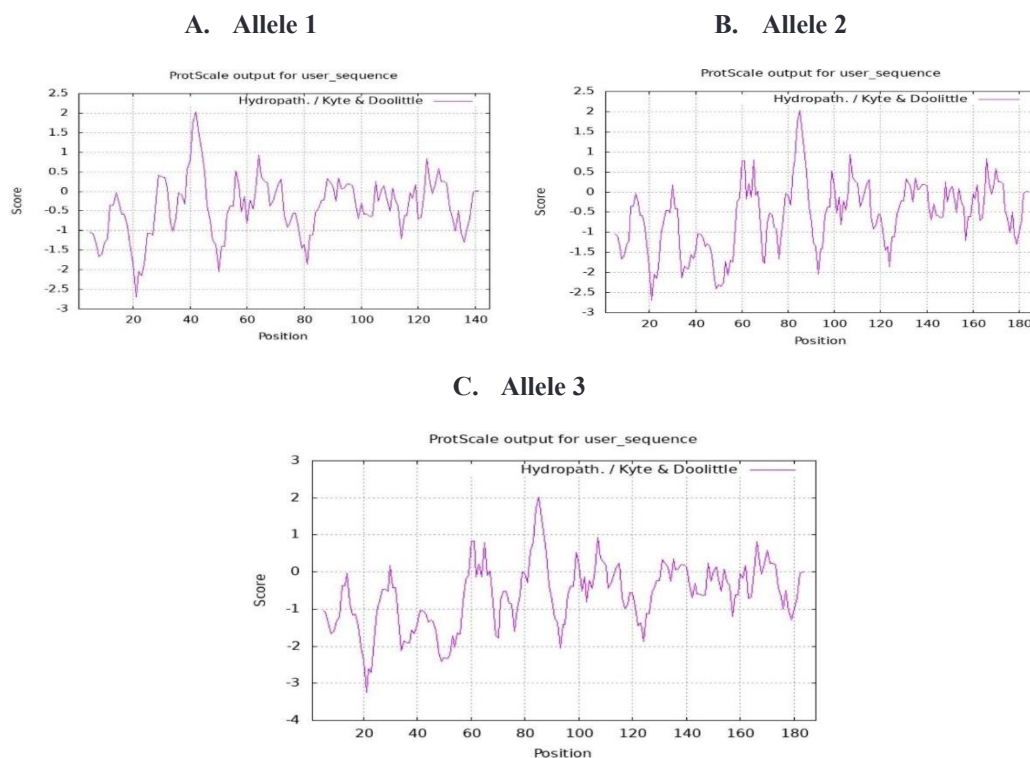


Figure 4. Hp Hob/Kyte & Doolittle hydropathy SpcB Graph Plot (a. Allele 1, b Allele 2 ,c. Allel 3).

To strengthen the results, its properties towards water are strengthened using the SOSUI program [SOSUI: submit protein sequences \(nagahama-i-bio.ac.jp\)](https://nagahama-i-bio.ac.jp/sosui/) which is shown in **Figure 5.**

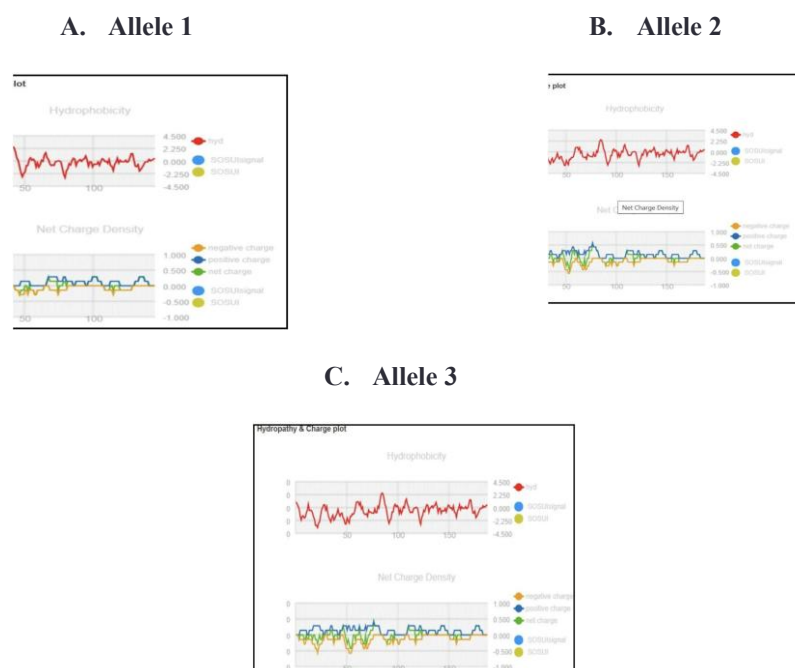


Figure 5. Graph Plot of SOSUI Program Prediction Results (a. Allele 1, b Allele 2 ,c. Allel 3).

The results obtained from using ProtScale and SOSUI are water-soluble proteins. The HpHob/Kyte & Doolittle hydropathy graph plot (**Figure 4**) shows many graphs that are below 0.

3. SpcB Secondary Structure

Secondary structure is used to study the structure and function of proteins. PSIPRED provides various secondary structure prediction methods. This site can also be used interactively via a web browser programmatically via the REST API [28,29]. Secondary structure was predicted using PSIPRED <http://bioinf.cs.ucl.ac.uk/psipred/> and use the SOPMA program. The secondary structure of the protein obtained is shown in **Figure 6**:

- The fold of the polypeptide chain is in the form of an alpha helix (Hh) which is a twist of the amino acid chain that has a spiral shape.
- Extended Strand (Ee), which is in the form of a sheet composed of amino acids linked to each other via hydrogen bonds,
- Random Coil (Cc), has a shape resembling a coil or coil that resembles a rope.



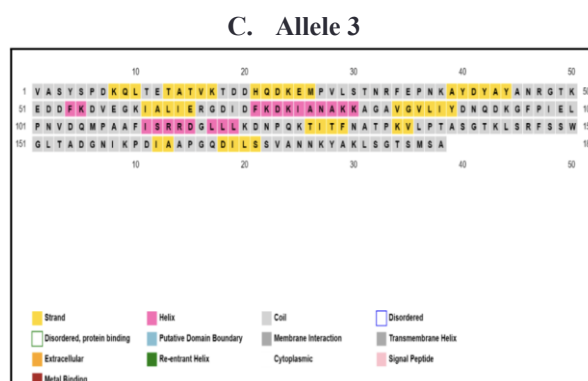


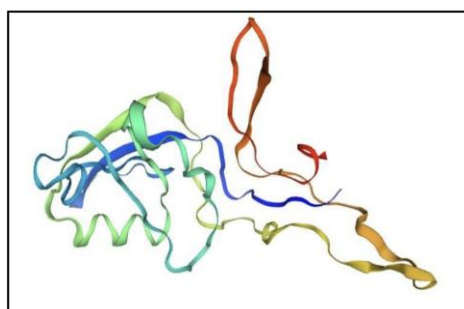
Figure 6. Secondary Structure of ScpB Protein (a. Allele 1, b Allele 2 ,c. Allel 3).

Table 6. Percentage of secondary structure prediction results for the ScpB protein using SOPMA

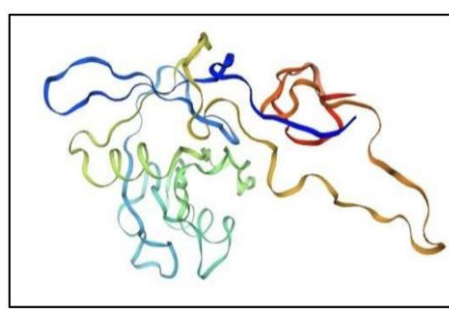
No.	Amino Acid Structure	Allele 1 (OR174997.1) Protein ID: WMQ58569.1		Allele 2 (OR174995.1) Protein ID: WMQ58567.1		Allel 3 (OR174990.1) Protein ID: WMQ58562.1	
		Percentage		Percentage		Percentage	
		Amount	(%)	Amount	(%)	Amount	(%)
1	Alpha helix (Hh)	22	15.17%	20	10.64%	20	10.64%
2	Extend strand (Ee)/ β -Sheet	26	17.93%	45	23.94%	43	22.87%
3	Beta turn (Tt)	5	3.45%	11	5.85%	12	6.38%
4	Random Coil (Cc)	92	63.45%	112	59.57%	113	60.11%

Strand dominates with the highest percentage of each allele among other secondary structure elements. Random coils or coiled strands have an important function in proteins for flexibility and conformational changes such as enzyme turnover [30]. Where the peptide bonds in the coil strand are not involved in intra- protein hydrogen bonds, this means the coil strand structure can freely interact with water molecules, small ligands or other proteins, this can be related to the enzymatic function of the protein [30,31]. The second most common component is the extended strand, alpha helix and then beta turn. If illustrated using the SWISS-MODEL program, you will get an example of a protein structure prediction in three dimensions (3D) as in **Figure 7**

A. Allele 1



B. Allele 2



C. Allele 3

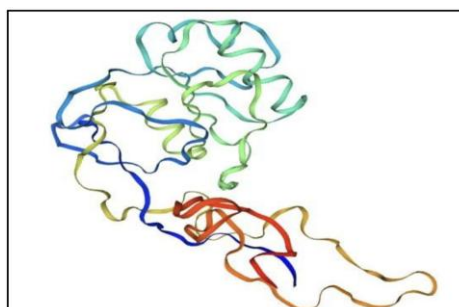


Figure 7. Example of SpcB protein structure prediction (a. Allele 1, b Allele 2 ,c. Allel 3).

4. Discussion

Streptococcus agalactiae has many virulence factors. See (Figure 8). Seeing the large number of virulence factors produced by GBS, it is necessary to carry out research on the characterization of virulence gene variations and the results of every country isolates so that later we will know more about the relationship of virulence genes to antibiotic resistance and their ability to produce biofilm.

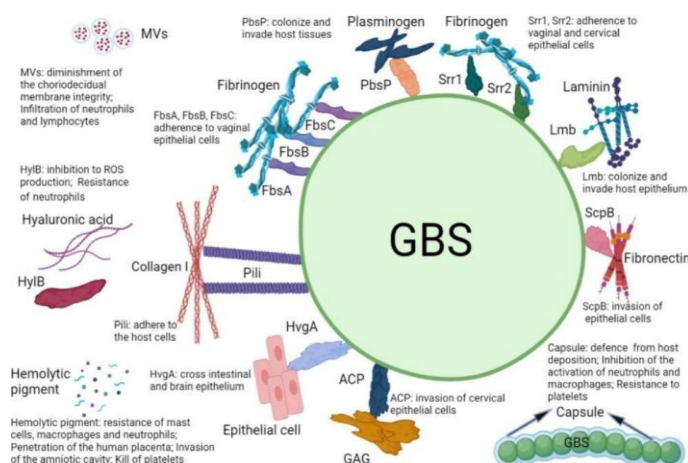


Figure 8. Virulence Factor Structure *Streptococcus agalactiae*.

The results of this *in silico* research need to be realized to study how far and how many variations of virulence-causing proteins are found in *Streptococcus agalactiae*. There is also a need for research related to antibiotic resistance other than those that have been studied (Table 3) so that patients suffering from *Streptococcus agalactiae* infections get the right and best.

4 Conclusion

Based on the results of the RFLP *in silico* it can be concluded that there are variations genetics by giving rise to 3 allelic variations in 10 ScpB gene sequences *Streptococcus agalactiae*. This shows that the level of virulence *Streptococcus agalactiae* different. The molecular characteristics of three

alleles have different results. Different results ranging from the amino acid composition, physico-chemical properties, primary and secondary structure.

References

1. de-Paris F, Machado AB, Gheno TC, Ascoli BM, Oliveira KR, Barth AL. Group B Streptococcus detection: comparison of PCR assay and culture as a screening method for pregnant women. *Braz J Infect Dis*. 2011 Jul- Aug;15(4):323-7. PMID: 21861001.
2. Hayati, Z. Infeksi Streptokokus Grup B (SGB) pada Ibu Hamil dan Neonatus : Diagnosis dan Pencegahan. *Jurnal Kedokteran Yarsi*. 2010. 18(1):079-085.
3. Park JS, Cho DH, Yang JH, Kim MY, Shin SM, Kim EC, Park SS, Seong MW. Usefulness of a rapid real-time PCR assay in prenatal screening for group B streptococcus colonization. *Ann Lab Med*. 2013 Jan;33(1):39-44. doi: 10.3343/alm.2013.33.1.39. Epub 2012 Dec 17. PMID: 23301221; PMCID: PMC3535195
4. Zanini da Rocha J, Feltraco J, Radin V, Vitola Gonçalves C, Almeida da Silva PE, Von Groll A. Streptococcus agalactiae colonization and screening approach in high-risk pregnant 2020 Apr 30;14(4):332-340. doi: 10.3855/jidc.12025. PMID: 32379709
5. Morozumi M, Chiba N, Igarashi Y, Mitsuhashi N, Wajima T, Iwata S, Ubukata K. Direct identification of Streptococcus agalactiae and capsular type by real-time PCR in vaginal swabs from pregnant women. *J Infect Chemother*. 2015 Jan;21(1):34-8. doi: 10.1016/j.jiac.2014.08.024. Epub 2014 Oct 5. PMID 25287153.
6. Guo D, Xi Y, Wang S, Wang Z. Is a positive Christie-Atkinson-Munch-Peterson (CAMP) test sensitive enough for the identification of Streptococcus agalactiae? *BMC Infect Dis*. 2019 Jan 3;19(1):7. doi: 10.1186/s12879-018-3561-3. PMID: 30606123; PMCID: PMC6318942
7. Helmig RB, Gertsen JB. Diagnostic accuracy of polymerase chain reaction for intrapartum detection of group B streptococcus colonization. *Acta Obstet Gynecol Scand*. 2017 Sep;96(9): 1070-1074. doi: 10.1111/aogs.13169. Epub 2017 Jun 22. PMID: 28504863
8. Cleary PP, Peterson J, Chen C, Nelson C. Virulent human strains of group G streptococci express a C5a peptidase enzyme similar to that produced by group A streptococci. *Infect Immun*. 1991 Jul;59(7):2305-10. doi: 10.1128/iai.59.7.2305-2310.1991. PMID: 2050400; PMCID: PMC258011.
9. Dai S, Long Y. Genotyping analysis using an RFLP assay. *Methods Mol Biol*. 2015;1245:91-9. doi: 10.1007/978-1-4939-1966-6_7. PMID: 25373751.
10. Aulia, N., Putri, C. K., Khairiyah, L., Khairunnisa, F. A., & Achyar, A. (2022, May). Analysis of Genetic Variation of Measles Virus Genotype and Isolate MF Intergenic Spacer Sequences PopSet: 2105287799 using RestrictionFragment Length Polymorphism (RFLP) In silico. In *Prosiding Seminar Nasional Biologi* (Vol. 1, No. 2, pp. 1139-1146).
11. Yeriska, F., Umar, M. Z., Hijriah, N. W., Fadhlurrohman, R., & Achyar, A. (2022, May). Analisis Variasi Genetik Sekuen Gen Pengkode Protein Spike Virus MERS-CoV (PopSet: 1843801421) Menggunakan RFLP Secara In silico. In *Prosiding Seminar Nasional Biologi* (Vol. 1, No. 2, pp. 898-906).
12. Pardede, S. O. Struktur Sel Streptokokus dan Patogenesis Glomerulonefritis Akut Pasca Streptococcus. *Jurnal Sari Pediatri* (Vol 11, No. 1, pp 56-66).
13. Zhu Z, Samuelson JC, Zhou J, Dore A, Xu SY. Engineering strand-specific DNA nicking enzymes from the type IIS restriction endonucleases BsaI, BsmBI, and BsmAI. *J Mol Biol*. 2004 Mar 26;337(3):573-83. doi: 10.1016/j.jmb.2004.02.003. PMID: 15019778.
14. Achyar et al. 2021. Analysis of Genetic Variations in POLY Gene Sequences in Dengue Virus 2 Using In-Silico RFLP. *Jurnal Bioscience*. Vol. 05. No. 1: 80-86.
15. Theodore, G. Schurr. 2000. Mitochondrial DNA and the Peopling of the New World Genetic variations among Native Americans provide further clues to who first populated the Americas and when they arrived. *American Scientist Online* (The Magazine of Sigma XI, The Scientific Research Society).
16. Fatchiyah, Arumingtyas, E. L., Widyarti, S., Rahayu, S. 2011. *Biologi Molekular, Prinsip Dasar Analisis*. Erlangga. Jakarta.

17. Hardjono S., 2013. Sintesis dan Uji Aktivitas Antikanker Senyawa 1-(2- Klorobenzoiloksi) urea dan 1-(4-klorobenzoiloksi) urea. *Berkala Ilmiah Kimia Farmasi*. 2 (1), 1.
18. Rosini R, Margarit I. Biofilm formation by *Streptococcus agalactiae*: influence of environmental conditions and implicated virulence factors. *Front Cell Infect Microbiol*. 2015 Feb 4;5:6. doi: 10.3389/fcimb.2015.00006. PMID: 25699242; PMCID: PMC4316791.
19. Miranda PSD, Lannes-Costa PS, Pimentel BAS, Silva LG, Ferreira-Carvalho BT, Menezes GC, Mattos-Guaraldi AL, Hirata R Jr, Mota RA, Nagao PE. Biofilm formation on different pH conditions by *Streptococcus agalactiae* isolated from bovine mastitic milk. *Lett Appl Microbiol*. 2018 Sep;67(3):235-243. doi: 10.1111/lam.13015. Epub 2018 Jul 10. PMID: 29856469.
20. Li G, Wei Y, Guo Y, Gong H, Lian J, Xu G, Bai B, Yu Z, Deng Q. Omadacycline Efficacy against *Streptococcus agalactiae* Isolated in China: Correlation between Resistance and Virulence Gene and Biofilm Formation. *Comput Intell Neurosci*. 2022 Apr 25;2022:7636983. doi: 10.1155/2022/7636983. PMID: 35510054; PMCID: PMC9061024.
21. Yang Y., Liu Y., Ding Y., et al. Molecular characterization of *Streptococcus agalactiae* isolated from bovine mastitis in Eastern China. *Plos One* . 2013;8(7) doi: 10.1371/journal.pone.0067755.e67755
22. Abrahamian F. M., Sakoulas G., Tzanis E., et al. Omadacycline for acute bacterial skin and skin structure infections. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* . 2019;69(1):S23–S32. doi: 10.1093/cid/ciz396.
23. Naga,W. S.,Adiguna, B., Retnoningtyas, E. S.,Ayucitra,A., & Daging, D.(2010). Koagulasi Protein Dari Ekstrak Biji Kecapir.*WidyaTeknik*, 9(1),1–11.
24. Pergande, M. R., & Cologna, S. M. (2017).Isoelectric Point Separations of Peptides and Proteins. <https://doi.org/10.3390/proteomes5010004>
25. Kaur, A., Kumar, P., Id, P., Pati, A. M., Kaur, A., & Id, N. (2020). Topological analysis ofpathogenesis-related proteins from *Arabidopsis thaliana* and *Oryza sativa* using in-silicoapproaches. 1 – 15. <https://doi.org/10.1371/journal.pone.0239836>
26. Osorio, D., Rondón villarreal,P. ,& Torres, R. (2012) Peptides: Apack agefordatamining ofantimicrobial peptides. XX, 1–11
27. Ali, M., Pandey, R. K., Khatoun, N., Narula,A., & Mishra,A. (2017). Exploring denguegenometo construct a multi- epitopebased subunit vaccine by utilizing immunoinformatics approach tobattle against dengue infection. *Scientific Reports*, July, 1–13.<https://doi.org/10.1038/s41598-017-09199-w>
28. Buchan,D.W.A.,&Jones,D.T.(2019).ThePSIP REDProtein AnalysisWorkbench:20 yearson. 47(April),402– 407.<https://doi.org/10.1093/nar/gkz297>
29. Moffat, L., & Jones,D.T. (2021). Structuralbioinformatics Increasing the accuracy ofsinglesequence prediction methodsusing a deep semi-supervised learningframework.37(July),37443751.<https://doi.org/10.1093/bioinformatics/bt ab491>
30. Filiz, E., & Koc, I. (2014). In silico sequence analysis and homology modeling of predicted beta-amylase 7-like protein in *Brachypodium distachyon* L . Brought toYou by CORE Provided byDirectory of Open Access Journals *J. BioSci. Biotech.*, 3(1): 61-6, 61–67
31. Nurani F. S., Sukenda S., Nuryati S. Maternal immunity of tilapia broodstock vaccinated with polyvalent vaccine and resistance of their offspring against *Streptococcus agalactiae*. *Aquaculture Research* . 2020;51(4):1513–1522. doi: 10.1111/are.14499.
32. Kamińska D, Ratajczak M, Nowak-Malczewska DM, Karolak JA, Kwaśniewski M, Szumala-Kakol A, Długaszewska J, Gajecka M. Macrolide and lincosamide resistance of *Streptococcus agalactiae* in pregnant women in Poland. *Sci Rep*. 2024 Feb 16;14(1):3877. doi: 10.1038/s41598-024-54521-y. PMID: 38366099; PMCID: PMC10873391.
33. Liu Y, Liu J. Group B *Streptococcus*: Virulence Factors and Pathogenic Mechanism. *Microorganisms*. 2022 Dec 15;10(12):2483. doi: 10.3390/microorganisms10122483. PMID: 36557736; PMCID: PMC9784991.
34. Hayes K, O'Halloran F, Cotter L. A review of antibiotic resistance in Group B *Streptococcus*: the story so far. *Crit Rev Microbiol*. 2020 May;46(3):253-269. doi: 10.1080/1040841X.2020.1758626. Epub 2020 May 2. PMID: 32363979.