CONSILIUM SANITATIS
JOURNAL OF HEALTH SCIENCE AND POLICY
Vol: 1, No. 2, April 2023, pp. 98 – 106
P-ISSN: 2985-9719 I E-ISSN: 2985-4628
Web: https://journals.eduped.org/index.php/JHSP

Primer Design and In Silico Amplification of *PhtD* Gene from *Streptococcus pneumoniae* as Serotype-independent Pneumococcal Vaccine Candidate

Desain Primer dan Amplifikasi Gen PhtD dari Streptococcus pneumoniae secara in silico Sebagai Kandidat Vaksin Pneumokokus Serotype-Independent

Risma Wiharyanti[1]*, Fera Aulia[2], Astiningdiah Merdekawati[3]

[1,2] Universitas Negeri Yogyakarta, Yogyakarta, Indonesia
[3]PT. Bio Farma, Bandung, Indonesia

*e-mail: rismawiharyanti@uny.ac.id

Abstract

Objective: This study aimed to obtain a primer capable of amplifying the Pneumococcal histidine triad protein D (PhtD) gene from various *S. pneumoniae* bacterial serotypes as a protein-based pneumonia vaccine candidate that does not depend on bacterial serotype.

Methods: Primer design was carried out through a bioinformatics approach using some software followed by using the primers in the amplification process, which was carried out in silico.

Results: The *PhtD* gene of the 6 serotypes of *S. pneumoniae* obtained from the genome centre has high similarity. The sequence also has 34 regions which are included in the conserved region. Based on the consensus of the six *S. pneumoniae* serotypes, a pair of primers that could amplify 17 *S. pneumoniae* strains from several serotypes contained in the reference sequence in silico was obtained.

Conclusion: The similarity of the *PhtD* gene sequence in *S. pneumoniae* makes the PhtD protein a potential candidate to be developed into a serotype-independent protein-based vaccine. The primers designed in this study can amplify target genes on various *S. pneumoniae* serotypes and are specific for the *PhtD* gene.

Keywords: In silico, pneumonia, primer, *Streptococcus pneumonia*e, vaccine.

Article History	Submitted	Revised	Accepted
	2023-04-07	2023-04-08	2023-04-09

Vol. 1, No. 2, April 2023, pp. 98 – 106
https://doi.org/10.56855/jhsp.v1i2.283

Introduction

Streptococcus pneumoniae is a gram-negative bacterium that causes most infections in humans' upper and lower respiratory tract. *S. pneumoniae* bacteria are capable of colonizing the nasopharynx, which results in otitis media, pneumonia conjunctivitis, meningitis and septicemia, especially in children, the elderly, and individuals with low immunity[1,2]. *S. pneumoniae* infection is the most common cause of death from pneumonia, especially in children under 5 years of age. Pneumonia is listed as the first cause of death due to infection in children worldwide. Approximately 14.5% of infant deaths and 5% of under-five deaths annually in Indonesia are caused by pneumonia. In developing countries, the mortality rate of 30% of cases of pneumonia due to pneumococcal infection affecting adults reaches 11% to 40%[3].

Prevention of transmission of *S. pneumoniae* bacteria can be done through vaccine administration. The importance of prevention and the magnitude of the impact caused by *S. pneumoniae* bacteria caused the Indonesian government to change the status of pneumonia immunization for toddlers from an additional vaccine to a mandatory vaccine. Pneumonia vaccines that have received FDA approval for use at this time are Pneumococcal Conjugate Vaccine (PCV) and Pneumococcal Polysaccharide Vaccine (PPSV)[3,4]. The PCV vaccine is formed from the polysaccharide capsule of *S. pneumoniae* bacteria, purified and conjugated with a non-toxic diphtheria toxin. At the same time, PPSV is the polysaccharide part of the capsule which is purified but without conjugation[4,5]. The two vaccines still depend on the serotype or type of *S. pneumoniae* bacteria used. The PCV-type vaccine is only available to protect up to 20 serotypes. At the same time, PPSV can protect up to 23 of the *S. pneumoniae* serotypes that most commonly infect humans, so the vaccine has low efficacy against non-invasive pneumococcal disease[5].

However, increasing the number of vaccine serotypes in PCV and PPSV formulations is complex and expensive because there are at least 98 pneumococcal serotypes[5,6]. Meanwhile, the need to develop a new pneumococcal vaccine with wider protection (more serotypes) is urgent. The limitations of polysaccharide-based vaccines have prompted the development of new strategies for the manufacture of pneumococcal vaccines. One approach that can be taken is to make vaccines from antigens based on conserved proteins found in various *S. pneumoniae* serotypes [2,7] .

Pneumococcal histidine triad protein D (PhtD) is an adhesin that helps the attachment of *S. pneumoniae* to the epithelial cells of the respiratory tract. The PhtD protein is also involved in metal ion homeostasis, avoiding the decomposition of complement C3b and assisting other virulence mechanisms in the lung[4,8]. These proteins are localized on the bacterial surface or cell wall and are accessible to antibodies. Experiments on animal models proved that PhtD could provide a protective effect during the colonization challenge test. In addition, PhtD is known to be a conserved protein in various *S. pneumoniae* serotypes, making it an optimal candidate as a protein-based vaccine[2].

One of the stages in researching the immunogenicity of a protein or research based on genetic material is to obtain the coding sequence for the protein, which can be carried out using the Polymerase Chain Reaction (PCR) method. The PCR process is a process of amplifying or amplifying DNA using primers. Primers are short nucleotide sequences that recognize specific sequences that are targets for amplification[9]. In carrying out the PCR of the *PhtD* gene, primers with high similarity values from different serotypes of *S. pneumoniae* are needed to be able to amplify the gene optimally. This study aims to analyze the similarity of the PhtD protein-coding gene sequences from various *S. pneumoniae* serotypes using a bioinformatics approach and then to design a *PhtD* gene primer that can be used to amplify this gene from various serotypes for further research.

Methods

PhtD Gene Searching from Several *S. pneumoniae* Serotypes at the Genome Data Center

The *PhtD* of *S. pneumoniae* gene database is accessed from the National Center for Biotechnology Information (NCBI) information data centre which integrates genomic data from several sources, such as GenBank, UniProt, Immune Epitop Database, and Protein Data Bank. These genomic data are then downloaded in FASTA format.

PhtD Gene Alignment Process of *S. pneumoniae*

The alignment process was performed using CLC Genomic Workbench 3.6.5 software. In this alignment process, which areas are the conserved regions of all serotypes will be seen.

Vol. 1, No. 2, April 2023, pp. 98 – 106
https://doi.org/10.56855/jhsp.v1i2.283

Determination of Conserved Region and Consensus of *PhtD* Gene Sequence Alignment Results

The determination of the conserved regions in the *PhtD* gene sequence is a continuation of the sequence alignment results using the Bioedit 7.0.5.3 software. The consensus formed from the alignment results is then downloaded and analyzed further.

Primer Design for *PhtD* Gene Sequence Amplification

The result of the previous alignment in the form of a consensus sequence is used as the basis for making primers. The designed primer must be used to obtain the target gene sequence and multiply the gene. Primers were made with the help of snapGene software according to the requirements of suitable primers. Furthermore, the primers were analyzed using a web-based program that can be accessed at idtdna.com.

In silico *PhtD* Gene amplification

The amplification or PCR process was carried out in silico using a web-based program which can be accessed on the insilico.ehu.es website. Amplification was carried out on several reference serotypes of *S. pneumoniae* contained in the program.

Results

***S. pneumoniae PhtD* Genome Data**

From the search results at the genomic data center, we obtained *PhtD* gene sequences from various strains and serotypes (Table 1) which were used as the basis for making *PhtD* gene primers.

Tabel 1. Genome Data of *S. pneumoniae* Used in the Determination of the *PhtD* Gene Sequence

No	Serotipe	Strain	*GenBank Acc. Number*
1	2	D39	CP000410
2	3	OXC141	FQ312027
3	4	Xen35	CP025256
4	4	TIGR4	CP089948
5	14	INV200	FQ312029
6	19F	ATCC49619	AP018938
7	23F	ATCC700669	FM211187

Vol. 1, No. 2, April 2023, pp. 98 – 106
https://doi.org/10.56855/jhsp.v1i2.283

Alignment and Determination of *PhtD* Gene Sustainable Areas

Following are the results of the alignment of the 7 PhtD gene sequences obtained from the NCBI genome data center using the CLC Genomic Workbench 3.6.5 software (Figure 1). Figure 2 is a conserved region in the *PhtD* gene from the results of data processing using Bioedit 7.0.5.3 software.

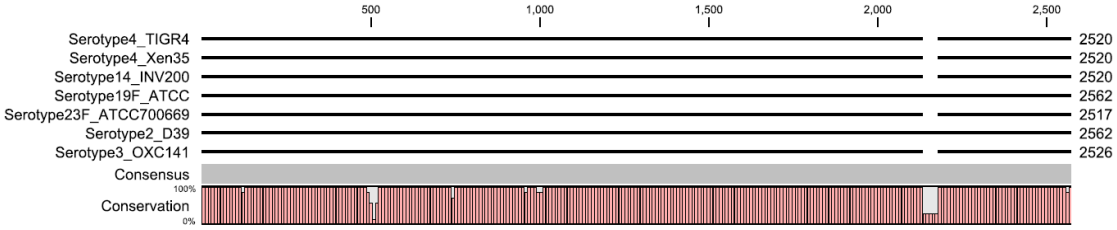

Figure 1. *PhtD* Gene Alignment Results of 6 *S. pneumoniae* serotypes

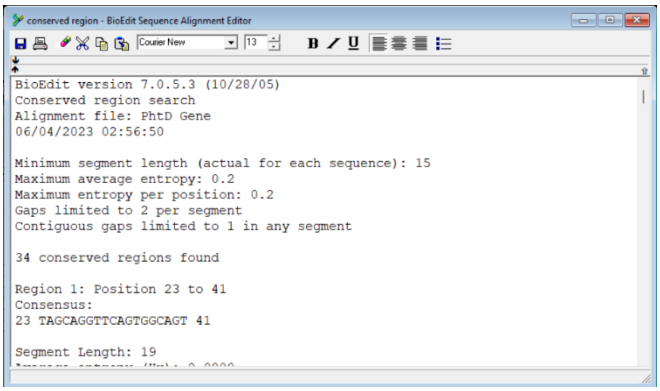

Figure 2. Conserved Region results of *PhtD* gene alignment of 6 serotypes of *S. pneumoniae*

Consensus Results of PhtD Gene Alignment

The Following are the results of the alignment of the initial 60 bases and the final 51 bases of the *PhtD* gene from various *S. pneumoniae* serotypes. The consensus formed is used as the basis for making primers.

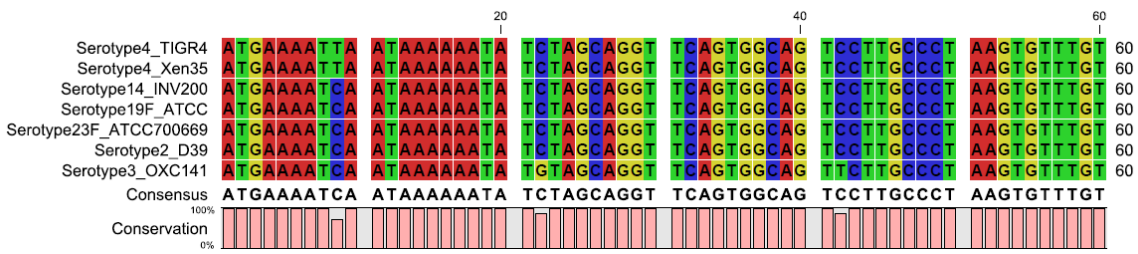

Figure 3. Alignment results of the initial 60 bases of the *PhtD* gene of *S. pneumoniae*

Vol. 1, No. 2, April 2023, pp. 98 – 106
https://doi.org/10.56855/jhsp.v1i2.283

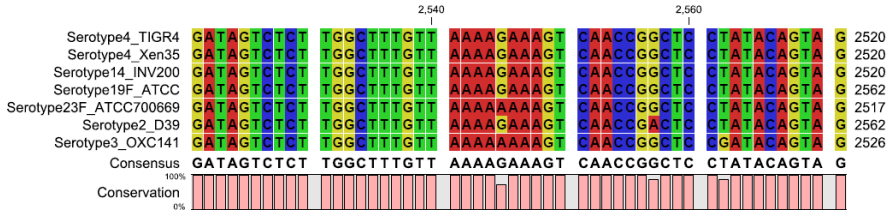

Figure 4. Alignment results of the final 51 bases of the *PhtD* gene of *S. pneumoniae*

PhtD Gene Primer Design and Analysis

The perimer design results using the SnapGene software are shown in Figure 5.

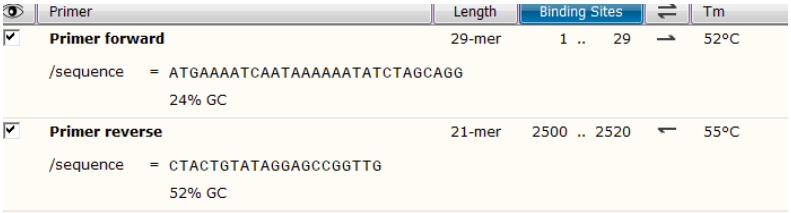

Figure 5. Design results of forward primer and reverse primer

In silico *PhtD* Gene Amplification Using Primers on Various *S. pneumoniae* Bacterial Serotypes

Primers were used to amplify several reference serotypes of *S. pneumoniae* on the site insilico.ehu.es. Figure 6 shows the results of the in silico PhtD gene amplification.

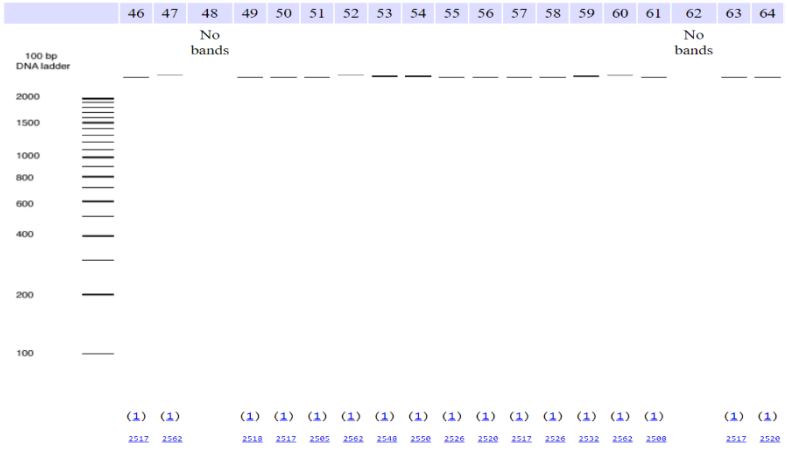

Figure 6. *PhtD* gene amplification from various *S. pneumoniae* reference serotypes in silico. Nos 46-64 indicate the *S. pneumoniae* (SP) strain used as a reference, namely 46: SP 670-6B, 47: SP 70585, 48: SP A026, 49: SP AP200, 50: SP ATCC 700669, 51: SP CGSP14, 52: SP D39, 53: SP G54, 54: SP Hungary19A-6, 55: SP INV104, 56: SPINV200, 57: SP JJA, 58: SP OXC141, 59: SP P1031, 60: SP R6, 61: SP SPNA45, 62: SP ST556, 63: SP TCH8431/19A, 64: SP TIGR4

Vol. 1, No. 2, April 2023, pp. 98 – 106
https://doi.org/10.56855/jhsp.v1i2.283

Discussion

From the search results in the genome database, 7 *PhtD* gene sequences were found from 6 serotypes of *S. pneumoniae*, namely serotypes 2, 3, 4, 14, 19F and 23F. These six serotypes are present in commercial vaccines currently used, namely PCV and PPSV4. The seven *PhtD* gene sequences from the six serotypes obtained were used to represent the various S. pneumoniae serotypes in observing sequence homogeneity and continuity through the alignment process.

The alignment results in Figure 1 shows that the *PhtD* gene has a high degree of similarity across all serotypes. The similarity of the gene sequences indicates the similarity of the structure of the protein formed and the function of the protein. This shows that the PhtD protein has an almost similar structure in all the serotypes tested. The protein shared by all serotypes is very essential because it can prevent serotype replacement, namely the emergence of disease due to one of the *S. pneumoniae* serotypes that are not found in polysaccharide and conjugate-based vaccines. Serotype replacement also increases the incidence of antibiotic resistance against bacterial serotypes that are not contained in vaccines[3].

Determination of the conserved region of the *PhtD* gene is also needed to identify areas with low gene sequence variation. Based on the analysis using the BioEdit Sequence Alignment Editor in Figure 2, 34 regions in the *PhtD* gene are conserved areas. These sustainable areas can be useful in developing pneumococcal vaccines through various approaches, such as virus-like particle (VLP) based protein or DNA and RNA vaccines[10].

The similarity of the *PhtD* gene sequence is significant in the basis of primer production because the primers made are expected to amplify the *PhtD* gene in all serotypes. Based on this, the consensus from the sequence alignment results was used as the basis for preparing primers (Figures 3 and 4). Several bases from the start and end of the sequence were analyzed using the SnapGene software because the expected protein targets were all parts of the protein from the start codon (AUG) to the stop codon (TAG). The overall protein expression of the *PhtD* gene aims for the protein to contain epitopes that will be recognized by the immune system so that it can trigger the body's immunity against *S. pneumococcus* infection.

The primer that has been designed (figure 5) has characteristics in accordance with the requirements of a suitable primer, including the primer having a melting temperature

Vol. 1, No. 2, April 2023, pp. 98 – 106
https://doi.org/10.56855/jhsp.v1i2.283

(Tm) between 50-60^{O}C. A primer with a Tm that is too high will easily experience mispriming or stick to the non-target area[11]. The *PhtD* gene primer has a Tm forward primer at 52^{O}C while the Tm reverse primer is at 55^{O}C, so the primer pairs have a temperature difference of 3^{O}C. A good pair of primers have a temperature difference between forward and reverse primers of less than 5^{O}C. This is based on the probability of failure of the amplification process. Primer pairs with a Tm difference of more than 5^{O}C can cause a decrease in the process so that no amplification process occurs[11,12].

The percentage of good GC from a primer range from 40-60%. The reverse primer that has been designed (figure 5) only has a rate of 24%. A low %GC value will reduce the stability of the primer to stick to the template. However, the primer can still be used if a similar sequence is not found in the genome. The analysis results of the primer and secondary structures show that both the forward and reverse primers do not have the possibility of secondary structures appearing in the temperature range of the amplification process.

Tests using primers in the amplification process or silico PCR showed that the *PhtD* gene primers made could amplify 17 strains of *S. pneumoniae* in the references, and only 2 strains were not amplified. In silico analysis is an amplification process carried out through computational prediction. The result of PCR product is quite uniform, namely in the range of 2500pb. Therefore it can be concluded that the primers that have been made can amplify target genes on various *S. Pneumoniae* serotypes and are specific for the *PhtD* gene.

Conclusion

The *PhtD* gene of the six serotypes of *S. pneumoniae* obtained from the genomic data centre has high similarity. The sequence has 34 regions which are included in the conserved region. These sequence similarities make the PhtD protein a potential candidate for use as a protein-based vaccine.

Based on the consensus of the six S. *pneumoniae* serotypes, a pair of primers that could amplify 17 *S. pneumoniae* strains from several serotypes contained in the reference sequence in silico was obtained. It could be concluded that the primers that have been made can amplify target genes on various *S. Pneumoniae* serotypes and are specific for the P*htD* gene.

References

1. Parameswarappa SG, Reppe K, Geissner A, et al. A Semi-synthetic Oligosaccharide Conjugate Vaccine Candidate Confers Protection against *Streptococcus pneumoniae* Serotype 3 Infection. *Cell Chem Biol*. 2016;23(11):1407-1416. doi:10.1016/j.chembiol.2016.09.016
2. André GO, Assoni L, Rodriguez D, et al. Immunization with PhtD truncated fragments reduces nasopharyngeal colonization by *Streptococcus pneumoniae*. *Vaccine*. 2020;38(26):4146-4153. doi:10.1016/j.vaccine.2020.04.050
3. Daniels CC, Rogers PD, Shelton CM. A review of pneumococcal vaccines: Current polysaccharide vaccine recommendations and future protein antigens. *J Pediatr Pharmacol Ther*. 2016;21(1):27-35. doi:10.5863/1551-6776-21.1.27
4. Feldman C. Review: Current and new generation pneumococcal vaccines. *J Infect*. 2014;69(4):309-325. doi:10.1016/j.jinf.2014.06.006
5. Santiesteban-Lores LE, Cabrera-Crespo J, Carvalho E. Development of a pneumococcal conjugate vaccine based on chemical conjugation of polysaccharide serotype 6B to PspA. *Microb Pathog*. 2021;158(April):105092. doi:10.1016/j.micpath.2021.105092
6. Geno KA. Pneumococcal capsules and their types: Past, present, and future. *Clin Microbiol Rev*. 2015;28(3):871-899. doi:10.1128/CMR.00024-15
7. Masomian M, Ahmad Z, Gew LT, Poh CL. Development of next generation *Streptococcus pneumoniae* vaccines conferring broad protection. *Vaccines*. 2020;8(1):1-23. doi:10.3390/vaccines8010132
8. Oliveira GS, Oliveira MLS, Miyaji EN, Rodrigues TC. Pneumococcal vaccines: Past findings, present work, and future strategies. *Vaccines*. 2021;9(11). doi:10.3390/vaccines9111338
9. Ayu Safitri T, Nurul Jannah Patty D, Saraswati H. 16 dan 18 Sebagai Dasar dalam Desain Primer untuk Deteksi Kanker Leher Rahim dengan In-house. *IJOBB*. 2018;2(2):67.
10. Saraswati H. Analisa Bioinformatika Gen E1 dan E2 Dari Virus Hepatitis C (HCV) Genotipe 1, 2, 3 dan 6 Sebagai Kandidat Vaksin Viral- Like Particles (VLP). *Indones J Biotechnol Biodivers*. 2017;2(2):48-57.
11. Seprianto, Wahyuni FD. Analisis Bioinformatika Gen Potensial Penyandi Halichondrin B Dari Spons Laut Sebagai Kandidat Anti. *Indones J Biotechnol Biodivers*. 2018;2(2):57-66.
12. Mardalisa, Suhandono S, Yanti N, Rozi F, Nova F, Primawati. Bioinformatic analysis in designing mega-primer in overlap extension PCR cloning (OEPC) technique. *Int J Informatics Vis*. 2021;5(2):139-143. doi:10.30630/joiv.5.2.459