



## Design of the *env-su* Gene Coding for Surface Unit Protein as a Vaccine Candidate for Jembrana Disease Virus in In Silico

Nur Asih Setiarini<sup>1\*</sup>, Indriawati<sup>2</sup>, R Susanti<sup>3</sup>, Endang Tri Margawati<sup>4</sup>

<sup>1,3</sup>Department of Biology, Faculty of Mathematics and Natural Sciences, State University of Semarang  
D6 Building, 1<sup>st</sup> Floor Campus Sekaran Gunungpati, Semarang, Central Java 50229, Indonesia

<sup>2,4</sup> Biotechnology Research Center, Indonesian Institute of Sciences

Jln. Raya Jakarta-Bogor No. Km 46, Cibinong, Bogor, West Java 16911, Indonesia

\*Corresponding Author:

e-mail: [Nurasihsetiarini02@gmail.com](mailto:Nurasihsetiarini02@gmail.com)

### Article History

Received : 19 February 2021  
Revised : 3 March 2021  
Approved : 22 March 2021  
Published : 31 March 2021

### Keywords

Vaccine candidate; env-su gene; sequence alignment; codon optimization; in silico

### ABSTRACT

Bali cattle are superior meat producers, but they are susceptible to Jembrana disease. The injection of the crude vaccine was considered ineffective, so the *env-su* gene was selected to express the Jembrana Surface Unit (JSU) protein as a candidate for the Jembrana vaccine. This study aimed to analyze the potential of the JSU protein as a candidate for the Jembrana vaccine and analyze the increase in the *env-su* gene expression which codon has been optimized in silico. Vaccine design was carried out through in silico including the selection of the SU protein sequences of the genus Lentivirus and sequence alignment of UniProt. The construction phylogeny tree of SU protein using MEGA-X program, optimization of *env-su* gene codon with preference codon *Esherichia coli* str. K-12 substr. MG1655 using *Optimizer*. The optimized *env-su* gene was inserted into the plasmid pET-21a (+) using *GenScript*. The result of sequence alignment showed that there is no SU protein that has a percent identity value of more than 30% with JSU protein. The SU JDV and BIV proteins are monophyletic groups and have a percent identity of 20.57%. Codon optimization showed an increase in CAI by 1,000 and GC 54.5%, and a decrease in ENc to 22 and AT 45.5%. EcoR1 and HindIII can recognize the gene target and MCS cut regions on the plasmid so that the *env-su* gene can be inserted into the pET-21a (+) plasmid. The JSU protein has the potential to be a candidate for the Jembrana vaccine, but it needs further research in vitro and in vivo.

**How to cite:** Setiarini, N.A., Indriawati., Susanti, R., & Margawati, E.T. (2021). Design of the *env-su* Gene Coding for Surface Unit Protein as a Vaccine Candidate for Jembrana Disease Virus in In Silico. *Jurnal Riset Biologi dan Aplikasinya*, 3(1):13-22. DOI: <https://doi.org/10.26740/jrba.v3n1.p13-22>.

### INTRODUCTION

Bali cattle are one of the original Indonesian cattle breeds with great potential as a meat producer because they have many advantages, including adaptability to critical environments, good reproduction, and growth rates (Astuti, 2018). However, the weakness of Bali cattle is that they are susceptible to blood sweat disease called Jembrana disease (Aryadrie et al., 2015). Jembrana disease is an endemic disease that first appeared and became epidemic in Sangkaragung Village, Jembrana District, Jembrana Regency, Bali Province in 1964 (Mardiatmi, 2015).

Along with technological advances in the field of genetic engineering, and the availability of genome sequence information for various types of viruses, recombinant DNA technology can be applied to produce functional proteins as vaccine ingredients (Ali, 2015). The Jembrana Disease Virus (JDV) genome in the GenBank can be used as a reference for Jembrana disease vaccine material, which is inserted in the expression plasmid and produced in *E. coli* host cells.

The Jembrana Surface Unit (JSU) protein plays a role at the beginning of the replication process by interacting in the binding process between the Jembrana Disease Virus particles and the surface of

the host cell so that it can trigger an antibody response that can block the recognition of receptors and prevent the process of entering the virus into the target cell, therefore genes *env-su* was selected as a candidate for the Jembrana vaccine to express the JSU recombinant protein (Indriawati et al., 2013; Kusumawati et al. 2015). The test for the potential of the JSU protein as a candidate for the Jembrana vaccine needs to be carried out with sequence alignment using UniProt to find out how much acidity is in the JSU protein with other SU proteins of the *Lentivirus* genus, the test results are determined in the percent identity calculation (Pearson, 2013).

The use of *E. coli* as a host cell in the production of recombinant protein because *E. coli* is easy to manipulate and has a short life cycle to minimize production costs and the resulting high and fast rate of target protein expression. However, the production of recombinant protein using *E. coli* host cells also have drawbacks, namely the occurrence of codon bias that affects the results of target protein expression by forming inactive protein aggregates (*inclusion bodies*) (Maksum et al., 2017; Silaban et al., 2017).

The strategy to overcome this is the use of synthetic gene technology from target genes to match the host cell preference codons (Silaban et al., 2017). The codon optimization method is a process that can adjust the *env-su* JDV gene codon coding for the JSU protein with preference codons in *E. coli* as the host (Lanza et al., 2014; Rosano & Ceccarelli, 2014). The discovery and development of the JSU recombinant protein vaccine design were carried out using the in silico test beforehand to reduce the risk of possible failure and loss and to save time and money (Ullah et al., 2020). Therefore, it is necessary to research to determine the JSU protein encoded in the JDV *env-su* gene as a candidate for the Jembrana disease vaccine based on the percent identity value of the sequence alignment with the SU protein of the *Lentivirus* genus and knowing the optimization results of the *env-su* JDV gene codon in silico will increase expression JSU recombinant protein.

## MATERIALS AND METHODS

The research was conducted from August to December 2020, at the Biology Laboratory, Faculty of Mathematics and Natural Sciences, Semarang State University. The research was conducted using the ASUS E202S series computer with Windows 7 system specifications and a dual-core Intel Celeron N3050 processor. The bioinformatics and

biocomputing tools used include NCBI, UniProt, GenScript, MEGA-X, OPTIMIZER. The materials used are various amino acid sequences of SU protein which are incorporated into the genus *Lentivirus*, such as Feline Immunodeficiency Virus (FIV), Simian Immunodeficiency Virus (SIV-mac), Bovine Immunodeficiency Virus (BIV), Caprine Arthritis Encephalitis Virus (CAEV-63), Jembrana Disease Virus (JDV), Equine Infections Anemia Virus (EIAV), dan Ovine Maedi Visna Related Virus (SA-OMVV).

### Selection of Sequences and Protein Sequence Alignment Processes SU

UniProt is used for the collection of various SU proteins of the genus *Lentivirus* and sequence alignment processes. The search was carried out with the keyword "envelope virus species name", selected entities that had the access code "ENV" and selected entries that had been reviewed by the SwissProt curator. Select entries that have the number of env amino acid sequences that correspond to the total env nucleotide sequences in NCBI. Each SU protein from various species is added to the "add basket" option available in the UniProt tool. The JSU protein sequence alignment process is carried out alternately with each other organism's SU protein sequence. Finally, a table of the results of the alignment of the JSU protein with each other SU protein gene includes the value of percent identity, identical positions, and similar positions.

### Construction of the *Lentivirus* Genus SU Protein Phylogenetic Tree

Each SU protein from various species is added to the "add basket" option available on the UniProt tool and downloaded in one file with the fasta format. MEGA-X software was used for the production of phylogenetic tree dendrogram for the SU protein genus *Lentivirus*. A file containing a set of genes encoding the *env-su* genus *Lentivirus* is opened in the MEGA software. Select the option "alignment by the cluster algorithm" to perform multiple sequence alignment processes. Select the "phylogenetic" menu and select the type of phylogenetic tree "construct/test neighbor-joining tree" with the "bootstrap method" 1000x. Click the "view" menu then select the option "branch lengths" and "starts/frequency". Finally, save the file as PNG.

### JDV *env-su* Gene Optimization

The OPTIMIZER website is used for the codon optimization process. The nucleotide sequence of the *env-su* gene was searched for in the complete JDV genome and nucleotides were selected in sequences—5197 to 6463 that encode the *env-su* JDV gene. *Env-su* gene nucleotides were inserted into the column, the host cell reference codon was selected from the HEG-DB data and selected "*Escherichia coli str. K-12 substr. MG1655*". Select the host cell genetic code, namely eubacteria, then the method used is "one AA-one codon". The resulting codon is optimized by the OPTIMIZER device and then the GC and ENc values are analyzed as well as the GC and AT content. To analyze the genome regions identified by the restriction enzymes using the GenScript website, *EcoRI* and *HindIII* enzymes were selected.

### Insertion of the *env-su* JDV gene into the Plasmid *pET-21a (+)*

The optimized *env-su* codon is inserted into the *pET-21a (+)* expression plasmid using the Genscript website. The first step, select the option "create construct", then given the name plasmid to be designed. Click the "the commons" option, then select "popular commercial vector". Select "bacterium" and look for the *pET-21a (+)* vector (plasmid), then click the "create construct" option under the column. The *env-su* gene is inserted through the MCS site on the plasmid by selecting the "sequence" option then "insert", then the "add" option is selected in the "annotation" menu and given the name *env-su*. The optimization process is carried out using the "gene optimization" option which is in the "sequence" menu by selecting the *E. coli* host cell as well as the *EcoRI* and *HindIII* restriction enzymes. To find out the plasmid design

errors made, you can select the "project" menu and select the "check design" option.

## RESULTS AND DISCUSSION

A total of seven *Lentivirus* genus SU protein sequences were taken from the UniProt website (Table 1). From the sequence alignment between these proteins, the evolutionary lines that culminate in a common ancestor can be studied (Bu'ulolo et al., 2010). The results of the alignment of the JSU protein with SU proteins of other species belonging to the genus *Lentivirus* are presented in Table 2. Based on the determination of the percent identity value limit used to compare between protein sequences, there is no SU protein encoded by other *env-su* genes in the genus *Lentivirus* that have percent identity more than 30% with Jembrana Surface Unit JDV protein. The results obtained in Table 2. are reinforced by the statement of Bu'ulolo et al (2010), that the farther the sequence equation is aligned, the more specific the type of material (sequence) is compared, on the contrary, the closer the equation between the sequences means the material is compared. the more common types. This shows that the JSU protein has the potential to be used as a vaccine candidate for Jembrana disease based on the percent identity value of the in-silico sequence alignment. Specific proteins are chosen as vaccine raw materials because they do not induce side effects, induce the formation of specific antibodies, and do not trigger an immune response against unwanted host cells (Gustiananda, 2011; Susmiarsih, 2018).

Percent identity on the alignment of sequences of amino acid and nucleotide sequences from two or more different species can be calculated by the number of identical positions divided by the total number of characters then multiplied by 100%

**Table 1.** A trace list of various proteins of the Env genus *Lentivirus* using the UniProt website

Access Number	Access Name	Virus Species	AA Length
Q82857	ENV_JEMBR	Jembrana Disease Virus (JDV)	781
P19557	ENV_BIV29	Bovine Immunodeficiency Virus (Strain R29) (BIV)	904
P16090	ENV_FIVPE	Feline Immunodeficiency Virus (Isolate Petaluma) (FIV)	856
P32541	ENV_EIAVC	Equine Infectious Anemia Virus (Isolate CL22) (EIAV)	859
P31627	ENV_CAEVG	Caprine Arthritis Encephalitis Virus (Strain 63) (CAEV-63)	942
P16899	ENV_OMVVS	Ovine Maedi Visna Related Virus (Strain South Africa) (SA-OMVV) (Ovine lentivirus)	990
P08810	ENV_SIVM2	Simian Immunodeficiency Virus (Isolate Mm251) (SIV-mac)	881

Note: AA = The total amino acid protein Env

(Ernawati et al., 2014). However, the percent identity calculation formed in each program has a different value even though the object of research is the same. This is influenced by the programming techniques for the tools used (Sunarto, 2015).

The sequence alignment process of various SU proteins of the *Lentivirus* genus using the UniProt website can only measure the similarity of the JSU protein sequences with other SU proteins without knowing the relationships between species in the *Lentivirus* genus. Therefore, it is necessary to use a multiple sequence alignment method with clustalW to construct a phylogeny tree that describes the kinship relationship based on genetic distances between sequences in the SU protein of the genus *Lentivirus* (Daugelaite et al., 2013).

SU protein belonging to the genus *Lentivirus* was reconstructed into a phylogenetic tree formed from neighbor-joining (NJ) analysis and tested by statistical tree evaluation using bootstrap 1000 times. A dendrogram-shaped phylogeny tree is presented in Figure 1 based on the value of the branch length and bootstrap test.

*Lentivirus* genus SU protein phylogeny tree constructs shown in Figure 1 form two clades with similar evolutionary distance values and a bootstrapping value of 70%. Clade 1 members consist of two distinct internal nodes, internal node 1 consists of SU EIAV and SIV-mac proteins while internal node 2 is composed of SU CAEV-63 and SA-OMVV proteins. Meanwhile, the members of clade 2 are also composed of two internal nodes, namely internal node 1 is occupied by SU FIV protein and internal node 2 contains SU JDV and BIV proteins.

The CAEV-63 species with SA-OMVV can be said to be closely related because they have almost the same branch length value, namely 0.23 for SA-OMVV and 0.22 for CAEV-63 calculated from the same node, which means that the two species are derived from one node ancestral unit. Besides, the bootstrap test results showed that both species were monophyletic. Monophyletic groups are shown in branches with a bootstrap value of 100% which are assumed to carry genetic and biochemical traits or patterns from one common ancestor (Akhsani et al., 2017; Retnaningati, 2017). A branch is monophyletic if all grouped OTUs have adjacent branches because the group shares the same ancestor compared to other groups with different lineages (Pangestika et al., 2015). Caprine Arthritis Encephalitis Virus (CAEV-63) and Visna Related Virus (SA-OMVV) have many genetic similarities in

the env structural genes because they belong to the SRLV (Small-ruminant lentiviruses) genome (Bartak et al., 2018). From the sequence alignment process using UniProt, the two proteins have a percent identity value of 67.864%.

In Figure 1, it can be seen that the SU JDV and BIV proteins are also monophyletic groups. The SU JDV and BIV proteins entered the same node with the branch length values between these species which were quite close, namely 2.47 in JDV and 1.19 in BIV. This branch length value strengthens the sequence alignment results that the two proteins have similarities in the env gene, which is thought to be inherited from one ancestral unit (Kusumawati et al., 2014).

SU SIV-mac and EIAV proteins are in one internal node but have far enough branch length values, namely 5.45 for SIV-mac and 1.96 for EIAV. This explains that the two species have different evolutionary distances even though they are formed at the same node, this statement is supported by the results of the bootstrap test on the SU SIV-mac and EIAV protein branches forming a value of 54% which is classified in the weak category, so there is still a possibility of changes in clade arrangement. The farthest genetic distance occurs in SU BIV and SIV-mac proteins, because the two proteins are separated by different clades.

The *env-su* JDV gene has a fragment length of 1266 nucleotide bases encoding the JSU protein with a molecular weight of 47 kDa and a length of 422 amino acids (Kusumawati et al., 2014). The nucleotide sequence of the *env-su* gene was then performed codon optimization with the OPTIMIZER device to ensure and increase the likelihood that the protein could be expressed in *E. coli* (strain K12) host cells (Mauro & Chappell, 2014). The results before and after codon optimization can be seen in Table 4 which includes changes in the values of CAI, ENc, and the content of GC and AT.

In Table 4, it can be seen that the CAI value of the *env-su* codon before the optimization is 0.217. This value is classified as low because the ideal CAI value range is 0.8–1.0, so codon optimization needs to be done (GenScript, 2019). The results of the *env-su* codon optimization show that the CAI value increases to 1, this value is categorized as the highest value level. The greater the CAI value, the stronger the use of host cell codons to increase the expression of target genes. If the CAI index in the codon optimization results is 1, it shows that the codon is always used for the synthesis of each coded

amino acid (Gun et al., 2018). The alignment results show that out of 1266 *env-su* nucleotide bases, 937 nucleotide bases match the *E. coli str K-12 substr MG1655* genome, 155 nucleotide bases are changed with their nitrogenous bases (transition), and 174 nucleotide bases are changed differently type of nitrogenous base (transversion).

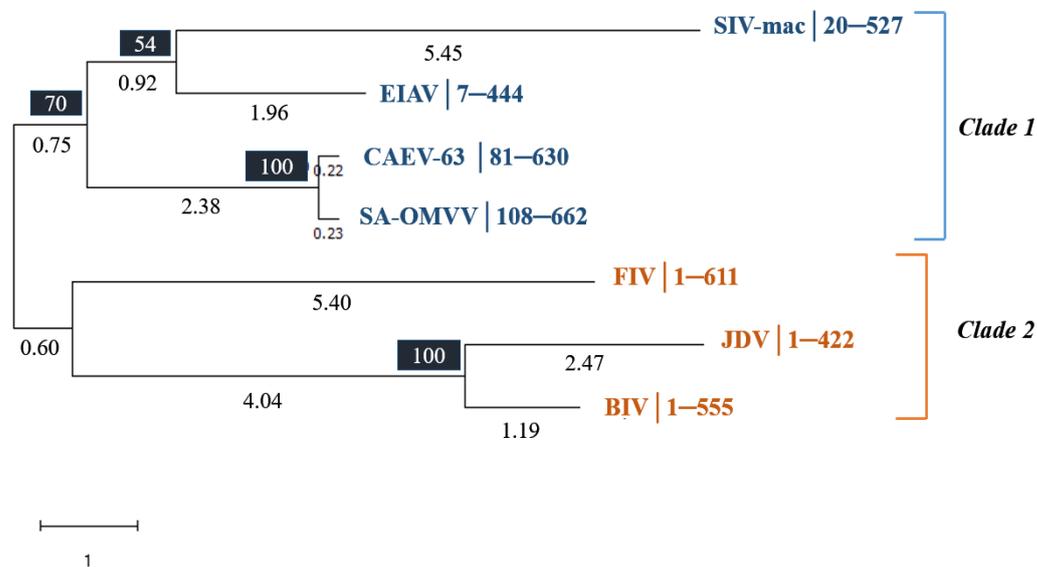
The ENc value listed in Table 4 before the optimization is 51, after optimization it drops to 22.

The index value of  $ENc \leq 35$  means that the preference for the host codon used is high, while the  $ENc$  value  $\geq 50$  indicates that there is still use of random codons so that gene expression is low. Similar to the results of research by Behura & Severson (2012), if the  $ENc$  value approaches 20, it indicates that there is an increase in the use of host codons in a target gene codon.

**Table 2.** Results of the alignment of the JSU protein with other SU proteins of the genus *Lentivirus*

A		Species		Percent Identity (%)	Similar Positions (AA)	Identical Positions (AA)
N		AA Length	AA Length			
1 <sup>st</sup> Species	2 <sup>nd</sup> Species	AA Length	AA Length			
JDV	FIV	422	611	12.78	135	81
JDV	EIAV	422	438	14.72	127	72
JDV	CAEV-63	422	550	10.15	113	66
JDV	SA-OMVV	422	555	11.47	105	75
JDV	SIV-mac	422	508	12.35	114	73
JDV	BIV	422	555	20.57	141	122

Note: A= The total amino acid protein SU



**Figure 1.** SU protein dendrogram based on branch length and bootstrap test (numbers in the column represent the bootstrap value)

The relative degree of adjustment for the use of the host codon in the target gene was determined by looking at the  $ENc$  value. This is by the statement of Uddin (2017), the smaller the  $ENc$  value, the higher the host cell codon used. The results of adjustment for the host cell codon *E. coli* on the *env-su* gene codon are shown in Table 3.

Based on Table 3 it can be seen that each amino acid making up the SU protein can be encoded by one or several different codons (synonym codons). After the codon optimization process was carried out, it was seen that several amino acids that were

initially coded by more than 1 codon became only 1 codon. The amino acid Alanine (A) in the *env-su* gene before optimization was coded by 4 codons, namely GCA, GCC, GCG, and GCT, after optimization, only the GCG codon was selected. Cysteine (C) amino acids were originally encoded by two codons namely TGC and TGT, after codon optimization, Cystein's amino acids were encoded only by the TGC codon.

Modification of the *env-su* gene nucleotide base sequence is carried out by replacing the adenine (A)

and thymine (T) nucleotide bases with guanine (G) and cytosine (C) in a codon without changing the amino acid sequence coding for the translated JSU protein, and equalization is obtained. The content of GC to all parts of the gene (Yulita et al., 2020). The *env-su* codon that had not been optimized had a GC content of 45.9%, and after optimization, the GC content increased to 54.5%. This is by research by Milo (2017), that the GC content in *E. coli str. K-12* ranges above 50.8%, while the ideal percentage of GC content is 30-70% (Mega et al., 2019).

Low GC content can affect the stability of the mRNA structure of the target protein by increasing the heterochromatin conditions in the DNA strand. Heterochromatin is a type of chromatin that makes up DNA that is solid (compact), because the structure of RNA polymerase cannot initiate the transcription and translation process so that the target gene is not expressed (Barahimipor et al., 2015; Ramadhani & Suvifan, 2012), while the GC content which is too high can reduce the level of expression, because it is thought to terminate the transcription process (Silaban et al., 2016). Protein coding genes with ideal GC content in an organism will be translated optimally to increase the expression of the target gene. The AT content before optimization was 54.1%, after optimization it fell to 45.5%. AT-rich codons will be translated inefficiently and can decrease the expression of target genes (Courel et al., 2019; Zhou et al., 2014).

Gene expression that was carried out with codon optimization with un-optimized genes, there was a significant difference in the level of expression of the recombinant protein formed. Codon-optimized genes produced higher recombinant protein than un-optimized genes and with western blotting analysis, protein expression in genes that were not codon-optimized could hardly be detected due to inclusion bodies (Liu et al., 2018).

Inclusion bodies are formed when the expression rate of recombinant protein is high, but not followed by folding of the appropriate protein into active protein and there is a reduction in disulfide bonds between amino acids (Glick et al., 2010). Proteins that are not completely folded will interact with each other, causing insoluble protein aggregates so that they cannot be detected (Maksum et al., 2019). The formation of the inclusion bodies in the form of target protein aggregates shows the inability of the host cell to produce recombinant proteins in their original form (Thenawidjaja et al., 2017).

The results of cutting the restriction endonuclease enzyme via the *GenScript* website shown in Figure 2 show that the two enzymes can recognize the cut regions at the MCS (multiple cloning site) site in the *pET-21a (+)* plasmid. The *EcoRI* restriction enzyme will cut DNA fragments in the sequence 5'-G'AATCC-3' and its complement 3'-CCTAA'G-5' with a length of 159–164bp on the plasmid, while the *HindIII* enzyme cuts through the 5'-A'AGCTT DNA sequence -3' and its complement 3'-TTCGA'A-5' with a length of 1431–1436bp on the plasmid. Cutting with two types of restriction enzymes at separate cutting points on different DNA was carried out simultaneously to produce 2 bands, namely a ribbon measuring 5443 bp (plasmid size *pET-21a (+)*) and a ribbon measuring 1266 pb (size of the *env-su* gene).

Optimization of DNA cutting was carried out using the double-digest endonuclease inhibition enzyme method, namely *EcoRI* and *HindIII*. The double-digest method is used to cut DNA at separate cutting points using two different types of endonuclease inhibition enzymes. This method is often used in directed cloning which will produce two DNA fragments, namely, vector DNA fragments, and target DNA fragments, which can reduce cloning costs and times, and facilitate ligation of different DNA (Shirasawa et al., 2016; Wang et al., 2017). The selection of the restriction enzyme cutting site was based on the presence or absence of the enzyme cutting site to be used on the nucleotide sequence of the *env-su* and plasmid *pET-21a(+)* genes (Priyatno et al., 2019).

The two enzymes are used to cut DNA encoding the target protein and plasmid DNA so that the results of the cutting of the two molecules are complementary so that the target protein DNA can be constructed into the plasmid by forming relatively stable hydrogen bonds and can be linked to the ligase enzyme through phosphodiester bonds (Choiriyah et al., 2013; Mega et al., 2019). The transcription process of *pET-21a (+)* plasmid is controlled by the T7 promoter which will only start transcription after it is initiated by RNA polymerase T7. The T7 promoter is commonly used in expression vectors because it has a high transcription power of target genes and can control gene expression so that it can suppress the toxic effect of proteins that are also expressed by *E. coli* host cells for their survival (Kusumaningsih, 2018).

The plasmid *pET-21a (+)* has been equipped with a lac operon to control the transport and metabolism of lactose in *E. coli* cells. The process of

target gene expression begins with the bonding between the inducer and the lac repressor protein. The bonding complex causes the lac repressor protein to be inactivated and not bind to the lac operon so that it will be able to activate the RNA polymerase enzyme in *E. coli* into T7-polymerization protein which can initiate the transcription and translation process. After the T7 polymerization protein is formed, the protein will bind to the T7 promoter contained in the *pET*-

*21a(+)* plasmid so that the JSU protein encoded by the *env-su* gene will be expressed (Nurjayadi et al., 2019; Yang et al., 2015).

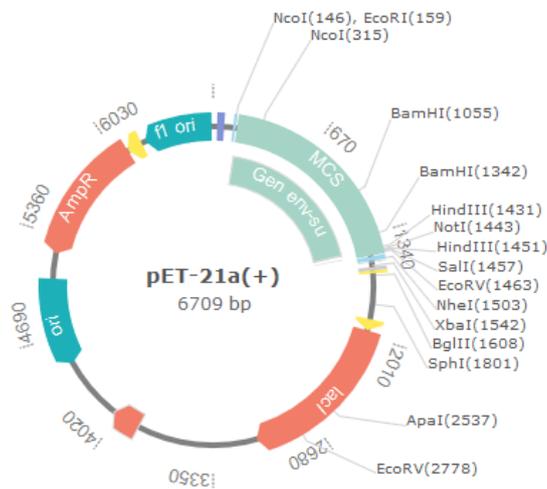
Validation of the results of cutting the endonuclease inhibition enzyme and the next step in producing the JSU recombinant protein from the *env-su* gene in silico requires further research in vitro and in vivo to ensure the safety and potential of the predicted vaccine candidates (Ullah et al., 2020).

**Table 3.** Adjustment for the use of *E. coli* cell codons in the *env-su* gene codon

Codons	Query	Optimized									
GCA(A)	5	0	GCC(A)	1	0	GCG(A)	2	17	GCT(A)	9	0
TGC(C)	7	13	TGT(C)	6	0	GAC(D)	10	21	GAT(D)	11	0
GAA(E)	18	26	GAG(E)	8	0	TTC(F)	7	17	TTT(F)	10	0
GGA(G)	18	0	GGC(G)	3	0	GGG(G)	15	0	GGT(G)	1	37
CAC(H)	6	11	CAT(H)	5	0	ATA(I)	16	0	ATC(I)	4	30
ATT(I)	10	0	AAA(K)	8	17	AAG(K)	9	0	TTA(L)	8	0
TTG(L)	7	0	CTA(L)	10	0	CTC(L)	2	0	CTG(L)	7	36
CTT(L)	2	0	ATG(M)	8	8	AAC(N)	10	21	AAT(N)	11	0
CCA(P)	13	0	CCC(P)	5	0	CCG(P)	2	26	CCT(P)	6	0
CAA(Q)	6	0	CAG(Q)	14	20	AGA(R)	12	0	AGG(R)	6	0
CGA(R)	1	0	CAG(Q)	1	0	CGG(R)	2	0	CGT(R)	3	25
AGC(S)	6	0	CGC(R)	5	0	TCA(S)	4	0	TCC(S)	3	0
TCG(S)	0	0	AGT(S)	0	18	ACA(T)	7	0	ACC(T)	5	22
ACG(T)	2	0	ACT(T)	8	0	GTA(V)	13	0	GTC(V)	4	0
GTG(V)	13	0	GTT(V)	4	34	TGG(W)	14	14	TAC(Y)	6	9
TAT(Y)	3	0	TAA(.)	0	0	TGA(.)	0	0	TAG(.)	0	0

**Table 4.** Optimization results of the *env-su* JDV gene codon with the optimizer

Type	CAI	ENc	%GC	%AT
Before optimization	0.217	51	45.9	54.1
After optimization	1.000	22	54.5	45.5



**Figure 2.** Construction of the *env-su* gene into the plasmid *pET-21a (+)*

Even though the in-silico test results in good predictions, it is not certain that when carried out in vitro and in vivo tests are directly proportional to the in-silico results, because various factors can affect the expected final results of the study.

## CONCLUSION

The JSU protein encoded by the JDV *env-su* gene has the potential as a candidate for the Jembrana disease vaccine based on the results of the percent identity sequence alignment value with the SU protein of other species in the genus *Lentivirus* less than 30%. Codon optimization results showed an increase in the CAI value and GC content and decreased the ENc value and AT content in the JDV *env-su* gene codon, thereby reducing codon bias and increasing the yield of JSU recombinant protein expression in silico.

## ACKNOWLEDGEMENT

Thank you to the Center for Biotechnology Research, the Indonesian Institute of Sciences for allowing and helping to carry out thesis research

## REFERENCES

- Akhsani, F., Hamidy, A., Farajallah, A., & Smith, E. N. (2017). Hubungan Filogenetik *Phrynella pulchra* Boulenger, 1887 berdasarkan Gen 16S rRNA. *Zoo Indonesia*, 26(2), 107–115. Retrieved from [https://e-journal.biologi.lipi.go.id/index.php/zoo\\_indonesia/article/view/3721](https://e-journal.biologi.lipi.go.id/index.php/zoo_indonesia/article/view/3721).
- Ali, M. (2015). Upaya Pengembangan Teknologi Cepat Transkripsi dan Translasi In Vitro dalam Sintesis Vaksin di Indonesia. *Wartazoa*, 25(4), 181–188. Retrieved from <https://onsearch.id/Record/IOS5428.18465>.
- Aryadrie, D. fitria, Santosa, P. E., & Suharyanti, S. (2015). Tingkat Infestasi Cacing Hati pada Sapi Bali di Kecamatan Sukoharjo Kabupaten Pringsewu Provinsi Lampung. *Jurnal Ilmiah Peternakan Terpadu*, 3(3), 134–139. <http://dx.doi.org/10.23960/jipt.v3i3.p%25p>.
- Astiti, N. M. A. G. R. (2018). *Sapi Bali dan Pemasarannya*. Warmadewa University Press.
- Barahimipor, R., Strenkert, D., Neupert, J., Schroda, M., Merchant, S. S., & Bock, R. (2015). Dissecting the Contributions of GC Content and Codon Usage to Gene Expression in the Model Alga *Chlamydomonas reinhardtii*. *Plant Journal*, 84(4), 704–717. doi: 10.1111/tbj.13033.
- Bartak, P., Simek, B., Vaclavek, P., Curn, V., Plodkova, H., Tonka, T., Farková, B., Vernerová, K., & Vejčík, A. (2018). Genetic Characterisation of Small Ruminant Lentiviruses in Sheep and Goats from The Czech Republic. *Acta Veterinaria Brno*, 87(1), 19–26. <https://doi.org/10.2754/avb201887010019>.
- Behura, S. K., & Severson, D. W. (2012). Comparative Analysis of Codon Usage Bias and Codon Context Patterns between Dipteran and Hymenopteran Sequenced Genomes. *Plos One*, 7(8), 1–11. doi: 10.1371/journal.pone.0043111.
- Bu'ulolo, I. C., Simamora, N., Tampubolon, S., & Pinem, A. (2010). Sequence Alignment Menggunakan Algoritma Smith Waterman. *Seminar Nasional Politeknik Batam*, 2(2), 2–7. Retrieved from <https://jurnal.polibatam.ac.id/index.php/JI/article/view/256>.
- Choiriyah, U., Nurjayadi, M., & Dewi, F. K. (2013). Subkloning dan Ekspresi Gen Fim-C S. Typhimurium. *Jurnal Riset Sains Dan Kimia Terapan*, 3(2), 280–291. <https://doi.org/10.21009/JRSKT.032.01>.
- Courel, M., Clément, Y., Bossevain, C., Foretek, D., Cruchez, O. V., Yi, Z., Bénard, M., Benassy, M. N., Kress, M., Vindry, C., Ernoult-Lange, M., Antoniewski, C., Morillon, A., Brest, P., Hubstenberger, A., Crollius, H. R., Standart, N., & Weil, D. (2019). Gc Content Shapes mRNA Storage and Decay in Human Cells. *ELife*, 8, 1–32. Retrieved from <https://elifesciences.org/articles/49708>.
- Daugelaite, J., O' Driscoll, A., & Sleator, R. D. (2013). An Overview of Multiple Sequence Alignments and Cloud Computing in Bioinformatics. *International Scholarly Research Notices*, 2013, 1–14. <https://doi.org/10.1155/2013/615630>.
- Ernawati, E., Puspitaningrum, D., & Pravitasari, A. (2014). Implementasi Algoritma Smith-Waterman Pada Local Alignment Dalam Pencarian Kesamaan Pensejajaran Barisan DNA (Studi Kasus: DNA Tumor Wilms). *Jurnal Pseudocode*, 1(2), 170–177. <https://doi.org/10.33369/pseudocode.1.2.170-177>.
- GenScript. (2019). *Gene Synthesis Handbook (4th ed)*. GenScript USA Inc.
- Glick, B. R., Pasternak, J. ., & Patten, C. (2010). *Molecular Biotechnology: Principles and Applications of Recombinant DNA (4th ed)*. ASM Press.
- Gun, L., Yumiao, R., Haixian, P., & Liang, Z. (2018). Comprehensive Analysis and Comparison on The Codon Usage Pattern of Whole Mycobacterium tuberculosis Coding Genome from Different Area. *Biomed Research International*, 2018(2), 1–8. <https://doi.org/10.1155/2018/3574976>.
- Gustiananda, M. (2011). Immunoinformatics Analysis of H5N1 Proteome for Designing an Epitope-derived Vaccine and Predicting the Prevalence of Pre-existing Cellular-mediated Immunity Toward Bird Flu Virus in Indonesian Population. *Immunome Research*, 7(3), 1–11. Retrieved from <https://www.longdom.org/abstract/immunoinformatics-analysis-of-h5n1-proteome-for-designing-an-epitoped-derived-vaccine-and-predicting-the-prevalence-of-pre-8638.html>.
- Indriawati, I., Margawati, E. T., & Ridwan, M. (2013). Identifikasi Virus Penyakit Jembrana pada Sapi Bali Menggunakan Penanda Molekuler Gen env SU. *Berita Biologi*, 12(2), 211–216. doi: 10.14203/beritabiologi.v12i2.534
- Kusumaningsih, P. (2018). *Evaluasi Konstruksi DNA Da+Lam Vektor Plasmid Berkaitan dengan Ekspresi Protein Rekombinan Rophtry-1 (ROP1) Toxoplasma Gondii pada Eschericia coli*. Universitas Dhyana Pura Bali.
- Kusumawati, A., Tampubolon, I. D., Hendarta, N. Y., Salasia, S. I. O., Wanahari, T. A., Mappakaya, B. A., & Hartati, S. (2015). Use of Reverse Transcription Loop-mediated Isothermal Amplification Combined with Lateral Flow Dipstick for an easy and Rapid Detection of Jembrana Disease Virus. *Virus Disease*,

- 26(3), 189–195. doi: 10.1007/s13337-015-0277-5.
- Kusumawati, A., Wanahari, T. A., Putri, R. F., Mappakaya, B. A., & Tampubolon, I. D. (2014). The Structure and Function of Jembrana Disease Virus Genome. *Journal of Infection and Molecular Biology*, 2(2), 26–29. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC103872/>.
- Lanza, A. M., Curran, K. A., Rey, L. G., & Alper, H. S. (2014). A Condition-Specific Codon Optimization Approach for Improved Heterologous Gene Expression in *Saccharomyces cerevisiae*. *BMC Systems Biology*, 8(1), 1–10.
- Liu, B., Kong, Q., Zhang, D., & Yan, L. (2018). Codon Optimization Significantly Enhanced the Expression of Human 37-kDa iLRP in *Escherichia coli*. *3 Biotech*, 8(4), 1–7. doi: 10.1007/s13205-018-1234-y.
- Maksum, I. P., Utama, E., & Subroto, T. (2017). Extracellular Secretion of Recombinant Human Epidermal Growth Factor by Using Trimethylamine N-Oxide Reductase A (TORA) Signal Peptide in *Escherichia coli* BL21 (DE3). *Journal of Pharmaceutical Sciences and Research*, 9(6), 1007–1016. Retrieved from <https://www.jpsr.pharmainfo.in/Documents/Volumes/vol9Issue06/jpsr09061746.pdf>.
- Maksum, I., Sriwidodo, & Yosua. (2019). *Strategi Peningkatan Ekspresi Protein Rekombinan Secara Intraselular pada Inang Escherichia coli*. Universitas Padjajaran.
- Mardiatmi, M. (2015). *Pedoman Pengendalian dan Penanggulangan Penyakit Jembrana*. Direktorat Kesehatan Hewan.
- Mauro, V. P., & Chappell, S. A. (2014). NIH Public Access. *Trends in Molecular Medicine*, 20(11), 604–613.
- Mega, O., Sumantri, C., Arief, I. I., & Budiman, C. (2019). Ekspresi Gen Lon-like Protease dari *Lactobacillus plantarum* IIA-1A5 pada *Escherichia coli* BL21(DE3). *Jurnal Agripet*, 19(2), 149–158. <https://doi.org/10.17969/agripet.v19i2.14904>.
- Milo, R. (2017). BioNumbers-GC Content of *E. coli* K12 Chromosome. *Nucleic Acids Research*, ID 100528.
- Nurjayadi, M., Afrizal, R., Hardianto, D., & Agustini, K. (2019). Variations of Binding, Washing, and Concentration of Imidazole on Purification of Recombinant Fim-C Protein *Salmonella typhi* with Ni-NTA Resin. *Journal of Physics: Conference Series*, 1402(2019), 1–7. <https://doi.org/10.1088/1742-6596/1402/5/055055>.
- Pangestika, Y., Budiharjo, A., & Kusumaningrum, H. P. (2015). Analisis Filogenetik *Curcuma zedoaria* (Temu Putih) Berdasarkan Gen Internal Transcribed Spacer (ITS). *Jurnal Biologi*, 4(4), 8–13. Retrieved from <https://ejournal3.undip.ac.id/index.php/biologi/article/view/19424>.
- Pearson, W. R. (2013). An Introduction to Sequence Similarity (“Homology”) Searching. *Current Protocols Bioinformatics*, 42(1), 1–9. doi: 10.1002/0471250953.bi0301s42.
- Priyatno, T. P., Winangsih, F., Manzila, I., & Bintang, M. (2019). Ekspresi dan Karakterisasi  $\beta$ -1,3-Glukanase Rekombinan dari Burkholderia cepacia (BiogenCC E76) yang Diekspresikan dalam Sistem Ekspresi *Escherichia coli*. *Jurnal AgroBiogen*, 15(1), 35–44.
- Ramadhani, D., & Suvifan, V. A. (2012). Aplikasi Teknik C-Banding sebagai Biodosimetri Radiasi. *Buletin Alara*, 14(1), 17–22.
- Retnaningati, D. (2017). Hubungan Filogenetik Intraspesies Cucumis melo L. berdasarkan DNA Barcode Gen matK. *Biota*, 2(2), 62–67. <https://doi.org/10.24002/biota.v2i2.1658>.
- Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant Protein Expression in *Escherichia coli*: Advances and Challenges. *Frontiers in Microbiology*, 5(172), 1–17. <https://doi.org/10.3389/fmicb.2014.00172>.
- Shirasawa, K., Hirakawa, H., & Isobe, S. (2016). Analytical Workflow of Double-Digest Restriction Site-Associated DNA Sequencing based on Empirical and In Silico Optimization in Tomato. *DNA Research*, 23(2), 145–153. doi: 10.1093/dnares/dsw004.
- Silaban, S., Maksum, I. P., Enus, S., Hasan, K., Subroto, T., & Soemitro, S. (2016). Kajian Ekspresi Gen Pretrombin-2 Manusia Sintetik pada *Escherichia coli* Secara In Silico Untuk Produksi Trombin sebagai Komponen Lem Fibrin. *Jurnal Pendidikan Kimia*, 8(1), 58–64. <https://doi.org/10.24114/jpkim.v8i1.4425>.
- Silaban, S., Maksum, I. P., Gaffar, S., Enus, S., Hasan, K., Subroto, T., & Soemitro, S. (2017). Desain, Optimasi, dan Kloning Gen Pretrombin-2 sintetik untuk Produksi Trombin sebagai Komponen Lem Fibrin. *Jurnal Ilmiah Ilmu Dan Teknologi Kimia*, 1(1), 69–81. Retrieved from <http://stakc.ac.id/2017/08/12/desain-optimasi-dan-kloning-gen-pretrombin-2-sintetik-untuk-produksi-trombin-sebagai-komponen-lem-fibrin/>.
- Sunarto, A. A. (2015). Perbandingan Program Sequence Alignment. *Jurnal Rekayasa Nusaputra*, 1(1), 1–5. Retrieved from <https://jurnal.nusaputra.ac.id/rekayasa/uploads/paper/55ed4-asril-adi-sunarto.pdf>.
- Susmiarsih, T. P. (2018). Kajian DNA Rekombinan pada Vaksin DNA dan Vaksin Subunit Protein. *Majalah Kesehatan Pharma Medika*, 10(2), 108–128.
- Thenawidjaja, M., Ismaya, W. T., & Retnoningrum, D. S. (2017). *Protein; Serial Biokimia Mudah dan Menggugah*. Grasindo.
- Uddin, A. (2017). Indices of Codon Usage Bias. *Journal of Proteomics & Bioinformatics*, 10(6), 4172. Retrieved from <https://www.longdom.org/open-access/indices-of-codon-usage-bias-jpb-1000e34.pdf>.
- Ullah, M. A., Sarkar, B., & Islam, S. S. (2020). Exploiting the Reverse Vaccinology Approach to Design Novel Subunit Vaccines Against Ebola Virus. *Immunobiology*, 225(3), 1–80. <https://doi.org/10.1016/j.imbio.2020.151949>.
- Wang, Y., Cao, X., Zhao, Y., Fei, J., Hu, X., & Li, N. (2017). Optimized Double-Digest Genotyping by Sequencing (ddGBS) Method with Highdensity SNP Markers and High Genotyping Accuracy for Chickens. *PLoS ONE*, 12(6), 1–19. <https://doi.org/10.1371/journal.pone.0179073>.
- Yang, X., Chen, B., Cai, Y., & Tseng, C. (2015). Understanding The Lac Operon with GeneAct. *International Journal of Computational Biology and Drug Design*, 8(2), 168–188. doi: 10.1504/IJCBDD.2015.071127.
- Yulita, D. S., Polosoro, A., Sisharmini, A., Apriana, A., Nurilmala, F., & Trijatmiko, R. (2020). Konstruksi Vektor Biner dan Transformasi Gen LcCsp Sintetis ke dalam Genom Padi Nipponbare dengan Bantuan *Agrobacterium tumefaciens*. *Jurnal AgroBiogen*, 16(1), 25–34. Retrieved from <http://ejournal.litbang.pertanian.go.id>

Zhou, H., Ning, L., Zhang, H., & Guo, F. (2014). Analysis of the Relationship between Genomic GC Content and Patterns of Base Usage , Codon Usage and Amino Acid Usage in Prokaryotes: Similar GC Content Adopts Similar Compositional Frequencies

Regardless of the Phylogenetic Lineages. *Plos One*, 9(9), 1–7.  
<https://doi.org/10.1371/journal.pone.0107319>.