The pET-Rep recombinant plasmid construction and
geminivirus Rep [C1] gene expression in Escherichia coli
strain BL21

E N Gozalia1, L Syukriani2, R Renfiyeni3, and J Jamsari1,2*

1 Department of Biotechnology, Postgraduate Program, Andalas University, Padang, West
Sumatera, 25163 Indonesia

2 Department of Agrotechnology, Agriculture Faculty, Andalas University, Padang, West
Sumatra, 25163 Indonesia

3 Faculty of Agriculture, University of Muhammad Yamin, 27317 Solok, West Sumatra,
Indonesia

E-mail: ajamsari@yahoo.com

Abstract. Geminivirus Rep [C1] gene produces a replication protein [Rep] that interacts with
Retinoblastoma-related protein [RBR] when Geminivirus infect the plants. That interactions
interfere with RBR and transcription factor function. So the mechanism of chili plant defense
against symptoms of Pepper Yellow Leaf Curl Disease is blocked. The NPR1 gene is a
transcriptional co-activator involved in the regulation defense mechanism of Systemic
Acquired Resistance [SAR] in chili plants. In silico, Rep protein blocks the SAR mechanism
that is regulated by the NPR1 gene. This study aimed to obtain a construction of pET-28a+
plasmid expression recombinant Rep [C1] gene and Rep protein that can be expressed in E.coli
BL21. This research is useful to study the interaction of Geminivirus with transcription co-
activator in chili plants. The construction was done by ligating plasmid and Rep genes which
have been cut using restriction enzymes BamHI and SacI. The construction of a recombinant
pET-Rep plasmid is confirmed by amplification and verification of the nucleotide sequence
through sequencing techniques. Nucleotide sequence verification results proved that the
Rep [C1] gene is successfully constructed into a pET-28a+ plasmid with proper position and
orientation. The pET-Rep recombinant proceeds to the expression test using the E.coli BL21
host. Expression was performed using the induction method with IPTG. Proteins Rep is
purified using the MagnehisTM Protein Purification System. Rep protein is visualized with
SDS-PAGE. Visualization showed that Rep protein was successfully expressed in E.coli BL21
where protein size is estimated of 41,03 kDa in size.

1. Introduction
Currently, pepper plants still have problems with yellow leaf curl disease. The disease is an epidemic
disease in pepper plants in West Sumatra, Indonesia. One of the major geminiviruses infecting many
plants in Indonesia is Pepper Yellow Leaf Curl Disease [PepYLCV] [Begomovirus]. The virus has
some components of functional protein-encoding genes designated as a coat of protein [CP],
movement protein [MP], Replication [C1] and satellite beta [β] [1].

One component of the functional protein-coding gene that plays a major role in the spread of virus
particles is the Rep [C1] gene. The Rep [C1] gene is a replication and transcription regulatory element
in geminivirus. The gene functions at the beginning of replication and initiates DNA replication [2].
The Rep [C1] gene has multifunctional characters including in the mechanism of rolling circle replication [RCR] and interacting with retinoblastoma proteins [3]. The Retinoblastoma-related protein [RBR] plays a significant role in cell regulation [4]. RBR suppresses rapid cell proliferation or development in adult tissues [5]. The Rep proteins that interact with RBR will interfere with some transcription factors. Interaction of Rep protein with RBR releases the transcription factor so that it activates the expression of Proliferating Cell Nuclear Antigen [PCNA] to produce DNA polymerase and the factors needed for viral replication [6].

Pepper plants have a Systemic Acquired Resistance [SAR] defense response. One of the transcription coactivators in pepper plants is the Non-expressor of Pathogenesis Related [NPR1]. The gene involved in defense mechanisms SAR in plants [7]. The NPR1 protein is the main regulator in the expression of Pathogen Related [PR] genes mediated by salicylic acid in SAR [8]. The NPR1 is also involved in basal resistance or when infected, Induced Systemic Resistance [ISR] and Pto mediated resistance [9]. The NPR1 gene expresses the PRI gene in response to pathogenic infections. Structurally, the NPR1 consists of BTB/POZ, ankyrin and transactivation domains [10]. Molecular interaction between the NPR1 gene and Rep protein in response to Geminivirus attacks on pepper plants has not been proven so far. Binding of Rep protein to the NPR1 gene will cause suppression of NPR1 gene function. So the NPR1 gene could not stimulate expression of the Pathogen Related 1 [PRI] resistance gene. Therefore, it is necessary to elucidate the mechanism of NPR1 gene interaction with Rep protein for yellow leaf curl disease regulation in pepper plants [11].

Several steps to find out such the above mechanism has to involve the recombinant plasmid construction of the Rep [C1] gene and its downstream expression. Recombinant plasmid construction could be done to produce Rep proteins that could be used in the study of interactions with NPR1 gene transcriptional coactivators. The plasmid commonly used as the expression vector for such a goal is a pET-28a+ plasmid. This plasmid belongs to the expression vector that is facilitated with His-tag in the upstream and downstream of the inserted target gene. The presence of His-tag facilitates the phase of protein purification using E.coli strain BL21 as its host. E.coli BL21 has high expression capability and T7 RNA polymerase corresponding with the T7 promoter on the pET-28a+ plasmid.

This study was aimed to construct the pET-28a+ plasmid recombinant containing Rep [C1] gene sequence. So, the Rep protein can be expressed in E.coli BL21.

2. Materials and methods

The main ingredients used in this research activity are 50 ng DNA, Rep [C1] gene, Kapa 2G, E. coli DH5α strain BL21, plasmid pET-28a+, Thermo scientific enzyme BamHI and SacI, tango buffer, 10x buffer ligation kit [Novagen], gel purification kit [Thermo scientific], plasmid purification kit [ATP™, Biotech Ink], IPTG, kanamycin antibiotic, CaCl₂, Coomassie Blue, 1x SDS running buffer, Magne-HisTM, Reagent Bradford dyes, protein ladder, and 1 KB DNA ladder. Some specific primers used in this study can be seen in Table 1.
Table 1. Primer data used in the study

| Primer Name                  | Sequence                                      |
|-----------------------------|-----------------------------------------------|
| T7                          | 5'-TAATACGACTCACTATAGGGCGA-3'                 |
| SP6                         | 5'-ATTTAGGTGACACTATAGAATAC-3'                |
| T7 Promotor Forward         | 5'-TAATACGACTCACTATAGGG-3'                   |
| T7 Terminator Reverse       | 5'-GCTAGTTATGCTCAGCGG-3'                     |
| C1-TD21-SmaINT             | 5'-CTAATCCCGGTCAGTCTCCTGCGA-3'               |
| C1-TD21-BamHINT             | 5'-CATGGGGGATCCATGCCTCCACCACGT-3'            |

2.1. *In silico* plasmid construction

The plasmid construction was carried out using the plasmid ApE editor win 2.0.36 application. The *in silico* construction was carried out according to the manual provided in the website. The task was started by downloading the *Rep* [C1] gene sequence available in the database https://www.ncbi.nlm.nih.gov/. The pET-28a+ plasmid sequences were retrieved from https://www.snapgene.com. The plasmid was cut with *SacI* and *BamHI* enzymes, and the short sequences were removed. The end of the *Rep* [C1] gene sequence in the near of the start codon was cut with the *BamHI* enzyme, while another end it is cut with the *SacI* enzyme.

2.2. *Rep* [C1] gene preparation

The *Rep* gene which is available in the *pGEM-T Easy* plasmid was verified using T7SP6 primer. Besides, verification was also performed using the *C1-TD21-SmaIBamHINT*. The verified pGEM-Rep stock is then transformed into *E. coli DH5α*.

2.3. pET-28a+ plasmid preparation

Plasmid pET-28a+ was obtained from laboratory stock which has been propagated in *E. coli DH5α*. The isolation process is carried out using ATP™ Plasmid Mini Kit [Biotech, USA].

2.4. *Rep* gene [C1] and pET-28a+ plasmid restriction

The restriction of the *Rep* [C1] gene from the pGEM-Rep and pET-28a+ plasmid was done to cut out the insert. The restriction step was done twice using different restriction enzymes. The first restriction was carried out using the *BamHI* Fast Digest enzyme, while the second restriction was carried out using *SacI*. The 30 µL of cocktail composed of 17 µL nuclease-free water, 10 µL pGEM-Rep PCR and pET-28a+ plasmid, 2 µL FD10x buffer and 1 µL *BamHI* Fast Digest restriction enzyme was prepared. Amplification was run under the condition as presented in Table 2.

Table 2. PCR conditions for restriction with *BamHI*

| Reaction        | Temperature [°C] | Duration [minute] |
|-----------------|-----------------|-------------------|
| Enzyme activation | 37              | 15                |
| Enzyme inactivation | 80           | 5                  |
| Pause            | 8               | ~                  |

After the first restriction, the restricted sequence was purified using isopropanol. Isopropanol was added to the restriction reaction using a 1:1 ratio. Incubated was done at -20°C for 1 hour and subsequently centrifuged for 10 minutes twice at 10,000 rpm. The supernatant was removed while the pellets were washed with 70% ethanol and finally was dried. The pellet is dissolved with nuclease-free
water before restricted with the \textit{Sac}I enzyme. The restriction was run using a PCR machine with the program presented in Table 3.

\textbf{Table 3. PCR conditions for \textit{Sac}I restriction}

| Reaction               | Temperature [$^\circ$C] | Duration [minute] |
|------------------------|-------------------------|------------------|
| \textit{Enzyme activation} | 37                      | 120              |
| \textit{Enzyme inactivation} | 80                     | 20               |
| \textit{Pause}         | 8                      | ~                |

The restriction of DNA was electrophoresed for 30 minutes with a voltage of 100 volts and visualized using gel documentation. The restricted \textit{Rep} [\textit{C1}] gene and \textit{pET-28a}+ plasmid were further purified using GeneJET Gel Extraction Kit [Thermo scientific].

2.5. \textit{Ligation of \textit{Rep} [\textit{C1}] gene with plasmid \textit{pET-28a}+}

The ligation activity was carried out based on the protocol provided by Novagen [USA] using a 10x ligation buffer kit.

2.6. \textit{Transformation of \textit{pET}-\textit{Rep} into \textit{E. coli BL21}}

Transformation of \textit{pET}-\textit{Rep} into \textit{E. coli} BL21 cell was done using the procedure described by Promega [Promega-USA] which was based using the heat shock technique.

2.7. \textit{Induction of \textit{pET}-\textit{Rep plasmid}}

A single colony of \textit{E. coli} BL21 was inoculated into 50 mL liquid LB media containing kanamycin 30 \textmu g/mL antibiotics. The bacteria were propagated under agitation at a speed of 160 rpm at 37\degree C until its OD [600] reached 0.6-1.0. The culture was stored at 4\degree C for overnight. Expression induction was used by adding \textit{isopropyl \beta-D-1-thiogalactoside} [IPTG] to a final concentration of 0.4 mM and followed by incubation for 2-3 hours at a temperature of 25\degree C - 30\degree C.

2.8. \textit{Purification \textit{Rep} protein}

Purification of \textit{Rep} protein was done using the MagneHis\textsuperscript{TM} Protein Purification System [Promega, USA]. The purified protein was visualized using SDS-PAGE.

3. \textit{Results and Discussion}

3.1. \textit{Construction of recombinant plasmids in silico}

Figure 1 shows the nucleotide sequences of the \textit{pET-28a}+ and the \textit{Rep} [\textit{C1}] gene in silico. The \textit{Rep} [\textit{C1}] gene and \textit{pET-28a}+ plasmid were digested using \textit{Bam}HI and \textit{Sac}I enzymes. The \textit{Bam}HI restriction site is adjacent to the codon start position [ATG], while the restriction site of \textit{Sac}I is adjacent to the stop codon position [TAA]. On the plasmid, the \textit{Bam}HI restriction site is adjacent to the T7 region of the promoter and the restriction side of \textit{Sac}I adjacent to the T7 terminator region.
Figure 1. Results of analysis in silico of pET-28a+ plasmid+ Rep [C1] gene. Grey color is the sequence of the Rep [C1] gene while red color is the restriction site of the enzyme used and the yellow is the sequence position of the T7 promoter and T7 terminator primers. ATG is start codon and TAA is the stop codon.

3.2. Transformation of pET-Rep into E.coli BL21

Figure 2. The results of plasmid restriction pET-28a+ and Rep [C1] genes. 1KB: DNA ladder marker. [A] Rep [C1] gene before restriction. [B] Rep [C1] gene after restriction. [C] DNA plasmid pET-28a+ before restriction. [D] DNA plasmid pET-28a+ after restriction.

The plasmid DNA of pET-28a+ and the amplified Rep [C1] sequence were successfully digested. Figure 2 shows the plasmid conformation before circular-shaped restriction resulting in two bands.
Restriction formed a single band indicating that the plasmid had been digested. Plasmid size before the restriction is 5369 bp and after the restriction is 5361 bp. The Rep [C1] gene size before the restriction was 1262 bp, while the size after being restricted was 1184 bp. Thus, ligation of both fragments should produce plasmid with 6545 bp in size.

The success of ligation between pET-28a+ and Rep gene sequence was verified using the PCR technique. Figure 3 indicates that the ligation between pET-28a+ and Rep [C1] gene did occur. Using primer T7_PFTR, PCR analysis could produce a single fragment with 1496 bp in size [B-C], while amplification using primer C1 Sma/BamHINT produced a single fragment with 1120 bp in size.

The success ligation between plasmid and target genes could be influenced by several factors. The ends of the plasmid and target gene must be compatible, therefore the same restriction enzyme should be used in the restriction process. The concentration and composition of the ligase and buffer enzymes in the cocktail must be appropriate. Furthermore, the ratio of plasmids compared to insert genes must be carefully measured. This generally is determined by the concentration and length of the nucleotides of both plasmid and target genes. Some external factors such as the incubation temperature during the ligation process could influence the success rate of ligation [12].

Further verification was performed using the PCR colony technique. The result is presented in Figure 4.

![Figure 4](image)

**Figure 4.** Product of PCR colony generated from BL21 transformant using T7_PFTR primer [a] C1 Sma/BamHINT primer [b].

Figure 4a shows that colony A produced some bands while colony C produced one single band. Colony B and colony D, unfortunately, produced no products. Based on that result colony A was re-amplified with specific primer C1 and finally exhibited a single fragment of 1120 bp in size [Figure 4b]. The positive colony was processed to further sequence analysis.

### 3.3. pET-Rep sequence analysis

Sequence analysis was performed in order to verify the orientation of the inserted gene target. Sequence data were further analyzed using Vecscreen software and the result is presented in Figure 5. Since reverse primer produced low-quality sequence data, the analysis was performed from the forward primer. Figure 6 shows that the codon start of the Rep [C1] gene is located on the downstream position of the T7 primer sequence. Meaning that its orientation is correct, theoretically enabling transcription initiation.
Further analysis was carried out to verify the integrity of the sequence. The nucleotide sequence in Figure 6 proves that the orientation of the Rep [C1] gene is correct. This fact is based on that the nucleotide sequence of the T7 promoter [TCGCATGCTCCAGCGCCCAT] followed by the restriction site of the BamHI enzyme is located between gene insert and start codon. Based on this it can be ascertained that the orientation of the Rep [C1] gene is correct. Moreover, to verify the identity of the sequence, BLAST and Clustal analysis were run. BLAST data using 692 nucleotides against nucleotide available in the NCBI database exhibited 99% identity with Rep [C1] nucleotide sequence isolated from PepYLCV strain Tanah Datar [TDWS]. Two base substitutions were found after comparing both nucleotide sequence data [see Figure 7].
The two nucleotides undergo substitution from base A to G in the position of 55th and 273rd. Despite we found two substitution events, the analysis was continued by comparing the amino acid sequence. The result is presented in Figure 8.

**Figure 8.** Amino acid sequence alignment between inserted Rep [C1] gene with the Rep [C1] gene isolated from PepYLCV_TDWS-21: amino acid variation.

Alignment of both amino acid sequences exhibited that only 1 nucleotide substitution [base A to G in the 55th] causes a change in amino acid composition from N [asparagine: asp] to D [aspartic acid: asn], while nucleotide substitution at position 273rd does not change any amino acid variation. However, asparagine is an analog of aspartic acid [13] so that the structure and function of Rep proteins could be almost similar. Analysis of protein structure using ProFunc software [http://www.ebi.ac.uk/thornton-srv/database/ProFunc] can be seen in Figure 9-10. Analysis of the secondary structure of the amino acid sequences of the two Rep proteins shows similar shape and arrangement of their α-helices and β sheets.

**Figure 9.** Secondary structure of the Rep [C1] gene from PepYLCV_TDWS-21 [A] and Rep [C1] gene inserted in pET-Rep.

A similar result is shown also from tertiary structure analysis [Fig 10]. They have also similar conformation shapes at each rotational degree of their structure. Based on this data, we assume that both proteins are analog and should have similar activity [14].
Figure 10. Tertiary structure comparison between Rep [C1] isolated from PepYLCV_TDWS-21 and the Rep [C1] inserted to the pET-Rep plasmid.

3.4. Induction of expression
To test, whether the inserted gene could be expressed or not, expression analysis was performed. The Rep [C1] gene expression was induced with 0.4 mM IPTG for 3 hours. Induction was performed at temperatures ranging from 25 to 30 °C [12]. Since the Rep protein doesn’t have a signal peptide domain, we assumed that Rep protein will be produced as intracellular protein. For that reason, the purification of Rep protein was performed using His-tagged approach [Figure 11].

Figure 11. SDS PAGE analysis of protein Rep. M: Protein Ladder. W1-W2: Protein trapped in the first and second washing steps. E: Protein Rep. N: Rep protein which is still bound to Ni particles.

SDS-PAGE analysis shows that Rep protein bands were larger than 40 kDa based on the unstained protein ladder. The protein produced from the expression is corresponding with the expected size which is about 41.03 kDa [Figure 11]. The expectation was performed using Promega Biomath Calculator software which is available online.

4. Conclusion
The Rep [C1] gene was successfully constructed into the plasmid expression pET-28a+ with the correct orientation. The gene was inserted between the SacI and BamHI cloning sites. Expression analysis showed that Rep protein was successfully expressed in the cell of E. coli BL21 conferring 41.03 kDa in size.
References

[1] Jamsari J and J Pedri. 2013. Complete nucleotide sequence of DNA A-like genome dan DNA of monopartite Pepper Yellow Leaf Curl Virus, a dominant begomovirus infecting Capsicum annum in West Sumatera Indonesia. Asian Journal of Plant Pathology. 5 1-14.

[2] Jamsari I, I Suliansyah, and I Manti, 2009. Gen-Gen Penentu Virulensi Virus Gemini dan Transformasi Genetik untuk Menghasilkan Tanaman Cabai Merah Tahan Penyakit Virus Kuning Keriting. Research Report. Padang.

[3] Laufs J, S Schumacher, N Geisler, I Junipin, and B Gronenborn. 1995. Identification of the nicking tyrosine of geminivirus Rep protein. FEBS Letters. 377. 258–262.

[4] Hull, R. 2002. Virus Replication. Plant Virology. 4. 293-372.

[5] Desvoyes B, M A De, T I Ruiz, and C Gutierrez, C. 2014. Novel roles of plant RETINOBLASTOMA-RELATED (RBR) protein in cell proliferation dan asymmetric cell division. Journal of Experimental Botany. 10. 2657-66.

[6] Egelkrou E M, L Mariconti, S B Settlage, R Cell, D Robertson, and L Hanley-Bowdoin, 2002. Two E2F elements regulate the proliferating cell nuclear antigen promoter differently during leaf development. The Plant Cell. 12. 3225-3236.

[7] Hanley-Bowdoin L, E R Bejarano, D Robertson, and S Mansoor, 2013. Geminiviruses: masters at redirecting dan Reprogramming plant processes. Nature Reviews. Microbiology. 11. 77-88.

[8] Nova B, S Syukur, and J Jamsari. 2017. Molecular Cloning dan Characterization of Npr1 Ankyrin Domain from Capsicum annum L. Der Pharmacia Lettre. 3. 148-55.

[9] Mou Z, W Fan, and X Dong, 2003. Inducers of plant systemic acquired resistance Regulate NPR1 function through redox changes. Cell. 113. 935-44.

[10] Dong X. 2004. NPR1, all things considered. Current Opinion in Plant Biology. 5. 547-52.

[11] Kuai X, B J MacLeod, and C Després, 2015. Integrating data on the Arabidopsis NPR1/NPR3/NPR4 salicylic acid receptors; a differentiating argument. Frontiers in Plant Science. 6. 1-5.

[12] Novagen. 2003. pET System Manual Novagen pET System Manual Novagen. 10th ed. Novagen. 68.

[13] Murphy M J and F J Stubbins, 1980. Synthesis and anticancer activity of asparagine analogs. Journal of Pharmaceutical Sciences. 69 553-55.

[14] Martin Y C, J I Kofron, and L M Traphagen, 2002. Do structurally similar molecules have similar biological activity?. Journal Medicinal Chemistry. 45. 4350-58.

Acknowledgments
This research was financially supported by the General Directorate Higher Education of the Republic of Indonesia through Professorship Cluster Research Grants of Andalas University contract number 43/UN.16.17/PP.HGB/LPPM/2017; 18/UN.16.17/PP.RGB/LPPM/2018 and T/22/UN.16.17/PP.KP-KRP1GB/LPPM/2019.