DOI: 10.31276/VJSTE.64(4).60-63

In-vitro synthesis of RNA fragments specifying envelope protein gene and *RdRp* gene of SARS-CoV-2 as positive standards for molecular diagnosis tests

Hang Thi Thu Dinh[1*], Mai Ngoc Nguyen[1, 2], Su Xuan Hoang[1]

[1]*Institute of Biomedicine and Pharmacy, Vietnam Military Medical University*
[2]*University of Science, Vietnam National University, Hanoi*

Received 22 July 2021; accepted 19 October 2021

Abstract:

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of Coronavirus Disease 2019 (COVID-19), has rapidly spread through the entire world and has become the worst pandemic from December 2019 until now. The establishment of positive standards for molecular diagnostic testing for SARS-CoV-2 plays a critical role in development and assessment of diagnostic assays associated with the shortage of positive specimens and viral culture fluids. This study aims to establish a novel assay for *in-vitro* RNA fragment synthesis based on amplicons of self-priming PCR targeting envelope protein gene and RNA-dependent RNA polymerase (*RdRp*) gene of SARS-CoV-2. The cDNA library of the targeted genes of SARS-CoV-2 was generated by using long-primers named sMn1 forward/reverse primer (*E* gene) and sMn2 forward/reverse primer (*RdRp* gene) for self-priming PCR assays. The synthesised amplicons that overlap the target sequence of the World Health Organization (WHO) assay were cloned into a pGEM-T easy vector, then transformed into *E. coli* competent cells by conventional methods. The recombinant plasmids were used as materials for *in-vitro* RNA transcription. Concentrations of the *in-vitro* transcribed RNA were 200-800 ng/µl with A260/A280 ratios of 2.0-2.2. Gel electrophoresis showed a single band of each RNA molecule with sizes of 216 and 214 bases for sMn1- E and sMn2- *RdRp* gene, respectively. Furthermore, we effectively evaluated the *in-vitro* transcribed RNA by a one-step, real-time RT-PCR assay according to the standard WHO protocol. The stability of in-vitro RNA over a 6-month storage period was then investigated. In conclusion, our assay for in-vitro synthesis of RNA fragments transcribed from self-priming amplicons were successfully established and thus these positive standards were useful for molecular diagnosing of SARS-CoV-2.

Keywords: ***in-vitro*** **transcribed RNA, one-step real-time RT-PCR, SARS-CoV-2, self-priming.**

Classification number: **3.2**

Introduction

COVID-19 is a severe acute respiratory syndrome (SARS) caused by a novel strain of coronavirus named SARS-CoV-2. This virus has been reported to be a highly transmissible and pathogenic coronavirus that arose in late 2019 from Wuhan, China [1] and has sparked a pandemic, posing a significant threat to human health, public safety, and many other aspects [2]. To date, more than 163 million cases have been confirmed with more than 3.37 million deaths related to COVID-19 worldwide as of May 16th, 2021 [3].

Similar to other coronaviruses, the genome of SARS-CoV-2 consists of a positive-sense single-stranded RNA molecule ranging from 26 to 32 kb in size, which was believed to have emerged from bats. However, there is no strong evidence to ensure its zoonotic origin [4]. Coronavirus particles are spherical in shape with diameters between 80 and 160 nm. The genetic structure of SARS-CoV-2 contains 4 structural proteins and 16 non-structural proteins with the arrangement of open reading frames (ORFs) as replicase and protease (1a-1b) and other major glycoproteins like Spike (S), Envelope (E), Membrane (M), and Nucleocapsid (N) proteins following a typical 5'-3' order [5]. Among these, the spike protein (S) coats the envelope surface and is responsible for virus penetration into host cells while membrane (M) and envelope (E) proteins cover the membrane. Also, in the SARS-CoV-2 genome, there is an essential gene for virus replication called the *RdRp* (RNA-dependent RNA polymerase) gene in the ORF1ab sequence, which is highly conserved and specific for this virus strain [6]. SARS-CoV-2 possesses many virulence factors that play an important role in the invasion, escape, and multiplication in host cells. This virus has the ability to stay in the human body for a very long time until it causes illness in the respiratory, digestive, and nervous systems of humans and other animals [7].

In order to manage COVID-19, the first step requires a rapid and accurate detection of SARS-CoV-2. Remarkably, the WHO has recommended reverse transcription-polymerase chain reactions (RT-PCR) targeting the *E* gene and *RdRp* gene for the confirmation of SARS-CoV-2 diagnosis. The establishment of positive standard

[*]*Corresponding author: E-mail: dinhhang@vmmu.edu.vn*

sources play a critical role in development and assessment of diagnostic kits as patients' specimens are not available for every laboratory. Besides, for virus culturing, these samples require handling in biosafety level 3 laboratories. Therefore, plasmids or RNA containing the target sequence of SARS-CoV-2 need to be prepared. Therefore, this study focused on the establishment of a novel assay for *in-vitro* transcribed RNA as positive standards based on amplicons of the self-priming PCR targeting envelope protein gene and *RdRp* gene of SARS-CoV-2.

Materials and methods

Envelope protein gene and RdRp gene cloning of SARS-CoV-2 in E. coli

E gene and RdRp gene construction and amplification: Using primer design software, the specific primers sMn1-Fwd/Rev (*E* gene) and sMn2-Fwd/Rev (*RdRp* gene) (Nucleotide position 26248 to 26366 and 15422 to 15538, GenBank accession no. NC_045512.2, respectively) were designed to amplify the target nucleotide sequence of the *E* and *RdRp* genes from SARS-CoV-2 (Table 1). Self-priming refers to the folding back of oligonucleotides. Given the advantage of self-priming that allows amplification initiation without the need for a DNA template, the forward and reverse primer were designed to interact with each other to form the double-stranded DNA fragments of interest. The self-priming PCR process consisted of denaturation (95°C/5 min), annealing and amplifying (94°C/30 s, 40°C/30 s, 72°C/40 s) for 35 cycles, followed by final extension step at 72°C/7 min. The PCR products (119 and 117 bp in length) were analysed by electrophoresis in a 1.2% agarose gel with ethidium bromide then purified by a PCR purification kit (Thermo Scientific, USA).

Table 1. Primers and probes used for this study.

No	Primer/ probe	Sequence (5'-3')	Target gene	Reference
1	sMn1-F	GAGACAGGTACGTTAATAGTTAATAGCGTACTTCTT TTTCTTGCTTTCGTGGTATTCTTGCTAGTTACAC	Envelope	This study
2	sMn1-R	ACAATATTGCAGCAGTACGCACACAATCGAAGCGC AGTAAGGATGGCTAGTGTAACTAGCAAGAATACCA	Envelope	This study
3	sMn2-F	AATATTGAGTGAAATGGTCATGTGTGGCGGTTCACT ATATGTTAAACCAGGTGGAACCTCATCAGGAGAT	*RdRp*	This study
4	sMn2-R	CAGCTGACAGCTTGACAAATGTTAAAAACACTATTAG CATAAGCAGTTGTGGCATCTCCTGATGAGGTTC	*RdRp*	This study
5	qPrimer1	ACAGGTACGTTAATAGTTAATAGCGT	Envelope	[8]
6	qPrimer2	ATATTGCAGCAGTACGCACACA	Envelope	[8]
7	nCoV-F	GTGAAATGGTCATGTGTGGCGG	*RdRp*	[9]
8	nCoV-R	CAAATGTTAAAAACACTATTAGCATA	*RdRp*	[9]
9	nCoVE_pr	/56-FAM/ACACTAGCCATCCTTACTGCGCTTCG /3BHQ_1/	Envelope	[8]
10	nCoV_pr	/56-FAM/ CAGGTGGAACCTCATCAGGAGATGC /3BHQ-1/	*RdRp*	[9]

Cloning of E gene and RdRp gene in E. coli DH5α: The purified PCR products were inserted directly into pGEM-T easy vector (Promega, USA) (Fig. 1) and transformed into *E. coli* DH5α competent cells by heat shock at 42oC/90 s. Then, recombinant plasmids were tested by a *SspI* restriction enzyme (Thermo Scientific, USA). The targeted sequences were then identified by Sanger method, and MegaX software was used to compare and analyse against other determined sequences on GenBank.

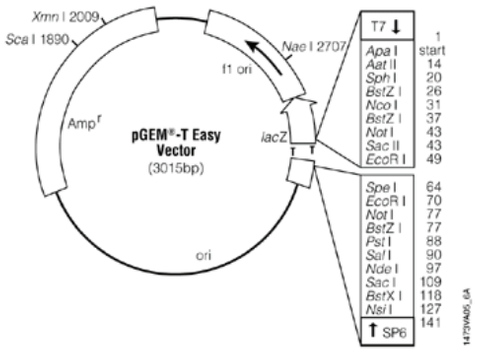

Fig. 1. The pGEM-T easy vector (pGEM-T easy vector systems, Promega, USA).

In-vitro synthesisation of E protein gene and RdRp gene of SARS-CoV-2

After analysis, the recombinant plasmid containing the sMn1/sMn2 sequence was cleaved at one point downstream of the T7 promoter and the segment to be transcribed using *NdeI* restriction enzyme creating a linear strand. The RNA product was then purified (PCR purification kit - Thermo Scientific, USA) for RNA *in-vitro* synthesis.

Synthesisation of the *E* and *RdRp* genes of SARS-CoV-2: Reaction components were 5X TranscriptAid Reaction buffer 4 µl; NTP mix (25 mM) 8 µl; TranscriptAid Enzyme Mix 2 µl (Thermo Scientific, USA); and DNA plasmid 6 µl; DEPC - treated water added up to the total volume of 20 µl. RNA *in-vitro* was synthesised at 37oC in 2 h. DNAseI was used to remove any remaining DNA at 37°C in 1.5 h. The RNA product was purified by QIAGEN RNeasy Mini kit (Qiagen, Germany) following the conventional protocol. Purified RNA were stored at -80°C for any further experiment (Fig. 2).

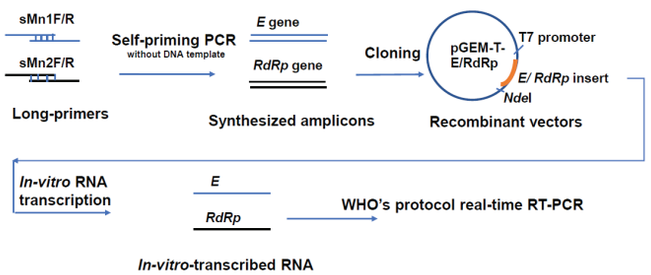

Fig. 2. Diagram of the experimental procedure for *in-vitro* transcribed RNA as positive standards based on amplicons of self-priming PCR targeting the *E* gene and *RdRp* gene of SARS-CoV-2.

A NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, USA) was used to quantify and assess the purity of the purified RNA. Also, their molecular size was visualised by conducting one-step RT-PCR (Qiagen, Germany) on 1.2% agarose gel electrophoresis. Following up with the WHO protocol

for one-step, real-time RT-PCR using Superscript III one-step RT-PCR system with Platinum *Taq* Polymerase (Thermo Scientific, USA), the primer qPrimer(1+2), probe nCoVE-pr and primer nCoV Fwd/Rev, probe nCoV_pr were used for the reaction of *in-vitro* synthesised RNAs sMn1 and sMn2, respectively (Table 1) [8]. After analysis by Sanger sequencing, the *in-vitro* transcribed RNAs were stored at -80oC to determine their stability after 1, 3, and 6 months using one-step real-time RT-PCR assays.

Results and discussion

Cloning envelope protein gene and RdRp gene of SARS-CoV-2 in E. coli

Observation from gel electrophoresis results of PCR products of primer sMn1 Fwd/Rev and sMn2 Fwd/Rev, the synthesising process succeeded with significant bands corresponding to sizes of 119 bp and 117 bp each (Fig. 3A).

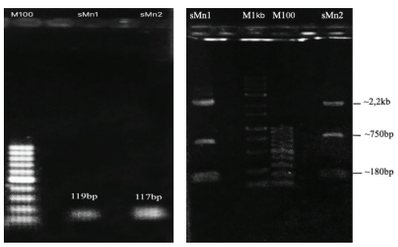

Fig. 3. (A) Self-priming PCR products of sMn1 (*E* gene) and sMn2 (*RdRp* gene) of SARS-CoV-2; (B) Recombinant plasmids of sMn1 and sMn2 after being cleaved by *Ssp* I restriction enzyme. Lane M100, 100 bp DNA ladder (Thermo Scientific, USA); Lane M1kb, 1kb DNA ladder (Thermo Scientific, USA).

As the SARS-CoV-2 genome contains positive-sense single-stranded RNA, forward and reverse primers were designed to create the gene fragment of interest for each envelope protein gene and *RdRp* gene. sMn1 (*E* gene) was synthesised using sMn1 Fwd/Rev primers and, similarly, sMn2 Fwd/Rev for sMn2 (*RdRp* gene).

The pGEM-T easy vector was chosen for its advantages in the cloning process such as its small size of 3015 bp and numerous restriction sites within the multiple cloning regions. Also, these are high-copy-number vectors in *E. coli* containing T7 and SP6 RNA polymerase promoters flanking a multiple cloning region within the α-peptide coding region of the enzyme β-galactosidase. Insertional inactivation of the α-peptide allows the identification of recombinants by blue/white screening on indicator plates. Recombinant plasmids were checked by *SspI* restriction enzyme and sequencing. Their sizes were calculated to be 3134 and 3132 bp. Digestion by the *SspI* restriction enzyme resulted in three specific bands equivalent to approximately 0.18, 0.75, and 2.2 kb as measured. Recombinant plasmids were then confirmed by Sanger sequencing.

Sequencing results proved that the cloned genes are exactly sMn1/2 belonging to the *E* and *RdRp* genes of SARS-CoV-2. The comparison consequences indicated two point mutations on the sMn2 sequence at positions of $G_{02} \rightarrow A_{02}$ and $T_{32} \rightarrow C_{32}$, while the sMn1 sequence was completely identical to all other 4 reference sequences. The important takeaway from sequence analysis is to verify if the sequences of interest were cloned into the exact position downstream to the T7 promoter of pGEM-T easy vector, therefore confirming that the transcribed *in-vitro* RNA afterward will be specifically a positive-sense, single-stranded RNA of SARS-CoV-2. The plasmids containing the targeted genes are named as pGEMT_sMn1_E and pGEMT_sMn2_P (Fig. 4).

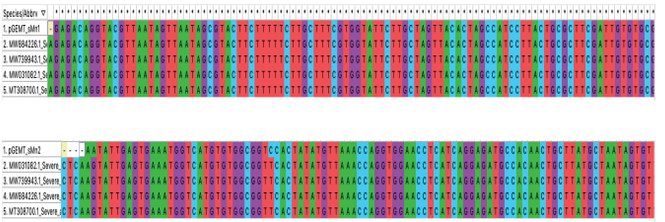

Fig. 4. Comparison between nucleotide sequence of sMn1/sMn2 of SARS-CoV-2 and other 4 reference Coronavirus SARS-CoV-2 strains on Genbank. Reference sequences are named in order of their code, name, and origin.

In-vitro synthesisation of envelope protein gene and RdRp gene of SARS-CoV-2

While the envelope protein gene (*E*) is a "universal" gene that codes for envelope proteins that are found in all coronaviruses, the RNA-dependent RNA polymerase gene (*RdRp*) of the ORF1ab sequence was found only in SARS-CoV-2. Given the importance of *RdRp* for viability and replication/transcription of RNA viruses, mutations in this gene are statistically less likely to occur [10]. Thus, numerous studies using these two genes have been provided to indicate SARS-CoV-2 based on real-time RT-PCR techniques. Our study focused on the highly conserved genomic region of the envelope protein and *RdRp* for cloning and transcription of *in-vitro* RNA, thus producing positive standards for further SARS-CoV-2 molecular diagnostic studies.

Transcription of *in-vitro* RNA was generated after the cleavage of the *NdeI* restriction enzyme to linearize the recombinant vectors. The *NdeI* enzyme was chosen as it has only one cleavage site on the recombinant plasmid at the position of 97 bp after the T7 promoter, which makes sure the process of transcription will result in a specific target RNA.

Theoretically, successfully transcribed *in-vitro* RNA sMn1 would have 216 bases and 214 bases for sMn2. Gel electrophoresis of the *in-vitro*-transcribed RNA showed clear bands and no by-products (Data not shown). The *in-vitro* transcribed RNA of the *E* and *RdRp* targets had concentrations of 200-800 ng/µl and A260/A280 ratios of 2.0-2.2. The number of copies was calculated by an online tool from http://www.scienceprimer.com [11] based on the concentration and the length of these single stranded *RNAs*, then RNA stocks were diluted into a series of 10-times concentration from 10^{-1} to 10^{6} copies/ml. All diluted RNA solutions were used as RNA templates for a real-time RT-PCR following WHO's protocol to confirm the *in-vitro*-transcribed RNA. Figure 5 shows that the targeted RNAs were successfully created. Additionally,

their stability during 1, 3, and 6 months of conservation shows equivalent results on one specific concentration (Data not shown).

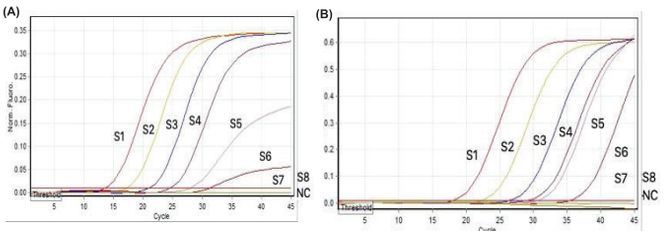

Fig. 5. Evaluation of quality of *in-vitro* transcribed *RNA* of sMn1 and sMn2 using the standard WHO protocol for SARS-CoV-2 detection by real-time RT-PCR. (A) sMn1 and **(B)** sMn2; S1-S8: in vitro RNA was diluted to the concentration of 10^{6}-10^{-1} copies/ul; NC: negative control.

In order to create a positive standard source enabling SARS-CoV-2 detection based on real-time RT-PCR, this study has proposed two cloned and transcribed *in-vitro* RNAs. Due to the properties of stability and ease of cloning into *E. coli* cells, plasmid-carrying target genes are normally used to establish positive standards. In spite of that, plasmids are only suitable for subjects with DNA as their genetic material. In this particular case, since SARS-CoV-2 uses RNA as its genetic material, the ideal positive standard must be RNA. Hence, *in-vitro*-transcribed RNA has been accessed by many studies [12-14].

Generally, two approaches are used to synthesise *in-vitro* RNA: direct transcription and indirect transcription through cloning. With direct transcription, RNA is produced using PCR products containing T7 promoter sequences at their 5'-end. This can be achieved by designing PCR primers having 5'-overhang ends including the T7 promoter [15, 16]. Meanwhile, for indirect transcription, no special design of the PCR primers is required, and plasmids are easily cloned with large amounts as the T7 promoter sequence was provided in the cloning vectors. On the other hand, through cloning techniques, target genes can be sustainably stored in plasmids as well as in *E. coli* cells. Our research had constructed an in vitro-transcribed RNA protocol of the *E* protein gene and *RdRp* gene of SAR-CoV-2 by indirect transcription resulting in RNA with relatively high purity when compared to direct assays [15, 16] and also higher than that of the in-vitro RNA synthesised according to Akyurek, et al. (2021) with about 10^{8} copies/µl (approx. 0.0536 ng/µl) [13] and L.T. Phan, et al (2020) [17] for 7.94 x10^{11} copies/µl.

Conclusions

This study successfully generated *in-vitro*-transcribed RNA coding for a specific fragment of the envelope protein gene (sMn1) and the *RdRp* gene (sMn2) of SARS-CoV-2 with concentrations of 200-800 ng/µl and A260/A280 ratios of 2.0-2.2 for sMn1 and sMn2. The *in-vitro* RNAs were tested for their content, purity, RNA electrophoresis, one-step real-time RT-PCR, and sequencing to confirm target fragments included in each RNA molecule. The stability of *in-vitro* RNA over a 6-month storage period was investigated. This experiment establishes a fundamental assay for the production of positive standards for RNA-based molecular techniques.

ACKNOWLEDGEMENTS

We would like to record our appreciation to everyone who participated in this work, especially the Institute of Biomedicine and Pharmacy, Vietnam Military Medical University for providing facilities and support toward our project. This work was supported by the Ministry of Science and Technology, Vietnam under grant number: DTDL.CN.29/20.

COMPETING INTERESTS

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

[1] J. Page, D. Hinshaw, B. McKay (2021), "In hunt for Covid-19 origin, patient zero points to second Wuhan market - The man with the first confirmed infection of the new coronavirus told the WHO team that his parents had shopped there", *The Wall Street Journal*, https://wsj.com/articles/in-hunt-for-covid-19-origin-patient-zero-points-to-second-wuhan-market-11614335404.

[2] C. Zimmer (2021), "The secret life of a Coronavirus", *The New York Times*, https://nytimes.com/2021/02/26/opinion/sunday/coronavirus-alive-dead.html.

[3] https://news.google.com/covid19/map?hl=vi&mid=%2Fm%2 F01crd5&gl=V&ceid=VN%3Avi.

[4] Z.W. Ye, et al. (2020), "Zoonotic origins of human coronaviruses", *International Journal of Biological Sciences*, **16(10)**, pp.1686-1697.

[5] A.A.T. Naqvi, et al. (2020), "Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach", *Biochim. Biophys. Acta Mol. Basis Dis.*, **1866(10)**, DOI: 10.1016/j.bbadis.2020.165878.

[6] R.H. Paul, I.I.G. Rather (2020), "The laboratory test for COVID-19, explained", *Science, the Wire*, https://science.thewire.in/the-sciences/rt-pcr-covid-19-rna-polymerase-thermal-cycler/.

[7] Y. Yang, et al. (2020), "SARS-CoV-2: Characteristics and current advances in research", *Virology Journal*, **17(1)**, DOI: 10.1186/s12985-020-01369-z.

[8] https://who.int/docs/default-source/coronaviruse/real-time-rt-pcr-assays-for-the-detection-of-sars-cov-2-institut-pasteur-paris.pdf.

[9] V.M. Corman, et al. (2020), "Diagnostic detection of 2019-nCoV by real-time RT-PCR", *Euro Surveill.*, **25(3)**, DOI: 10.2807/1560-7917.ES.2020.25.3.2000045.

[10] M. Pachetti, et al. (2020), "Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant", *Journal of Translational Medicine*, **18**, DOI: 10.1186/s12967-020-02344-6.

[11] http://scienceprimer.com/copy-number-calculator-for-realtime-pcr.

[12] J. Bland, et al. (2021), "Development and validation of viral load assays to quantitate SARS-CoV-2", *Journal of Virological Methods*, **291**, DOI: 10.1016/j.jviromet.2021.114100.

[13] S. Akyurek, et al. (2021), "The production and characterization of SARS-CoV-2 RNA reference material", *Analytical and Bioanalytical Chemistry*, **413(13)**, pp.3411-3419.

[14] T.G.W. Graham, et al. (2021), "Open-source RNA extraction and RT-qPCR methods for SARS-CoV-2 detection", *PLOS ONE*, **16(2)**, DOI: 10.1371/journal.pone.0246647.

[15] E. Keyaerts, et al. (2006), "Viral load quantitation of SARS-coronavirus RNA using a one-step real-time RT-PCR", *International Journal of Infectious Diseases*, **10(1)**, pp.32-37.

[16] M.L. Choudhary, et al. (2020), "Development of in vitro transcribed RNA as positive control for laboratory diagnosis of SARS-CoV-2 in India", *Indian J. Med. Res.*, **151(2&3)**, pp.251-254.

[17] L.T. Phan, et al. (2020), "Development of standardized specimens with known concentrations for severe acute respiratory syndrome coronavirus 2 Realtime-RT-PCR testing validation", *Bulletin of the World Health Organization*, DOI: 10.2471/BLT.20.259630.