Vietnam Journal of Biotechnology **22**(2): 382-390, 2024. DOI: 10.15625/vjbt-19874

CODON OPTIMIZATION, EXPRESSION, AND PURIFICATION OF THE REVERSE TRANSCRIPTASE ENZYME VARIANT MMLV RT2 IN *Escherichia coli*

Nguyen Van Sang[1,✉], Le Tra My[1], Nguyen Thi Uyen[1], Dang Thi Lan Anh[1]

[1]*Faculty of Biology, University of Science, Vietnam National University, Ha Noi, Vietnam*

✉To whom correspondence should be addressed. E-mail: nvsangvnu@hus.edu.vn

Received: 09.01.2024
Accepted: 03.06.2024

ABSTRACT

Moloney leukemia virus reverse transcriptase (MMLV RT) is an important enzyme widely utilized in molecular biology research and various diagnostic testing kits. A previous study demonstrated that variants of MMLV RT containing specific mutations contribute to enhanced stability and increased capability for RNA detection. However, the optimization and increased expression level of MMLV RT variants for large-scale production have not been extensively studied. In this research, the gene segment encoding the MMLV RT variant (E286R/E302K/L435R/D524A) was codon optimized and subcloned into pET-M expression vector. The obtained pET-M MMLV RT2 was transformed into *E. coli* BL21(DE3) host cell system for over-expression of the MMLV RT2 enzyme. The recombinant enzyme was then purified using nickel affinity chromatography and quantified using the Bradford method. The results revealed significant increase in recombinant enzyme yield, achieving 0.25 g/L in the culture medium. This represents a 1.4-fold increase compared to the 0.175 g/L yield reported in the previous study. Enzyme MMLV RT2 was also evaluated for its activity and was found to have comparable activity to the commercial enzyme Promega and another variant, enzyme MMLV RT1, which carries the D524A mutation and has previously been evaluated for activity. This result suggests that a high expression level of a stable MMLV-RT enzyme variant can be achieved which is beneficial for further improving the production of this important enzyme. The high yield and sustained activity of the MMLV RT2 variant indicate its potential as a reliable and cost-effective alternative for various applications in research and diagnostics. The successful optimization and expression in *E. coli* highlight the feasibility of large-scale production, potentially lowering costs and improving accessibility for molecular biology applications.

Keywords: affinity chromatography, codon optimization, expression, *Moloney leukemia*, reverse transcriptase.

INTRODUCTION

Moloney murin leukemia virus MMLV reverse transcriptase enzyme (MMLV RT), derived from the Moloney murine leukemia virus, is widely used in scientific research (Rein, 2011; Schultz, Champoux, 2008; Wong *et al.*, 1998). It is applied for various

purposes, such as cDNA synthesizing, real-time RT-PCR, microarray analysis, and RNA sequencing (Arezi, Hogrefe, 2009; Gerard, D'Alessio, 1993). However, the RNase H activity of MMLV RT has been indicated to reduce the efficiency of its cDNA synthesis *in vitro*. To resolve this challenge, mutations have been strategically developed to inhibit RNase H activity while maintaining the enzyme's original size (Arezi, Hogrefe, 2009; Mizuno *et al.*, 2010). The mutation at the D524A position has abolished RNase H activity, leading to an improvement in cDNA synthesis efficiency (Coté, Roth, 2008; Mizuno *et al.*, 2010). In 2014, Atsushi et al. indicated that E286R, E302K, and L435R mutations not only contribute to increasing the thermal stability of the MMLV RT enzyme but also result in the loss of RNase H activity (Konishi *et al.*, 2014; Yasukawa *et al.*, 2010). The combination of these mutations enhanced the catalytic activity and thermal stability of the enzyme (Baranauskas *et al.*, 2012). Therefore, optimizing the expression level and solubility of the enzyme in the protein expression system in *E. coli* is necessary to achieve the best economic efficiency.

In this study, we constructed the pET-M MMLV RT2 vector containing four mutations (E286R, E302K, L435R, and D524A). The MMLV RT2 enzyme was expressed in *E. coli* BL21 (DE3) and exhibited better solubility compared to the MMLV RT1 (carrying a single mutation, D524A) (Nguyen *et al*., 2023). The MMLV RT2 protein was purified using nickel affinity chromatography, yielding a substantial amount of up to 0.25 g/L. The M-MLV RT2 enzyme demonstrated good activity compared to both the GoScript enzyme (Promega) and the M-MLV RT1 enzyme, as evaluated using the one-step RT-PCR method.

MATERIALS AND METHODS

Materials

The *E. coli* strain BL21 (DE3) and the pET-M MMLV RT WT vector, which contains a DNA fragment encoding the MMLV RT wide-type enzyme, were provided by the Molecular Cell Biology Laboratory at the Center for Life Science Research, Faculty of Biology, Vietnam University of Science. The pET-M vector is a variant of the pET-32a vector with the S-tag and TrxA regions removed. The MMLV RT1 enzyme was prepared as described in a previous study (Nguyen *et al*., 2023). The reagents and chemicals were purchased from reputable suppliers, including Sigma, Qiagen, Merck, and Fermentas.

Methods

Construct the MMLV RT2 using overlap extension PCR

The pET-M MMLV RT WT vector was used as a template for PCR reactions to generate point mutations (Edelheit *et al.*, 2009). The primers used in the PCR reactions are presented in Table 1. The PCR reactions used Phusion DNA polymerase (Thermo Scientific). Fragments with small sizes containing point mutations were used as templates for overlapping PCR reactions to create the MMLV RT2 gene carrying the desired mutations. The PCR full-length products were then amplified using the RT2-*Bam*HI F and RT2-*Eco*RI R primers to create the MMLV RT2 gene containing recognition sequences for the restriction enzymes *Bam*HI and *Eco*RI for insertion

into the expression vector. The PCR reaction includes 10 µl of 5x HF buffer, 2 µl of F primer, 2 µl of R primer, 1 µl of Phusion DNA Polymerase, 2 µl of template, and additional water to bring the total reaction volume to 50 µl. The PCR cycling program consists of 35 cycles: 98°C for 1 minute, 55°C for 15 seconds, and 72°C for 15 seconds to 1 minute (depending on the size of the gene fragment). The PCR products were analyzed on a 1% TAE agarose gel and purified using the MEGAquick-spin™ Plus Total Fragment DNA Purification Kit from iNtRON. The PCR products, and pET-M vector were cut by FastDigest *Eco*RI and *Bam*HI retriction enzymes (Thermo Scientific). One (1) µg of pETM plasmid or 0.2 µg of PCR product was digested in a reaction mixture containing 2 µl of 10x FastDigest buffer, 1 µl of *Bam*HI enzyme, 1 µL of *Eco*RI enzyme, and water to a final volume of 30 µl at 37°C for30 mins. The fragments then were ligated into the pET-M expression vector using T4 DNA ligase (Thermo Scientific) following the manufacturer's instructions. The ligated products were transformed into *E. coli* DH5α cells by the heat shock method. The plasmid, after screening, was sequenced at 1st BASE, Malaysia, to determine the mutations and reading frame on the vector. The obtained vector was transformed into *E. coli* BL21 (DE3) for expression of the enzyme.

Table 1. The PCR primer sequences used in this study

Site	Primer	Sequence (5' – 3')
E286R	*E286R*_F	ACCGAAGCTCGTAAACGTACCGTTATGGGTCAGCCGACCCCG
	*E286R*_R	CGGTACGTTTACGAGCTTCGGTCAGCC
E302K	*E302K*_F	GCGTAAATTTCTGCCTACCGCTGGTTTCTGC
	*E302K*_R	CCCAGAAATTTACGCAGCTGACGCGGGGT
L435R	*L435R*_F	CTGGTTATCCGTGCTCCGCACGCTGTTGAAGC
	*L435R*_R	TGCGGAGCACGGATAACCAGCGGCTGACCCATGGTCAG
D524A	*D524A*_F	CCACACCTGGTACACCGCCGGTTCTTCTCTGC
	*D524A*_R	GCAGAGAAGAACCGGCGGTGTACCAGGTGTGG
RT2-*Bam*HI F		GCggatccATGACCCTGAACATCGAAGAC
RT2-*Eco*RI R		CGgaattcTTACAGCAGGGTAGAGGTGTCC

Expression and purification of recombinant proteins

The BL21 DE3 cells containing pET-M MMLV RT2 were cultured in LB media containing ampicillin 100 µg/mL (designated as LBA medium) at 37°C on a rotating shaker at 200 rpm. When the OD_{600} of the culture reached ~ 0.6, IPTG was added at the final concentration of 0.3 mM. The cell pellets obtained from 500 mL of LB were collected by centrifugation at 4000 rpm for 15 minutes at 4°C. The cell pellets were then resuspended in 50 mL of lysis buffer (50 mM Na_2HPO_4/NaH_2PO_4, pH 8.0; 300 mM KCl; 5% glycerol; 0.4% triton-X; 5 mM imidazole) and disrupted by sonication for 30 minutes on ice using a

BIOBASE ultrasonic cell disruptor (China). The soluble protein in the supernatant was collected by centrifugation at 13,000 rpm for 30 minutes at 4°C. The recombinant enzyme was then purified using Ni-NTA affinity chromatography. The lysate was applied to a nickel column equilibrated with lysis buffer and incubated for 10 minutes. The column was then washed with wash buffer (50 mM Na2HPO4/NaH2PO4, pH 8.0; 300 mM KCl; 5% glycerol; 0.4% triton-X; 5 mM imidazole) to remove impurities. The protein was eluted from the column using elution buffer (20 mM Tris-HCl, pH 8.0; 300 mM KCl; 5% glycerol; 0.2% triton-X; 250 mM imidazole), then dialyzed into the buffer (40 mM Tris-HCl, pH 8.0; 400 mM KCl; 5% glycerol).

After purification, the protein was analyzed using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The enzyme concentration was determined using the Bradford method. The protein was transferred to a storage buffer (40 mM Tris-HCl, pH 8.0; 400 mM KCl; 50% glycerol; 0.01% NP-40; 0.1 mM EDTA; 1 mM DTT) and stored at -80°C.

Determination of the MMLV RT2 enzyme activity

The enzyme activity was assessed using RT-PCR. A one-step RT-PCR reaction was performed with the GAPDH primers targeting a 495 bp gene, using total RNA extracted from Hela cells as a template (which has a stock concentration of 0.5 µg/µL) (Barber *et al.*, 2005). The enzyme activity was compared with a commercial enzyme (Promega) and the MMLV RT1 enzyme (previously assessed for good activity and produced in a previous study) (Nguyen *et al*., 2023).

RESULTS AND DISCUSSION

Plasmid pETM-MMLV RT2 construction

To create mutations in the MMLV RT2 sequence, the wild-type MMLV RT DNA fragment was used as a template for mutagenesis PCR. The PCR reaction employed four pairs of primers designed with specific mutations, generating five gene fragments labeled from F1 to F5 in **Figure 1A**. Subsequently, an overlap PCR reaction was used to produce a 2016 bp gene fragment with the desired point mutations, as shown in **Figure 1B**. This fragment was inserted into the pET-M vector, resulting in the recombinant plasmid pETM-MMLV RT2 (**Figure 1C)**. Figure **1D** illustrates the structure of this vector. The plasmid was confirmed to carry the desired mutations and maintain the correct open reading frame based on Sanger sequencing results.

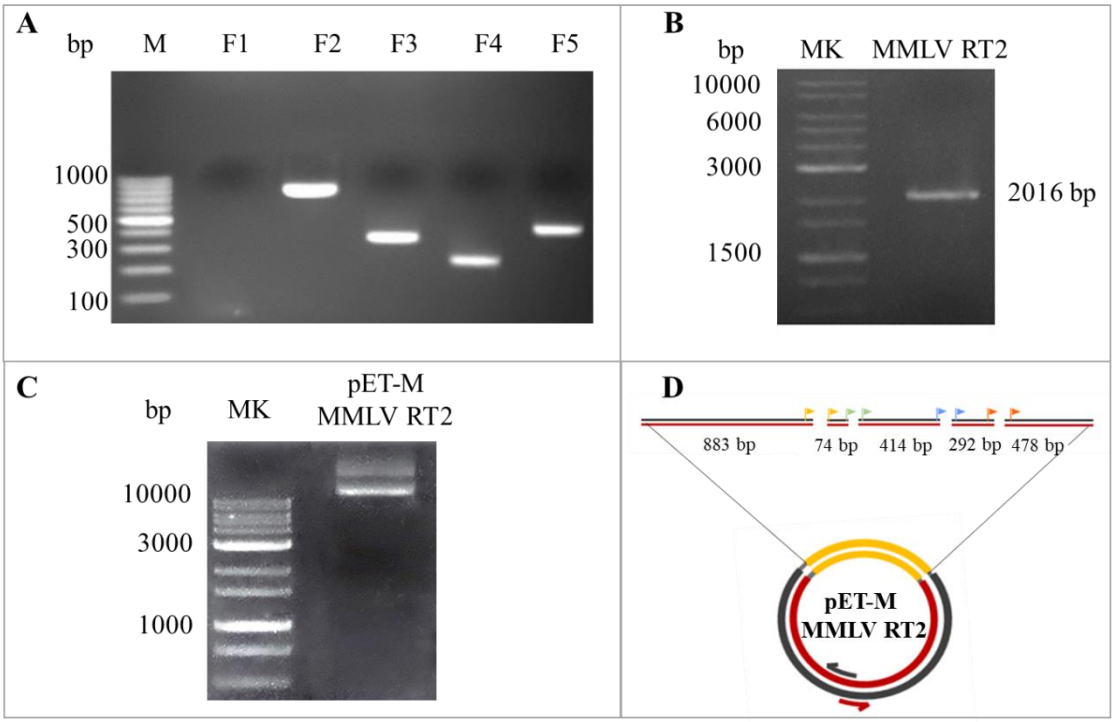

Figure 1. Construction of pET-M MMLV RT2. **A-C**: The electrophoresis analysis of DNA products; M:100 bp marker (Thermo Scientific); MK: Marker 1 kb Plus (Cleaver Scientific). **A**: PCR products amplifying component fragments of the gene *MMLV RT2* include F1, F2, F3, F4, and F5. These gene fragments have point mutations and respective sizes of 74 bp, 883 bp, 292 bp, and 414 bp. **B**: The PCR product of the amplified full MMLV RT2 gene of 2016 bp generated by overlap PCR of F1, F2, F3, F4, and F5 PCR fragments. **C**: pET-M MMLV RT2 plasmid. **D**: Structure of the pET-M MMLV RT2 vector.

The expression and solubility level of MMLV RT2 in E. coli BL21(DE3) cells

The MMLV RT2 enzyme has been successfully expressed in *E. coli* BL21 (DE3) cells. As shown in Figure 2A, with the same amount of protein (30 µg each) loaded into each well, the expression level of the MMLV RT2 protein (of 75 kDa) is higher compared to MMLV RT1. After sonication, the protein quantity in the soluble phase of MMLV RT2 (lane 5) is also greater than that of MMLV RT1 (lane 9). Therefore, it can be observed that the mutations (E286R/E302K/L435R/D524A) have increased the expression level and improved the solubility of the protein compared to MMLV RT1, which only carries the D524A mutation. Using ImageJ software to quantify enzyme protein bands and compare, the results show that the expression level of MMLV RT2 increased 1.8-fold, and the solubility increased 1.9-fold compared to the MMLV RT1 variant (Nguyen *et al*., 2023).

Thus, in addition to codon optimization for enhanced protein expression in *E. coli*, specific mutations (E286R/E302K/L435R/D524A) in the protein can also improve the solubility of the MMLV RT enzyme. Codon

optimization is a crucial strategy for enhancing protein expression. Previous studies have demonstrated that codon optimization has significantly increased the expression levels of various proteins (Frumkin *et al.*, 2018; Yu *et al.*, 2015). Furthermore, the mutations (E286R/E302K/L435R/D524A) can also help the protein maintain a more stable structure under various temperature and pH conditions, minimizing degradation and loss of activity (Konishi *et al.*, 2014; Mizuno *et al.*, 2010; Yasukawa *et al.*, 2010). These mutations notably improve the solubility of the protein compared to the variant containing only the D524A mutation.

Purification of MMLV RT2

The MMLV RT2 enzyme was purified using the nickel affinity chromatography method. The results are shown in Figure 3B, in lanes 8 – 9, where the protein was purified. Subsequently, the MMLV RT2 enzyme was dialyzed in storage buffer (20 mM Tris HCl pH 8.0; 200 mM KCl; 5% glycerol; 0.4% triton-X) and stored at −80°C.

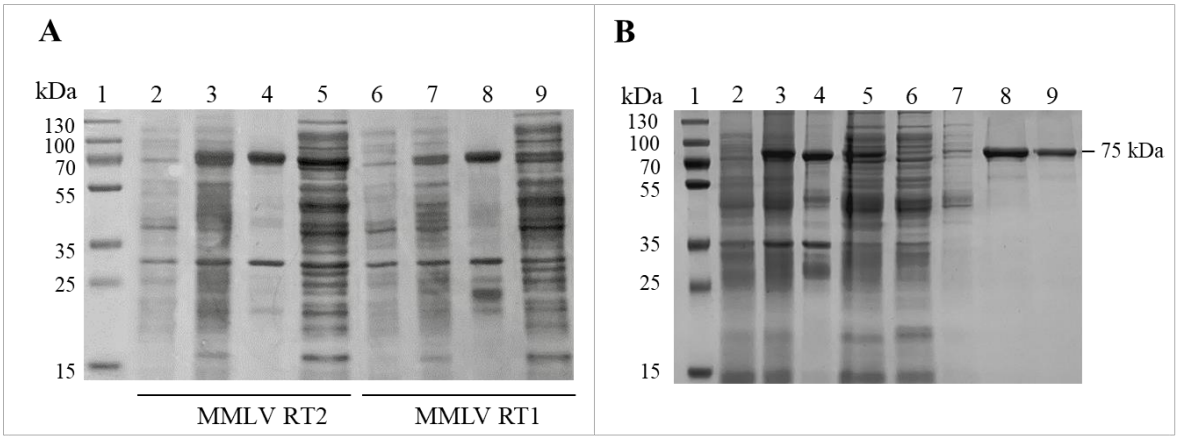

Figure 2. Analysis of protein expressed in *E. coli* (**A**) and protein fractions during the purification process (**B**) by SDS-PAGE. **A**: 1: the protein ladder (Thermo Scientific); 2&6: The recombinant *E. coli* BL21 cells were not induced by IPTG (negative control); 3-5, 7-9 proteins from the recombinant *E. coli* BL21 cells induced by IPTG, in which 3&7 were total proteins; 4&8 were insoluble proteins and the others were soluble proteins. **B**: 1: the protein ladder (Thermo Scientific); 2: negative control, 3-5: total, soluble, and insoluble protein fractions, respectively; 6: Flow through after incubation in the nickel column; 7: wash fraction; 8 – 9: elution fractions.

Previous studies on the RT enzyme by the Kiyoshi group in 2014 (Konishi *et al.*, 2014) and the Katano group in 2017 (Katano *et al.*, 2017) indicated the necessity of multiple purification steps to achieve high purity. Such multi-step purification processes resulted in significant protein loss at each step. In our study, using a single purification step with nickel affinity chromatography yielded bands of comparable purity. This method shortens the purification process while maintaining the quality of the purified enzyme.

Evaluation of the MMLV RT2 enzyme activity

The activity of the MMLV RT2 enzyme

was evaluated by an RT-PCR reaction. The one-step RT-PCR used Hela mRNA as a template and the GAPDH gene as primers. The activity of the MMLV RT2 enzyme was compared with the Goscript enzyme from Promega and the MMLV RT1 enzyme, which had previously demonstrated good activity in a previous study.

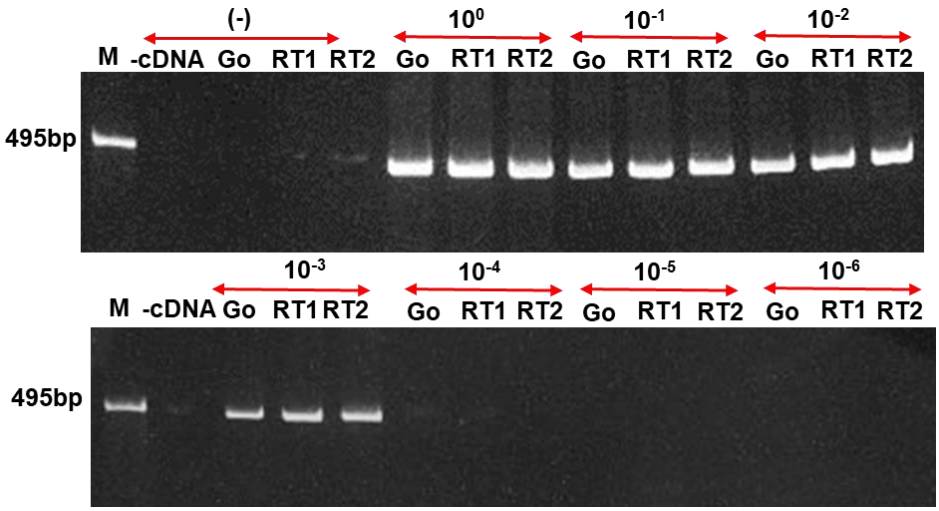

Figure 3. The result enzyme activity assay compared the activity of MMLV RT2 with MMLV RT1 and Goscript commercial enzyme. **M**: The standard sample has a size of 495 bp; **-cDNA**: Control sample without cDNA; **Go**: the commercial enzyme Goscript (Promega); **RT1**: MMLV RT1 enzyme; **RT2**: MMLV RT2 enzyme; **10^0 – 10^{-6}**: Hela RNA diluted from 1 to 10^6 folds.

As illustrated in Figure 3, the RT2 enzyme exhibits activity comparable to that of both the commercial Goscript enzyme and the MMLV RT1 enzyme. Notably, all three enzymes (MMLV RT2, Goscript, and MMLV RT1) are capable of detecting as little as 0.5 ng of total HeLa RNA in a 10 µl one-step RT-PCR reaction. In these RT-PCR reactions, the final concentration of RNA was approximately 0.05 ng/µl, corresponding to a 10^{-3} fold dilution from the stock total HeLa RNA template with an initial concentration of 0.5 µg/µl. At template dilution concentrations ranging from 10^0 to 10^{-1} and 10^{-2}, all three enzymes exhibited good detection ability, resulting in bands of similar brightness. This result indicated that the activity of the RT2 enzyme obtained from this study is as good as that of the commercial enzyme, which can be used for reverse transcription applications. Furthermore, previous studies on RT enzymes used RT activity assays or Two-step RT–PCR assays to assess enzyme activity (Arezi, Hogrefe, 2009). In this study, we employed an One-step RT-PCR assay using HeLa RNA, a common technique for assessing RT enzyme activity (Gerard, D'Alessio, 1993). These results demonstrated high and reliable enzyme activity. Based on these, it can be concluded that the enzyme expressed and purified in this study is of high quality and has potential applications for RT reactions. This high level of activity, combined with the single-step purification method we employed, underscores the efficiency of our approach. The quality of the enzyme, as evidenced by its robust activity, suggests that the mutations and codon optimization

strategies we implemented not only enhanced expression and stability but also maintained the functional integrity of the enzyme.

CONCLUSION

The pETM-MMLV RT2 vector has been constructed, carrying four mutations (E286R/E302K/L435R/D524A). The MMLV RT2 enzyme is well-expressed in *E. coli* BL21 (DE3) bacterial cells and exhibits better solubility compared to the MMLV RT1 enzyme carrying a single mutation D524A. The MMLV RT2 enzyme is purified using nickel affinity chromatography with a substantial yield of up to 0.25 g/L of culture medium and high purity. The activity of the MMLV RT2 enzyme is assessed to be equivalent to that of the commercial enzyme from Promega and the MMLV RT1 enzyme.

ACKNOWLEDGMENTS

This work was supported by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 106-NN.02-2016.58.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

Arezi, B., & Hogrefe, H. 2009. Novel mutations in Moloney Murine Leukemia Virus reverse transcriptase increase thermostability through tighter binding to template-primer. *Nucleic acids research*, 37(2): 473-481. https://doi.org/10.1093/nar/gkn952.

Baranauskas, A., Paliksa, S., Alzbutas, G., Vaitkevicius, M., Lubiene, J., Letukiene, V., Burinskas, S., Sasnauskas, G., & Skirgaila, R. 2012. Generation and characterization of new highly thermostable and processive M-MuLV reverse transcriptase variants. *Protein Engineering, Design & Selection*, 25(10): 657-668. https://doi.org/10.1093/protein/gzs034.

Barber, R. D., Harmer, D. W., Coleman, R. A., & Clark, B. J. 2005. GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. *Physiological genomics*, 21(3): 389-395. http://doi.org/10.1152/physiolgenomics.00025.2005.

Coté, M. L., & Roth, M. J. 2008. Murine leukemia virus reverse transcriptase: structural comparison with HIV-1 reverse transcriptase. *Virus research*, 134(1-2): 186-202. https://doi.org/10.1016/j.virusres.2008.01.001.

Edelheit, O., Hanukoglu, A., & Hanukoglu, I. 2009. Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC biotechnology*, 9: 1-8. https://doi.org/10.1186/1472-6750-9-6.

Frumkin, I., Lajoie, M. J., Gregg, C. J., Hornung, G., Church, G. M., & Pilpel, Y. 2018. Codon usage of highly expressed genes affects proteome-wide translation efficiency. *Proceedings of the National Academy of Sciences*, 115(21): E4940-E4949. https://doi.org/10.1073/pnas.1719375115.

Gerard, G. F., & D'Alessio, J. M. 1993. Reverse Transcriptase (EC 2.7. 7.49) The Use of Cloned Maloney Murine Leukemia Virus Reverse Transcriptase to Synthesize

DNA from RNA. *Enzymes of molecular biology*: 73-93. https://doi.org/10.1007/BF02762340.

Katano, Y., Li, T., Baba, M., Nakamura, M., Ito, M., Kojima, K., Takita, T., & Yasukawa, K. 2017. Generation of thermostable Moloney murine leukemia virus reverse transcriptase variants using site saturation mutagenesis library and cell-free protein expression system. *Bioscience, Biotechnology, and Biochemistry*, 81(12): 2339-2345. https://doi.org/10.1080/09168451.2017.1394790.

Konishi, A., Hisayoshi, T., Yokokawa, K., Barrioluengo, V., Menéndez-Arias, L., & Yasukawa, K. 2014. Amino acid substitutions away from the RNase H catalytic site increase the thermal stability of Moloney murine leukemia virus reverse transcriptase through RNase H inactivation. *Biochemical and biophysical research communications*, 454(2): 269-274. https://doi.org/10.1016/j.bbrc.2014.10.044.

Mizuno, M., Yasukawa, K., & Inouye, K. 2010. Insight into the mechanism of the stabilization of Moloney murine leukaemia virus reverse transcriptase by eliminating RNase H activity. *Bioscience, biotechnology, and biochemistry*, 74(2): 440-442. https://doi.org/10.1271/bbb.90777.

Nguyen Van Sang, Nguyen Thi Thu Huyen, Ho Thi Yen, & Nguyen Thi Uyen. 2023. Expression, purification, and activity determination of M-MLV reverse transcriptase enzyme with RNase H domain mutations in *E. coli* host. *Scientific Report at the National Biotechnology Conference* 2023, (34): 216-222.

Rein, A. 2011. Murine leukemia viruses: objects and organisms. *Advances in virology*, 2011. https://doi.org/10.1155/2011/403419.

Schultz, S. J., & Champoux, J. J. 2008. RNase H activity: structure, specificity, and function in reverse transcription. *Virus research*, 134(1-2): 86-103. https://doi.org/10.1016/j.virusres.2007.12.007.

Wong, L., Pearson, H., Fletcher, A., Marquis, C. P., & Mahler, S. 1998. Comparison of the efficiency of moloney murine leukaemia virus (M-MuLV) reverse transcriptase, RNase H--M-MuLV reverse transcriptase and avian myeloblastoma leukaemia virus (AMV) reverse transcriptase for the amplification of human immunoglobulin genes. *Biotechnology Techniques*, 12: 485-489. https://doi.org/10.1023/A:1008871608947.

Yasukawa, K., Mizuno, M., Konishi, A., & Inouye, K. 2010. Increase in thermal stability of Moloney murine leukaemia virus reverse transcriptase by site-directed mutagenesis. *Journal of biotechnology*, 150(3): 299-306. https://doi.org/10.1016/j.jbiotec.2010.09.961.

Yu, C.-H., Dang, Y., Zhou, Z., Wu, C., Zhao, F., Sachs, M. S., & Liu, Y. 2015. Codon usage influences the local rate of translation elongation to regulate co-translational protein folding. *Molecular cell*, 59(5): 744-754. https://doi.org/10.1016/j.molcel.2015.07.018.