Production of ORF8 protein from SARS-CoV-2 using an inducible virus-mediated expression-system in suspension-cultured tobacco BY-2 cells

Tomohiro Imamura[1], Noriyoshi Isozumi[2], Yasuki Higashimura[3], Shinya Ohki[2], and Masashi Mori[1*]

[1] Research Institute for Bioresources and Biotechnology, Ishikawa Prefectural University, 308-1 Suematsu, Nonoichi, Ishikawa 921-8836, Japan

[2] Center for Nano Materials and Technology (CNMT), Japan Advanced Institute of Science and Technology (JAIST), 1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan

[3] Department of Food Science, Ishikawa Prefectural University, 308-1 Suematsu, Nonoichi, Ishikawa 921-8836, Japan

***Corresponding authors:**

Masashi Mori

Laboratory of Plant Gene Function, Research Institute for Bioresources and Biotechnology, Ishikawa Prefectural University, 308-1 Suematsu, Nonoichi, Ishikawa 921-8836, Japan

**Phone:** +81-76-227-7527

**Fax:** +81-76-227-7557

**E-mail:** mori@ishikawa-pu.ac.jp

**Abstract**

COVID-19, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which spread worldwide in 2019, is an urgent problem to be overcome. The ORF8 of SARS-CoV-2 has been suggested to be associated with the symptoms of COVID-19, according to reports of clinical studies. However, little is known about the function of ORF8. As one of the ways to advance the functional analysis of ORF8, mass production of ORF8 with the correct three-dimensional structure is necessary. In this study, we attempted to produce ORF8 protein by chemical-inducible protein production system using tobacco BY-2 cells. An ORF8-producing line was generated by the *Agrobacterium* method. As a result, the production of ORF8 of 8.8 ± 1.4 mg/L of culture medium was confirmed. SDS-PAGE and Nuclear magnetic resonance (NMR) analysis confirmed that the ORF8 produced by this system is a dimeric form with a three-dimensional structure, unlike that produced in *Escherichia coli*. Furthermore, it was suggested that the ORF8 produced by this system was *N*-glycosylated. Through this study, we succeeded in producing ORF8 with the correct three-dimensional structure in a chemical-inducible protein production system using tobacco BY2 cells. It is expected that the functional analysis of ORF8 will be advanced using the ORF8 produced by this system and that it will greatly contribute to the development of vaccines and antibodies against ORF8.

**Keywords:** Chemical-inducible virus-mediated protein production; COVID-19; NMR; ORF8; SARS-CoV-2; Tobacco BY-2 cells; ToMV

Coronavirus disease 2019 (COVID-19), caused by SARS-CoV-2 and characterized by severe pneumonia, spread worldwide in 2019. To overcome the threat of SARS-CoV-2, development of therapeutic agents and vaccines based on functional elucidation is urgently needed. The SARS-CoV-2 genome was sequenced recently (Acc. No.: NC_045512.2, Fig. 1a). Interestingly, ORF8 which is an accessory protein, has a lower homology than the closely related SARS virus, unlike other SARS-CoV-2 proteins. Instead, SARS-CoV-2 ORF8 showed high homology with bat coronavirus ORF8 (Lu et al., 2020). SARS-CoV-2 mutants lacking ORF8 were reported to be less likely to cause severe disease (Young et al., 2020). ORF8 is reported to be highly immunogenic and anti-ORF8 antibodies are formed in the early stage of infection (Hachim et al., 2020). Furthermore, it has been reported that a significant T-cell response to ORF8 is observed in recovered patients (Grifoni et al., 2020). Thus, ORF8 is predicted to be closely related to the symptoms caused by SARS-CoV-2. X-ray crystallization study has revealed the three-dimensional structure of ORF8 (Flower et al., 2020), and that ORF8 affects the immune system (Zhang et al., 2020). Further studies on ORF8 function are warranted.

To analyze ORF8 function, a large amount of homogeneous and structurally correct sample needs to be produced, which can also be used for developing antibodies and vaccines to treat COVID-19. A plant-based production system can mass produce eukaryotic proteins that are physiologically active. We have developed a highly efficient chemical-inducible protein production system using tobacco BY-2 cells, utilizing the

extremely high protein production capacity of tobamovirus (ToMV) (Dohi et al., 2006, Fig. 1b). Protein production using plant-cultured cells is large-scale, uniform, aseptic, and GMP-compliant. Our system can mass produce eukaryotic proteins with multiple disulfide-bonds with sufficient quality for structural and functional studies. We have been successful in the production and structure determination (by solution NMR) of proteins that were difficult to produce uniformly in *E. coli.*, including plant-bioactive (Costa et al., 2014; Ohki et al., 2011) and animal (Ohki et al., 2008) proteins. Mature ORF8 has seven cysteines, six of which form three intramolecular disulfide-bond pairs and one forms an intermolecular disulfide-bond for dimerization (Flower et al., 2020, Fig. 1c). In addition, one *N*-glycosylation site is predicted for ORF8 (Fig 1c). Our production system is expected to be suitable for uniform ORF8 protein production with these structural features.

For producing ORF8, we designed an amino acid sequence in which the signal peptide extracellular of translocation derived from *Arabidopsis* chitinase and 8× His-Tag were added to the N-terminal of mature ORF8. Next, artificial *ORF8* was synthesized by optimizing codons in tobacco and introducing restriction enzyme sites for cloning at both ends (IDT, Coralville, USA; Acc. No.: LC586256). The artificially synthesized *ORF8* was introduced into chemical-inducing tobamovirus vector with a restriction enzyme (pBICLBSER-ToMV-SP-His-ORF8, Fig. 1d). Next, pBICLBSER-ToMV-SP-His-ORF8 and pBICHgLBSXVE (Fig. 1d) expressing the artificial transcription factor XVE, which activates transcription by binding with 17β-estradiol (Dohi et al., 2006), were introduced

into tobacco BY-2 cells by the *Agrobacterium* method, and the ORF8-producing line was selected (Fig. 1e). The ORF8-producing line was suspension-cultured in normal MS medium and labeled MS medium using an $^{15}N$ nitrogen source for NMR analysis. ORF8 production was induced by adding 10 μM 17β-estradiol (Fig. 1e). Four days later, the induced tobacco BY-2 cells had significantly suppressed proliferation due to viral vector replication compared to non-induced cells (Fig. 1f). The production of ORF8 was confirmed in the induced cell line, and ORF8 was the major protein found in the culture medium (Fig. 1g). This production system could produce 8.8 ± 1.4 mg of ORF8 per liter culture medium. Following purification using a nickel column, ORF8 in the culture medium could be obtained in the form of a single protein (Fig. 1h). In *E. coli*, ORF8 did not form the correctly folded three-dimensional structure and required refolding; moreover, a mixture of monomers and dimers is produced (Flower et al., 2020), However, ORF8 produced in tobacco BY-2 cells was dimeric due to the formation of disulfide-bonds (Fig. 1h, j). Furthermore, $^{1}H$ and $^{1}H$-$^{13}C$ HSQC NMR analysis suggested that ORF8 was glycosylated (Fig. 1j, k), indicating that ORF8 causes *N*-glycosylation in animal cells. Moreover, distribution, line-shape, and intensity of the peaks in $^{1}H$ and $^{1}H$-$^{15}N$ HSQC strongly indicated that ORF8 prepared using the BY-2 system forms a stable β-sheet-rich three-dimensional structure (Fig. 1j, l). The number of peaks in $^{1}H$-$^{15}N$ HSQC was less than expected probably because the larger molecular size due to *N*-glycosylation and dimerization caused line broadening.

In this study, we succeeded in producing structurally intact ORF8 in a ToMV-mediated

chemical-induced protein production system using tobacco BY-2 cells. Furthermore, the purification process could be simplified by releasing ORF8 into the culture medium that was low in plant-derived contaminants. The advantage of protein production using plants is that there is no risk of contamination with endotoxins and pathogens compared to production using other organisms. Our finding shows that it is possible to produce structural ORF8 protein in large quantities, with high efficiency. Because ORF8 had an intact three-dimensional structure, it can be used for functional analysis and the production of antibodies and vaccines targeting ORF8.

**Acknowledgments**

This work was partly supported by the Nanotechnology Platform of MEXT, Japan. The authors thank Akiko Mizuno, Hiroko Hayashi, and Akio Miyazato for technical assistance.

**Author contributions**

MM conceived this study. MM and TI designed the experiments. All authors performed the experiments. TI, SO and MM wrote the manuscript. All authors have read and approved the final manuscript.

**Conflict of interest**

The authors have no conflicts of interest to declare.

## References


Costa, L.M., Marshall, E., Tesfaye, M., Silverstein, K.A., Mori, M., Umetsu, Y., Otterbach, S.L., et al. (2014) Central cell-derived peptides regulate early embryo patterning in flowering plants. *Science* **344**, 168-172.

Dohi, K., Nishikiori, M., Tamai, A., Ishikawa, M., Meshi, T. and Mori, M. (2006) Inducible virus-mediated expression of a foreign protein in suspension-cultured plant cells. *Archives of virology* **151**, 1075-1084.

Flower, T.G., Buffalo, C.Z., Hooy, R.M., Allaire, M., Ren, X. and Hurley, J.H. (2020) Structure of SARS-CoV-2 ORF8, a rapidly evolving coronavirus protein implicated in immune evasion. *bioRxiv* (preprint) doi: https://doi.org/10.1101/2020.08.27.270637

Grifoni, A., Weiskopf, D., Ramirez, S.I., Mateus, J., Dan, J.M., Moderbacher, C.R., Rawlings, S.A., et al. (2020) Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* **181**, 1489-1501.e1415.

Hachim, A., Kavian, N., Cohen, C.A., Chin, A.W.H., Chu, D.K.W., Mok, C.K.P., Tsang, O.T.Y., et al. (2020) ORF8 and ORF3b antibodies are accurate serological markers of early and late SARS-CoV-2 infection. *Nature immunology* **21**, 1293-1301.

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., et al. (2020) Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565-574.

Ohki, S., Dohi, K., Tamai, A., Takeuchi, M. and Mori, M. (2008) Stable-isotope labeling using an inducible viral infection system in suspension-cultured plant cells. *Journal of biomolecular NMR* **42**, 271-277.

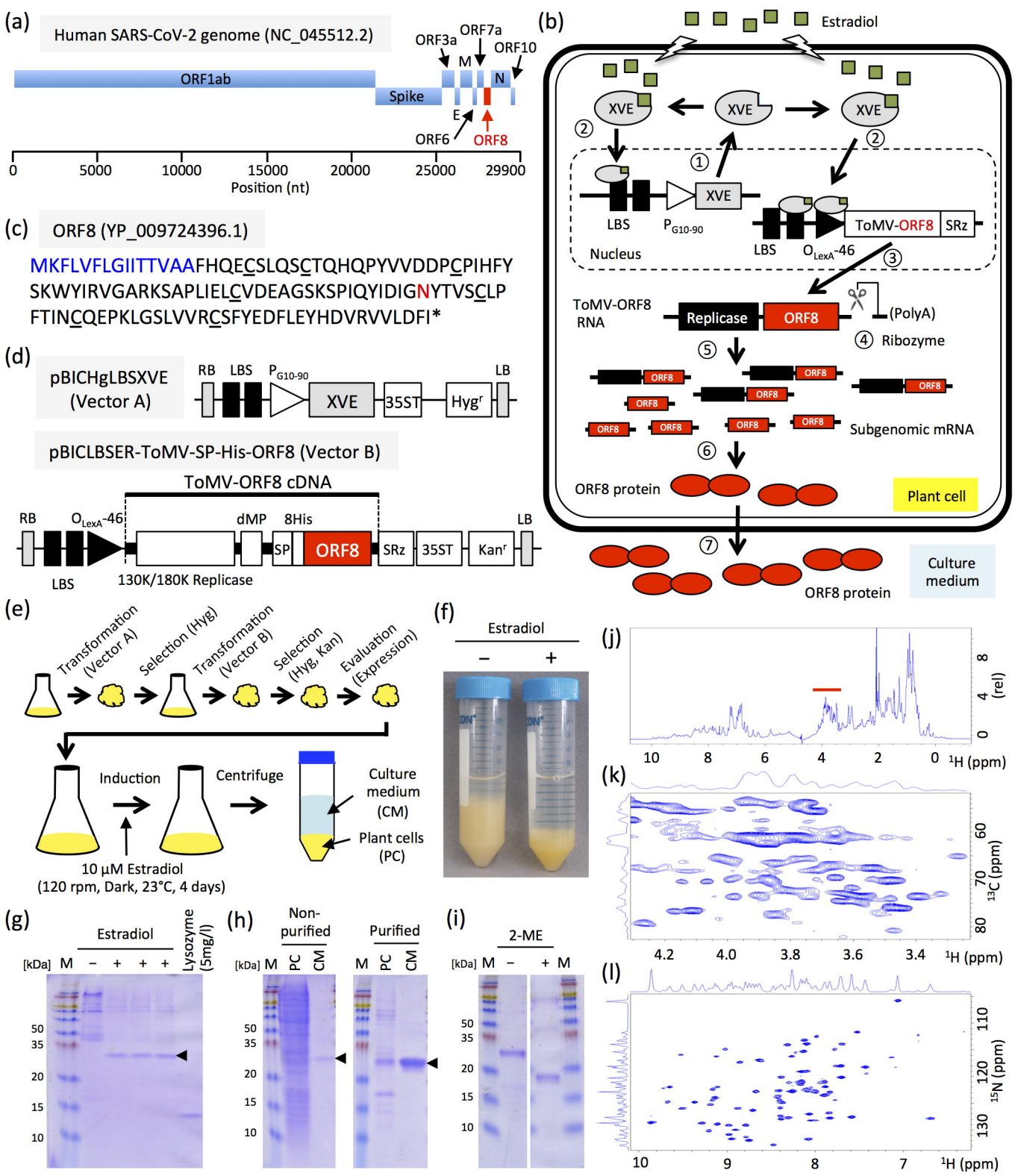
Ohki, S., Takeuchi, M. and Mori, M. (2011) The NMR structure of stomagen reveals the basis of stomatal density regulation by plant peptide hormones. *Nature communications* **2**, 512.

Young, B.E., Fong, S.W., Chan, Y.H., Mak, T.M., Ang, L.W., Anderson, D.E., Lee, C.Y., et al. (2020) Effects of a major deletion in the SARS-CoV-2 genome on the severity of infection and the inflammatory response: an observational cohort study. *Lancet* **396**, 603-611.

Zhang, Y., Zhang, J., Chen, Y., Luo, B., Yuan, Y., Huang, F., Yang, T., et al. (2020) The ORF8 protein of SARS-CoV-2 Mediates Immune Evasion through Potently Downregulating MHC-I. *bioRxiv* (preprint) doi: https://doi.org/10.1101/2020.05.24.111823

## Figure legend

**Figure 1.** Production and characterization of ORF8. (a) Schematic representations of SARS-CoV-2 genome. E, M, and N indicate envelope protein, membrane glycoprotein, and nucleocapsid phosphoprotein, respectively. (b) Schematic overview of the chemical-inducible ToMV-mediated expression-system. Step 1: Constitutive expression of XVE, a transcriptional activator. Step 2: XVE binds to estradiol and gets activated. Step 3: Activated XVE binds to the $O_{LexA-46}$ promoter or LBS to promote transgene transcription. Step 4: Non-viral sequences that interfere with the replication of the viral vector are removed by ribozymes. Step 5: Viral vector replication. Step 6: Translation of ORF8 from subgenomic RNA. Step 7: Translocation of ORF8. (c) Amino acid sequence of ORF8. Blue and red letters indicate signal peptide and *N*-glycosylation site, respectively. Underbars indicate cysteine. (d) Schematic representations of ToMV-mediated chemical-induced expression plasmids. XVE, transcriptional activator that responds to estrogen; SP, signal peptide of *Arabidopsis* chitinase; 8 His, 8× His-tag; LBS, LexA binding site; $P_{G10-90}$, synthetic constitutive promoter; $P_{LexA-46}$, fusion promoter controlled by XVE; dMP, partial movement protein; SRz, ribozyme sequence from tobacco ringspot virus satellite RNA; 35ST, 35S terminator; RB, right border; and LB, left border. $Hyg^r$ and $Kan^r$ indicate expression cassette of hygromycin and kanamycin resistance genes, respectively. (e) Generation of ORF8-producing cell line and induction of ORF8 production. (f) Four days after 17β-estradiol induction of suspension-cultured tobacco BY-2 cells; + and − indicate with or without 17β-estradiol, respectively. (g) Detection of ORF8 from 7.5 μL cultured medium by Coomassie Brilliant Blue staining. + and − indicate with or without 17β-estradiol, respectively. Arrowhead indicates ORF8. Lysozyme (5 mg/L) indicates loading control. (h) Non-purified and purified ORF8. PC and CM indicate crude extracts of plant cells and culture medium, respectively. (i) Molecular property of ORF8. + and − indicate with or without 2-mercaptoethanol (2ME), respectively. M, protein size marker. Numbers indicate molecular weights (kDa). (j) $^1H$ NMR spectra of unlabeled ORF8. Red line indicates the glycan signal appearing region. (k) $^1H$-$^{13}C$ HSQC of unlabeled ORF8 protein. An expanded view of the glycan signal appearing region is shown. (l) $^1H$-$^{15}N$ HSQC of $^{15}N$-labeled ORF8.


(a)
Human SARS-CoV-2 genome (NC_045512.2)
ORF1ab
Spike
ORF3a
ORF7a
M
ORF10
N
E
ORF6
ORF8
0
5000
10000
15000
20000
25000
29900
Position (nt)
(b)
Estradiol
XVE
LBS
P G10-90
Nucleus
ToMV-ORF8
SRz
O LexA-46
ToMV-ORF8 RNA
Replicase
ORF8
(PolyA)
Ribozyme
Subgenomic mRNA
ORF8 protein
Plant cell
Culture medium
(c)
ORF8 (YP_009724396.1)
MKFLVFLGIITTVAAFHQECSLQSCTQHQPYVVDDPCPIHFY
SKWYIRVGARKSAPLIELCVDEAGSKSPIQYIDIGNYTVSCLP
FTINCQEPKLGSLVVRCSFYEDFLEYHDVRVVLDFI*
(d)
pBICHgLBSXVE
(Vector A)
RB
LBS
P G10-90
XVE
35ST
Hyg r
LB
pBICLBSER-ToMV-SP-His-ORF8 (Vector B)
ToMV-ORF8 cDNA
RB
O LexA-46
dMP
8His
SP
ORF8
SRz
35ST
Kan r
LB
LBS
130K/180K Replicase
(e)
Transformation (Vector A)
Selection (Hyg)
Transformation (Vector B)
Selection (Hyg, Kan)
Evaluation (Expression)
Induction
Centrifuge
Culture medium (CM)
Plant cells (PC)
10 µM Estradiol
(120 rpm, Dark, 23°C, 4 days)
(f)
Estradiol
–
+
(g)
Estradiol
Lysozyme (5mg/l)
[kDa]
M
(h)
Non-purified
Purified
PC
CM
(i)
2-ME
(j)
(rel)
¹H (ppm)
(k)
¹³C (ppm)
(l)
¹⁵N (ppm)