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

ORCID: 0000-0002-9021-8104

Abstract

COVID-19, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which spread worldwide in 2020, 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 chemically 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 single structure, unlike that produced in *Escherichia coli*. Furthermore, it was suggested that the ORF8 produced by this system was glycosylated. Through this study, we succeeded in producing ORF8 folded into a single structure in a chemically inducible protein production system using tobacco BY-2 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 antibodies and therapeutic agents targeting ORF8.

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

Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and characterized by severe pneumonia, spread worldwide in 2020. To overcome the threat of SARS-CoV-2, the 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, open reading frame 8 (ORF8), an accessory protein of SARS-CoV-2, has lower homology than that of the closely related SARS virus, unlike other SARS-CoV-2 proteins. Instead, ORF8 of SARS-CoV-2 shows high homology with ORF8 of bat coronavirus (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 associated with the symptoms caused by SARS-CoV-2. An 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 protein folded into a single structure needs to be produced, which can also be used for developing antibodies and therapeutic agents to treat COVID-19. A plant-based production system can be used to mass produce eukaryotic proteins that are physiologically active. We have developed a highly efficient chemically inducible protein production system using tobacco BY-2 cells,

utilizing the extremely high protein production capacity of tobamovirus (ToMV) (Dohi, et al. 2006). In this study, we used an improved expression system that provided positive feedback on transcription factor (XVE) expression (Fig. 1b). Protein production using plant-cultured cells is a large-scale process that is 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 successfully produced proteins that were difficult to produce uniformly in *Escherichia coli*. including plant-bioactive (Costa, et al. 2014; Ohki, et al. 2011) and animal (Ohki, et al. 2008) proteins and have determined their structures using solution nuclear magnetic resonance (NMR). Mature ORF8 contains 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 has been predicted for ORF8 (Fig 1c). Our production system is expected to be suitable for uniform ORF8 protein production with these structural features.

To produce ORF8, we designed the amino acid sequence for the 8 His-tagged ORF8 fusion protein with an extracellular signal peptide added to the N-terminus. 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 a chemically inducible tobamovirus vector (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 cells were suspension-cultured in normal MS medium and MS medium labeled with an ^{15}N nitrogen source for NMR analysis. ORF8 production was induced by adding 10 μM 17β-estradiol (Fig. 1e). After 4 days, the induced tobacco BY-2 cells had significantly suppressed proliferation due to viral vector replication compared to non-induced cells (Fig. 1f). ORF8 production was confirmed in the induced cells. 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 was produced (Flower, et al. 2020). In contrast, the molecular weight of ORF8 produced by tobacco BY-2 cells was changed by approximately 50% in the presence of 2-mercaptoethanol (2ME) compared with that in the absence of 2ME. This result indicates that ORF8 produced in tobacco BY-2 cells are dimers due to the formation of disulfide bonds (Fig. 1i). Furthermore, ^{1}H and ^{1}H-^{13}C HSQC NMR analysis suggested that ORF8 was glycosylated (Fig. 1j, k), indicating that ORF8 is glycosylated in infected individuals. Moreover, distribution, line-shape, and intensity of the peaks in ^{1}H and ^{1}H-^{15}N HSQC strongly indicated that ORF8 produced 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 glycosylation and dimerization caused line broadening.

In this study, we successfully produced ORF8 that is folded into a single structure in a ToMV-mediated chemically 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. Our finding shows that it is possible to produce ORF8 folded into a single structure in large quantities with high efficiency. Because ORF8 has an intact three-dimensional structure, it can be used for functional analysis and the production of antibodies and therapeutic agents 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 LM, Marshall E, Tesfaye M, Silverstein KA, Mori M, Umetsu Y, Otterbach SL, Papareddy R, Dickinson HG, Boutiller K, VandenBosch KA, Ohki S, Gutierrez-Marcos JF (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, Mori M (2006) Inducible virus-mediated expression of a foreign protein in suspension-cultured plant cells. Archives of virology 151:1075-1084

Flower TG, Buffalo CZ, Hooy RM, Allaire M, Ren X, Hurley JH (2020) Structure of SARS-CoV-2 ORF8, a rapidly evolving coronavirus protein implicated in immune evasion. bioRxiv the preprint server for biology https://doi.org/10.1101/2020.08.27.270637

Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR, Rawlings SA, Sutherland A, Premkumar L, Jadi RS, Marrama D, de Silva AM, Frazier A, Carlin AF, Greenbaum JA, Peters B, Krammer F, Smith DM, Crotty S, Sette A (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 CA, Chin AWH, Chu DKW, Mok CKP, Tsang OTY, Yeung YC, Perera R, Poon LLM, Peiris JSM, Valkenburg SA (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, Song H, Huang B, Zhu N, Bi Y, Ma X, Zhan F, Wang L, Hu T, Zhou H, Hu Z, Zhou W, Zhao L, Chen J, Meng Y, Wang J, Lin Y, Yuan J, Xie Z, Ma J, Liu WJ, Wang D, Xu W, Holmes EC, Gao GF, Wu G, Chen W, Shi W, Tan W (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, Mori M (2008) Stable-isotope labeling using an inducible viral infection system in suspension-cultured plant cells. Journal of

biomolecular NMR 42:271-277

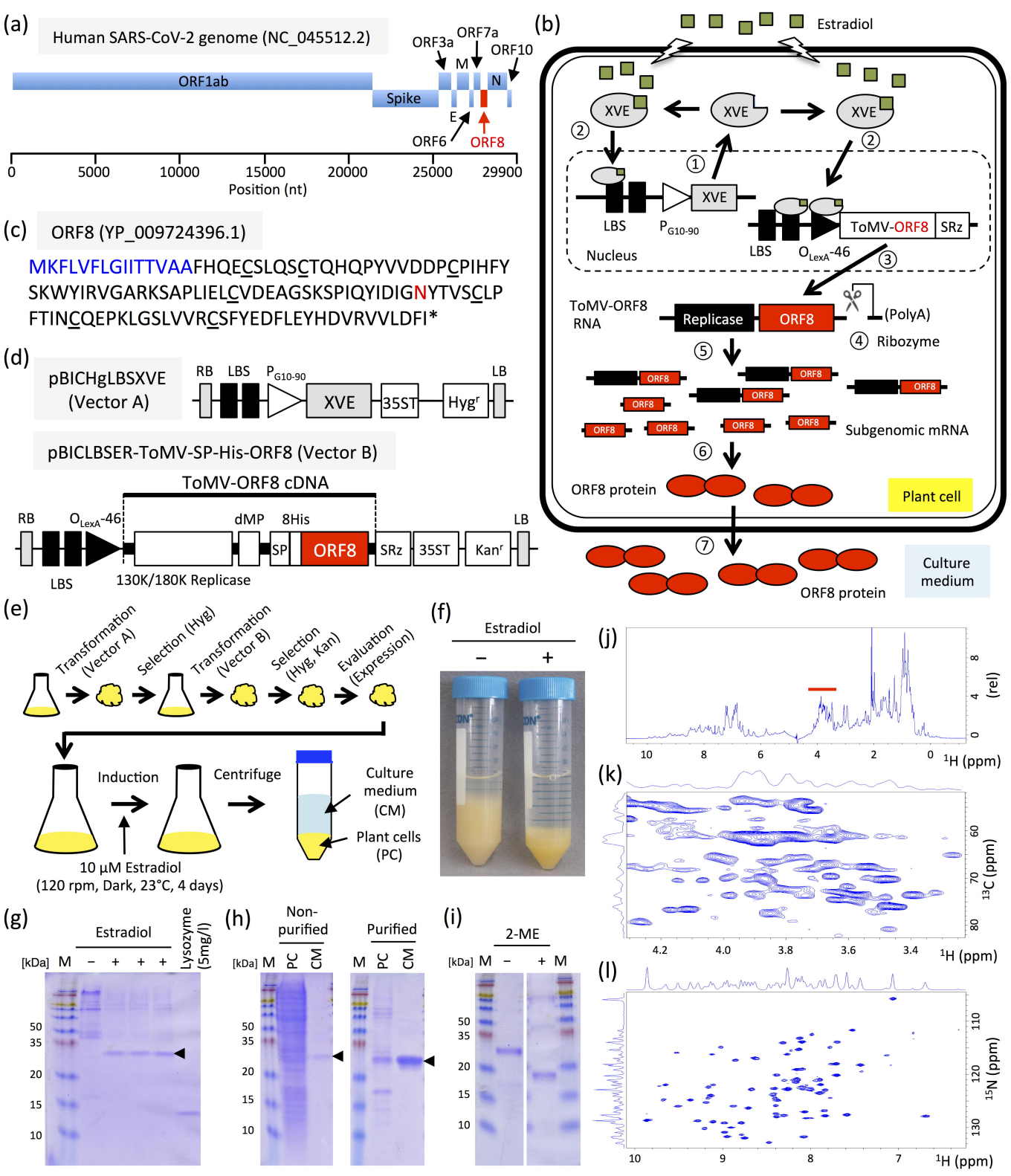
Ohki S, Takeuchi M, Mori M (2011) The NMR structure of stomagen reveals the basis of stomatal density regulation by plant peptide hormones. Nature communications 2:512

Young BE, Fong SW, Chan YH, Mak TM, Ang LW, Anderson DE, Lee CY, Amrun SN, Lee B, Goh YS, Su YCF, Wei WE, Kalimuddin S, Chai LYA, Pada S, Tan SY, Sun L, Parthasarathy P, Chen YYC, Barkham T, Lin RTP, Maurer-Stroh S, Leo YS, Wang LF, Renia L, Lee VJ, Smith GJD, Lye DC, Ng LFP (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, Yu F, Liu J, Liu B, Song Z, Chen J, Pan T, Zhang X, Li Y, Li R, Huang W, Xiao F, Zhang H (2020) The ORF8 protein of SARS-CoV-2 mediates immune evasion through potently downregulating MHC-I. bioRxiv the preprint server for biology https://doi.org/10.1101/2020.05.24.111823

Figure caption

Fig. 1 Production and characterization of ORF8. (a) Schematic representation of SARS-CoV-2 genome. E, M, and N indicate envelope protein, membrane glycoprotein, and nucleocapsid phosphoprotein, respectively. (b) Schematic overview of the chemically inducible ToMV-mediated expression system. Step 1: Constitutive expression of XVE, a transcriptional activator. Step 2: XVE binds to estradiol and is activated. Step 3: Activated XVE binds to $O_{LexA-46}$ 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 representation of ToMV-mediated chemically 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 of 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) ^{1}H NMR spectra of unlabeled ORF8. Red line indicates the glycan signal appearing region. (k) ^{1}H-^{13}C HSQC of unlabeled ORF8 protein. An expanded view of the glycan signal appearing region is shown. (l) ^{1}H-^{15}N HSQC of ^{15}N-labeled ORF8.


(a)
Human SARS-CoV-2 genome (NC_045512.2)
ORF3a
ORF7a
ORF10
M
N
ORF1ab
Spike
E
ORF6
ORF8
0
5000
10000
15000
20000
25000
29900
Position (nt)
(b)
Estradiol
XVE
LBS
P G10-90
ToMV-ORF8
SRz
Nucleus
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
O LexA -46
dMP
8His
SP
ORF8
SRz
35ST
Kan r
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
–
+
(j)
(rel)
1H (ppm)
(k)
13C (ppm)
1H (ppm)
(l)
15N (ppm)
1H (ppm)
(g)
Estradiol
Lysozyme (5mg/l)
[kDa]
M
(h)
Non-purified
Purified
PC
CM
(i)
2-ME