Virus Genes 27:2, 189–196, 2003

High-level Expression of the *ORF6* Gene of Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) in *Pichia pastoris*

PING QIAN,[1] XIANGMIN LI,[1] GUANGZHI TONG[2] & HUANCHUN CHEN[1,*]

[1]*Laboratory of Animal Virology, College of Vet. Med. Science, Huazhong Agricultural University, Wuhan, 430070*
[2]*National Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute, Harbin 150001, P.R. China*

Received April 14, 2003; Accepted June 3, 2003

Abstract. High-level expression of the *ORF6* gene of porcine reproductive and respiratory syndrome virus (PRRSV) has been proved very difficult. In this work, we cloned and sequenced the *ORF6* gene of PRRSV and found that it could not be expressed in *Pichia pastoris* strain GS115. Then, the *ORF6* gene was modified and synthesized based on the codon bias, poly (A) signal of yeast expression system and secondary structure of 5′-end mRNA of foreign gene. The modified gene was inserted into the yeast expression vector pPICZαA, induced and expressed by the same methods. The recombinant protein with a molecular mass of approximately 23 kDa was screened by SDS-PAGE and identified by Western blot with convalescent sera of animals infected with CH-1a strain of PRRSV. The results indicated that it was similar to the native protein. The expression level of the recombinant protein could attain 2.0 g/L. In the meanwhile, the optimal conditions for expression were determined. It provides an additional means for studying the structural and functional characteristics of PRRSV *ORF6* gene.

Key words: codon bias, gene modification, high-level expression, *ORF6* gene, *Pichia pastoris*, PRRSV

Introduction

Porcine reproductive and respiratory syndrome virus (PRRSV) is a member of the *Arteriviridae* family which includes lactate dehydrogenase-elevating virus (LDV) of mice, equine arteritis virus (EAV), and simian hemorrhagic fever virus (SHFV) [1,3,5,18]. PRRSV is a relatively small, enveloped virus containing single-stranded positive-sense RNA genome of approximately 15 kb [5,18]. The viral genome contains 8 open reading frames (ORFs) that are transcribed in cells as a nested set of subgenomic mRNAs [5,18,19]. The ORF1a and ORF1b situated at the 5′ end of the genome represent nearly 75% of the viral genome and encode proteins with apparent RNA-dependent RNA polymerase activities [18,20].

* Author for all correspondence.
E-mail: qianping004@sina.com

In addition to the polymerase, the genomic sequence encodes 4 envelope glycoproteins named GP_2 (ORF2), GP_3 (ORF3), GP_4 (ORF4), and GP_5 (ORF5), as well as an 18–19 kDa unglycosylated membrane protein M (ORF6) and a 15 kDa nucleocapsid protein N (ORF7) [5,15,18,19,27]. The M protein contains 1–2 putative glycosylation sites and 3 highly hydrophobic regions in its N-terminal half [15,17], which are assumed to represent potential membrane spanning domains by analogy to the coronavirus M protein [23]. Another, 18 amino acid residues of the protein is probably exposed on the virion surface which relates to viral absorbtion and propagation [1]. The M protein is the most conserved structural protein between North American and European isolates. It is highly antigenic and can elicit a detectable antibody response as early as 10 days following infection [10,14,16,32,33]. Therefore it is suggested that the recombinant protein can be

used as antigen for sera test and subunit vaccine. Because the structure and function of M protein are still unclear so far, it is necessary to prepare a large quantity of the recombinant M protein by the gene engineering method for studying the protein. *Pichia pastoris* has been developed into an expression system for high-level expression of recombinant proteins. It not only shares many of the advantages of higher eukaryotic expression systems such as protein processing, protein folding, and posttranslational modification, but is also as easily manipulated as *E. coli* [7,8,11]. Many heterologous proteins have been highly expressed in *P. pastoris* such as TNF, EGF [4,28].

In the present study, we extracted the PRRSV total RNA, cloned the *ORF6* gene by RT-PCR and sequenced the whole gene. Meanwhile, we inserted the native and modified *ORF6* gene into the pPICZαA by the correct ORF and transformed the appropriate *P. pastoris* strain GS115. The modified gene was expressed on the higher level and more activity of secreted expression, but the native gene could not be expressed.

Materials and Methods

Virus and Strains

PRRSV CH-1a strain (obtained from Harbin Veterinary Research Institute, China) was propagated in the African green monkey kidney permissive cell line (MARC-145). Cells were cultured in Eagle's minimum essential medium (EMEM, Gibco, USA). *Pichia* expression vector pPICZαA and *pichia* strain GS115 came from invitrogen.

Isolation of PRRSV RNA and Cloning of ORF6 *Gene*

The PRRSV total RNA was isolated from MARC-145 cells infected with PRRSV CH-1a strain according to the manufacturer's recommendation (TaKaRa, Dalian). Based on the published complete nucleotide sequences of PRRSV VR2332 strain (GenBank accession No. UOO153), a pair of primers was designed for the cloning of *ORF6* gene. Sequences of the primers were as follows: the sense primer was 5′-GAAGGATCCGTTTCAGCGGAACAATGG-3′ and the antisense primer was 5′-TGGAGCGGCCGCAG-CTGATTGACTGGCTGG-3′. Reverse transcription (RT) was performed as follows: 10 μl total RNA, 1 μl corresponding antisense primer (10 μmol/L), 10 μl 5 × RT-buffer, 1 μl RNase inhibitor (TaKaRa, Dalian), 16 μl 2.5 mmol/L dNTPs, 2 μl AMV Rtase (TaKaRa, Dalian) were mixed and incubated at 25°C for 10 min, 42°C for 1 h, and finally chilled on ice. To amplify the *ORF6* gene, a mixture of the first stand cDNA, the corresponding sense and antisense primers, EX-Taq DNA polymerase, 10 × PCR buffer, dNTPs, and sterile distilled water were added to a total volume of 50 μl. PCR was performed in Thermocycler with the following conditions: 94°C 4 min; 94°C 1 min, 56°C 1 min, 72°C 1.5 min, 35 cycles; extension at 72°C for 10 min. The purified specific fragment was cloned into pMD18-T vector and identified by PCR and specific restriction enzyme digestion. The positive recombinant plasmids were sequenced.

Modification of ORF6 *Gene and Construction of Recombinant Yeast Transfer Vector*

Based on the *P. pastoris* codon bias, poly (A) signal of yeast expression system and secondary structure of 5′-end of foreign gene [12,25,26], the *ORF6* gene was synthesized, *Eco*RI and *Not*I sites were added to both ends of the modified *ORF6* gene. The modified *ORF6* gene and pPICZαA were digested with the restriction enzyme *Eco*RI and *Not*I, respectively. The interesting fragments were recovered and ligated. In the meantime, the native *ORF6* gene was also inserted into the pPICZαA vector within correct ORF by the same methods.

Yeast Transformation and Identification

Yeast vector pPICZαA/pPICZαA-ORF6/pPICZαA-M were digested with *Sac*I enzyme and transformed competent *pichia* cells GS115 according to Easy-select *pichia* Expression kit instruction manual. Positive recombinant yeasts were selected on the Yeast Extract Peptone Dextrose Medium with Sorbitol (YPDS) plates with 100 μg Zeocin and the mut phenotype was confirmed on the selection Minimal Methanol Medium and Histidine (MMH) plates and Minimal Dextrose Medium and Histidine (MDH) plates. At the same time, colonies were picked into Yeast Extract Peptone Dextrose (YPD) medium and incubated overnight. The yeasts were lysed with Zymolyase (Sigma, USA) and gently

extracted with an equal volume of pheno: chloroform (1 : 1, v/v) followed by an equal volume of chloroform: isoamyl alcohol (24 : 1, v/v). The aqueous layer was split into microcentrifuge tubes and added 1/2 volume of ammonium acetate and 2 volumes of ethanol to each tube, centrifuged at 10,000*g* for 20 min at 4°C. The pellets were washed once with 1 ml of 75% ethanol. After brief air-drying, the pellets were resuspended with 50 µl of TE buffer (5 mmol/L Tris-HCl, pH 7.5, 1 mmol/L EDTA) each tube for PCR amplification. PCR conditions were: 95°C 4 min; 94°C 1 min, 55°C 1 min, 72°C 1.5 min, 35 cycles; extension at 72°C for 10 min.

Induction and Expression of the Foreign Gene

A single colony was incubated in 25 ml of Buffered Glycerol-complex Medium (BMGY) in a 250 ml baffled flask, grown at 28–30°C in a shaking incubator for approximately 16–18 h until OD_{600} of the culture reached 2–6. The cells were harvested and the cell pellets were resuspended to an OD_{600} of 1.0 in Buffered Methanol-complex Medium (BMMY) and cultured at 28–30°C for 4 days. Methanol (100%) was added to a final concentration of 0.5% every 24 h. 1 ml of the expression culture was moved to a 1.5 ml micro centrifuge tube, centrifuged at maximum speed for 2–3 min at room temperature and the supernatant was transferred to a separate tube and analyzed for protein expression by Coomassie Blue staining SDS-PAGE and Western blot.

SDS-PAGE and Western Blot

SDS-PAGE and Western blot were carried out by the procedure of Sambrook [24] Five Microliter of the supernatant was lysed in 5.0 µl of 2 × SDS-PAGE loading buffer (0.2% bromophenol, 20% glycerol, 4% SDS, 100 mmol/L Tris-HCl, pH 6.8, 200 mmol/L dithiothreitol) by boiling for 10 min. Then samples were separated in the 15% polyacrylamide gel. Mock-yeast supernatant was similarly processed. Following SDS-PAGE, the gels were stained by Coomassie Blue or the proteins were transferred to nitrocellulose, using a semidry apparatus (Bio-Rad, Hercules, CA). The nitrocellulose blots were saturated with PBS containing 3% (W/V) BSA and 0.05% Tween20 overnight at 4°C. Recombinant proteins were incubated with porcine anti-PRRSV serum (1 : 50 dilution) in blocking solution for 1 h at 37°C, and followed by washing three times of 15 min each with PBS containing 0.05% Tween20. The blots were then incubated with alkaline phosphatase conjugated rabbit anti-swine IgG for 1 h at 37°C. After washing, the blots were developed using 5′-bromo-4-chloro-3-indolyphosphate p-Toluidine salt (BCIP) and nitroblue tetrazolium chloride (NBT) (Gibco, USA).

Determination of Optimal Expression

Samples collected at different time points were analyzed for the concentration of both the total protein and interesting proteins to determine the optimal harvesting time. The total protein of each sample was measured, and the concentration of the interesting protein was calculated by scanning the Coomassie-stained SDS-PAGE and determined the optimal time post-induction to harvest. Based on the optimal harvesting time post-induction, the recombinant *pichia* strains were induced by methanol with a final concentration of 0.125%, 0.25%, 0.5%, 1.0%, respectively, to determine the optimal methanol concentration for inducible expression.

Results

Design and Chemical Synthesis of the ORF6 *Gene*

Nucleotide sequencing analysis indicated that the inserted fragment contained the entire coding sequence of PRRSV *ORF6* gene. Furthermore, yeast has an optimal codon bias for the high-level expression. Yeast bias codons were scattered in the *ORF6* gene. Since further site-directed mutagenesis was not feasible, we decided to chemically synthesize the entire *ORF6* gene. In the design of the synthetic *ORF6* gene, three aspects were considered. First, G + C content was increased and special attention was paid to eliminate A + T-rich runs for avoiding transcript stop. Second, the redundancy of the genetic code was used to insert unique *Kpn*I site and identify modified or native gene. Third, the synthetic gene was designed to have an optimal codon bias for high-level expression in yeast. No amino acid changes were made compared to the native sequence. In addition, *Eco*RI and *Not*I sites were added to both

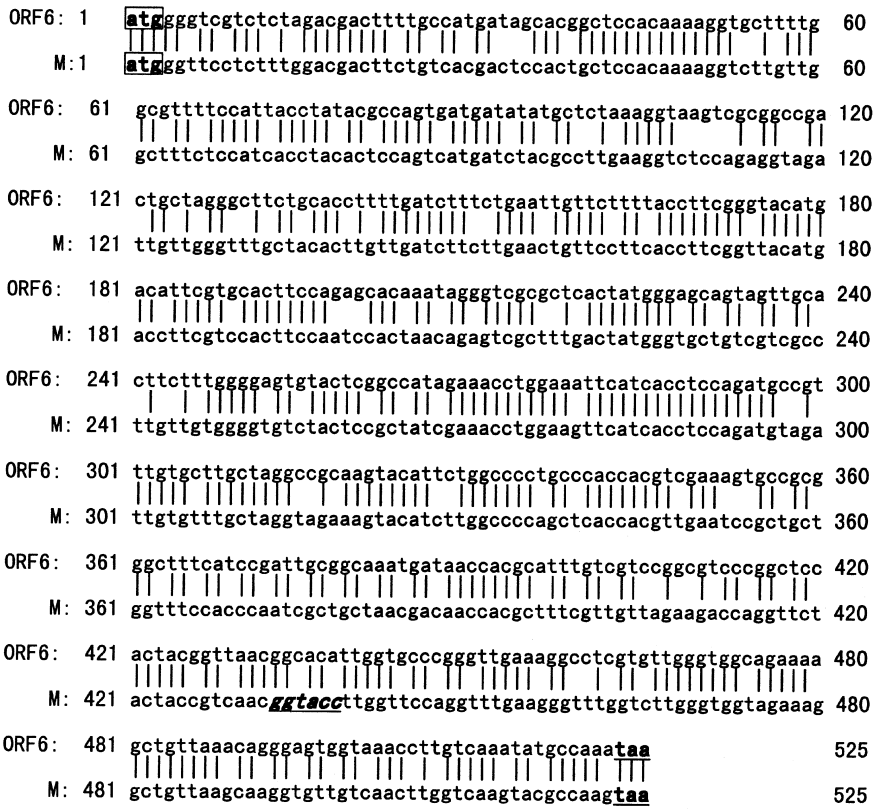

Fig. 1. Comparison of nucleotide sequence of native and modified *ORF6* gene. *ORF6* represented native *ORF6* gene, which was cloned and sequenced (see Materials and Methods); M represented modified *ORF6* gene. atg and ***taa*** were initiative and terminal codons, respectively. The *KpnI* restriction enzyme ***ggtacc*** was introduced into the modified *ORF6* gene which no animo acid was changed.

ends of the gene, respectively. Native and modified *ORF6* genes were shown in Fig. 1.

Strategy of Construction of the Yeast Expression Vector

Modified *ORF6* gene was inserted into the yeast expression vector pPICZαA by the unique *Eco*RI and *Not*I sites with the correct ORF to construct the yeast transfer vector. Native gene was subcloned into the pPICZαA vector by the same methods Fig. 2

Identification and Transformation of the Recombinant yeast

Since Zeocin gene exists in recombinant yeast but not in the parent strain GS115, 10 transforms were picked up to the YPD culture with 25 μg/ml Zeocin and grown at 28–30°C overnight. The total DNA of yeast was extracted and the interesting fragments were amplified by 2 pairs of primers: 5′AOX1/3′AOX1 and α-Factor/3′AOX1. Results showed that native and modified *ORF6* gene had been correctly integrated into the yeast genome.

Detection of Induced Expression Production of the Recombinant Yeast

A distinct polypeptide with a molecular mass (Mr) of approximately 23 kDa was detected in the recombinant supernatant and no other polypeptides were secreted (Fig. 3A); however the native *ORF6* gene could not be expressed (data not shown). Recombinant M protein had immunogenic activity by analysis of Western blot (Fig. 3B).

Determination of Optimal Expression

Five Microliter of the supernatant at different time points were analyzed by SDS-PAGE and the optimal

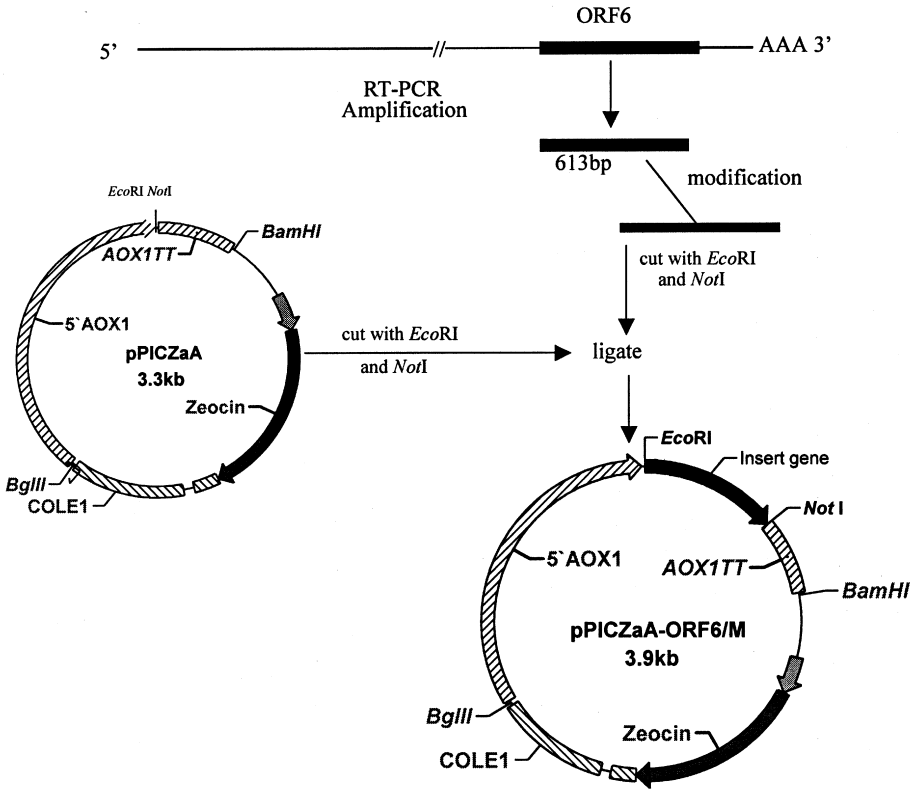

Fig. 2. Schematic representation of the cloning strategy. Total RNA from PRRSV CH-1a strain-infected cells was used in the RT-PCR reaction to amplify *ORF6* gene and the *ORF6* gene was modified (see Materials and Methods). α-factor sequence was located between 5′AOX1 and multiple cloning site (MCS). 5′AOX1 = A 942 bp fragment containing the AOX1 promoter that allows methanol-inducible, high-level expression in *pichia*; AOXTT = AOX1 transcription termination. It is native transcription termination and polyadenylation signal from *AOX1* gene (~260 bp) that permits efficient 3′mRNA processing including polyadenylation for increased mRNA stability.

induction time was found to be approximately 84 h. Protein band of lane 4 was about 84% of the total protein and the level of expression was approximately 2.0 g/L (Fig. 4). Based on the optimal inducible time, the recombinant yeast was induced by methanol with a final concentration of 0.125%, 0.25%, 0.5%, 1.0%, respectively, and the results showed that the optimal methanol concentration was about 0.5%. The recombinant yeast could almost not be expressed with the final concentration of 0.125% and 0.25%. The expression levels were the same with 1% methanol concentration and 0.5% methanol concentration (Fig. 5).

Discussion

It is necessary to express M protein *in vitro* in order to research its function and develop possibility as subunit vaccine. At present, GP3, GP5, and N proteins were expressed, but M protein could not be expressed in the baculovirus system [22] or its expression level was very low because the M protein was a highly hydrophobic protein [13]. As a eukaryote, *P. pastoris* has been developed as an excellent host for the production of foreign proteins [8]. The organism has the potential for high level of expression, efficient secretion, easily grown to very high cell densities and the advantages of higher eukaryotic expression systems such as protein processing, protein folding, posttranslational modification while being as easily as to be manipulated as *E. coli* or *Saccharomyces cerevisiae* [4,7,8,11,28]. It is faster, easier, and less expensive to be used than other eukaryotic expression systems such as baculovirus or mammalian tissue culture. As a yeast expression system, it uses the promoter from the methanol-induced alcohol oxidase gene—Aoxl,

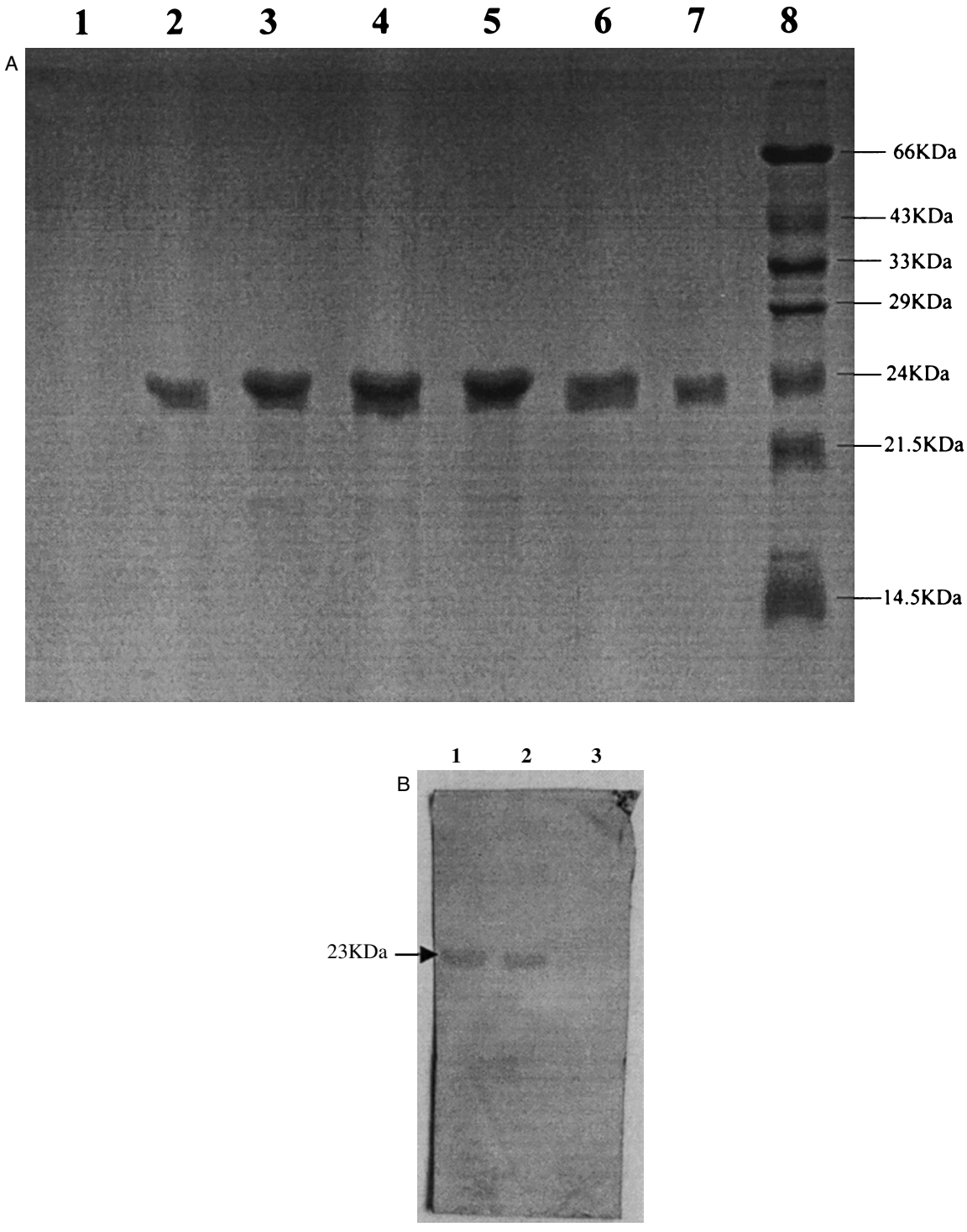

Fig. 3. Analysis of the recombinant proteins by the 15% SDS-PAGE gel. (A) Coomassie blue straining of SDS-PAGE from the supernant of recombinant yeasts collected at different times. Mocked-recombinant yeast (lane 1), collected at 108, 96, 84, 72, 60 h p.i., respectively (lane 2–7), protein marker (lane 8). Molecular weight markers in kDa were shown on the right. (B) Western blot of the supernant of the recombinant yeast collected at 84, 96 h p.i., respectively (lane 1–2), mocked-recombinant yeast (lane 3).

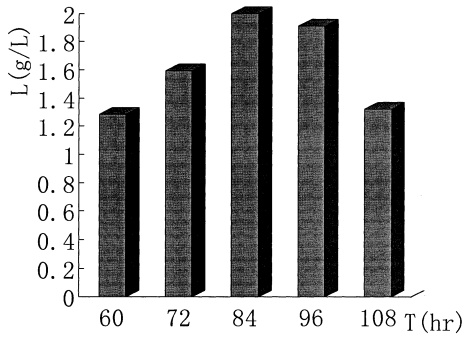

Fig. 4. Optimal induction time of the recombinant yeast. Total protein concentration of 5 μl supernatant were calculated and SDS-PAGE was scanned, analysis percent of the interesting protein, and calculated the concentration of these proteins. L, expression level; T, induce time.

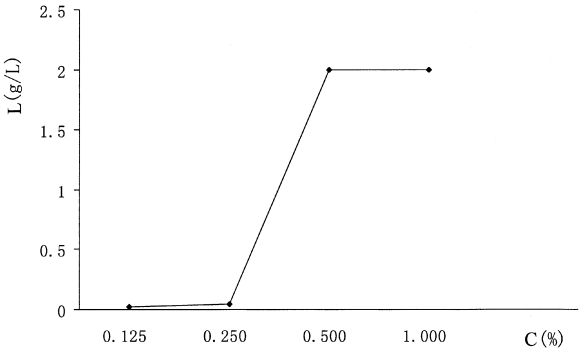

Fig. 5. Optimal methanol concentration. Recombinant yeast was induced with a final methanol concentration of 0.125%, 0.25%, 0.5%, 1.0%, respectively and collected supernatant at 84 h. L, expression level; C, a final methanol concentration.

shares the advantages of molecular and genetic manipulation with *saccharomyces* and has the added advantage of 10–100-fold higher heterologous protein expression level. These features make *pichia* very useful as a protein expression system [2,11,31]. In this paper, we have highly expressed the M protein with the yeast expression system, while native *ORF6* gene was not expressed (data not shown).

We found that there were many factors drastically influencing protein production in this system including the copy number of the expression cassette, site, and model of chromosomal integration of expression cassette, mRNA 5′- and 3′-untranslated regions, translational start codon (AUG) context, A + T composition of cDNA, transcriptional and translational blocks, nature of secretion signal, endogenous protease activity, host strain physiology, medium and growth conditions, and fermentation parameters, etc. [29]. In addition, an apparent difference of bias between yeast and bacteria was found by analyzing the codon characteristics of 110 yeast's genes. The usage frequency of codon AGA and CGC were 86.6% and 1.25%, respectively, and CGA and CGG were not used within the six codons of arginine [10,25,26]. So codons of arginine in the *ORF6* gene were changed into AGA. Because codons of other amino acids were similar to arginine, we decided to modify the whole *ORF6* gene. All codons of *ORF6* gene were mutated according to yeast bias codons. In the meanwhile, rich ATT region was removed to avoid the undesirable termination signal of yeast. Results confirmed that the modified *ORF6* gene was highly expressed (approximately 2.0 g/L), indicating that *ORF6* gene was successfully modified.

In this study, yeast-expressed protein M (Fig. 3A and B) was about 23 kDa which is 4 kDa larger than the calculated molecular mass of 19 kDa. It is possibly glycoslated in yeast system because *ORF6* gene of CH-1a strain has two latent glycoslated sites that are un-functional [30]. So the recombinant M protein was cut with Endoglycosylase H (Roche, USA), but M protein mass was not reduced (data not shown), Drew et al. (1995) also identified a non-glycosylated M protein in the transcription and translation studies of *ORF6* gene [9]. Our studies further confirmed that the M protein expressed in the yeast system was a non-glycosylated protein.

Recombinant baculovirus-expressed M protein [13] was also larger than M detected in cell lysates [21], in purified LV particles [19] or by *in vitro* translation of the Canadian and European strain of PRRSV [15,19]. But *in vitro* translation of the LV M protein from infected cell lysate yielded a peptide of Mr 21.5 kDa [19] that was similar to the yeast-expressed M protein reported herein (Fig. 3A and B).

In conclusion, a secreted yeast engineering strain was constructed, and the optimal induction conditions were determined for M protein of PRRSV. It is very convenient to separate and purify the M protein and product a large amount of the recombinant M protein by the fermentation and it provides a tool for studying its structural and functional characterization.

Acknowledgment

This research was supported by the grants for the 973 National Program (strategic grant G1999011902). We thank Dr. Guo Aizhen for critical modification of the manuscript.

References

1. Andreyev V.G., Scherbakov A.G., Pylnov V.A., and Gvsev A.A., in *Proceedings of the International Symposium on PRRS and Aujeszky's Disease, Genetic heterogeneity of PRRSV in Russia*, Ploufragan, France, (in press) 1999, pp. 211–212.
2. Buckholz R.G. and Gleeson M.A., Biotechnology *9*, 1067–1072, 1991.
3. Cavanagh D., Arch Virol *142*, 629–633, 1997.
4. Clare J.J., Romanos M.A., Rayment F.B., Rowedder J.E., Smith M.A., Sreekrishna K., and Henwood C.A., Gene *105*, 205–212, 1991.
5. Conzelmann K.K., Visser N., Van Wvensel P., and Thiel H.J., Virology *193*, 329–339, 1993.
6. Cregg J.M., Vedvick T.S., and Raschke W.C., Biotechnology *11*, 905–910, 1993.
7. Cregg J.M., Barringer K.J., Hessler A.Y., and Madden K.R., Mol Cell Biol *5*, 3376–3385, 1985.
8. Cregg J.M., Madden K.R., Barringer K.J., Thill G.P., and Stillman C.A., Mol Cell Biol *9*, 1316–1323, 1989.
9. Dea S., Gagnon C.A., Mardassi H., Pirzazdeh B., and Rogan D., Arch Virol *145*, 659–688, 2000.
10. Gagnon C.A. and Dea S., Can J Vet Res *62*, 110–116, 1998.
11. Gellissen G. and Hollenberg C.P., Gene *190*, 87–97, 1997.
12. Bennetzen J.L. and Hall B.D., J Biol Chem *257*, 3026–3031, 1982.
13. Kreutz L.C. and Mengeling W.L., Vet Microbiol *59*, 1–13, 1997.
14. Loemba H.D., Mounir S., Mardassi H., Archambault D., and Dea S., Arch Virol *141*, 751–761, 1996.
15. Mardassi H., Mounir S., and Dea S., Arch Virol *140*, 1405–1418, 1995.
16. Meng X.J., Paul P.S., Halbur P.G., and Lum M.A., Arch Virol *140*, 745–755, 1995.
17. Mengeling W.L., Lager K.M., and Vorwald A.C., J Vet Diagn Invest *7*, 3–16, 1995.
18. Meulenberg J.J., Hulst M.M., de Meijer E.J., Moonen P.L., den Besten A., de Kluyver E.P., Wensvoort G., and Moormann R.J., Virology *192*, 62–72, 1993.
19. Meulenberg J.J., Petersen-den Besten A., de Kluyver E.P., Moormann R.J., Schaaper W.M., and Wensvoort G., Virology *206*, 155–163, 1995.
20. Meulenberg J.J., Bos-de Ruijter J.N., van de Graaf R., Wensvoort G., and Mooramann R.J., J Virol *72*, 380–387, 1998.
21. Nelson E.A., Christopher-Hennings J., and Benfield D.A., J Vet Diagn Invest *6*, 410–415, 1994.
22. Plana Duran J., Climent I., Sarraseca J., Urniza A., Cortes E., Vela C., and Casal J.I., Virus Genes *14*, 19–29, 1997.
23. Rottier P.J., Welling G.W., Welling-Wester S., Niesters H.G., Lenstra J.A., and van der Zeijst B.A., Biochemistry *25*, 1335–1339, 1986.
24. Sambrook J., Fritsch E.F., and Maniatis T., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.
25. Scorer C.A., Buckholz R.G., Clare J.J., and Romanos M.A., Gene *136*, 111–119, 1993.
26. Sharp P.M., Tuohy T.M., and Mosurski K.R., Nucleic Acids Res *14*, 35125–35143, 1986.
27. Snijder E.J., and Meulenberg J.J., J Gen Virol *79*, 961–979, 1998.
28. Sreekrishna K., Nelles L., Potenz R.H., Cruze J., Mazzaferro P., Fish W., Holden K., Phelps D., and Wood P., Biochemistry *28*, 4117–4125, 1989.
29. Sreekrishna K., Brankamp R.G., Kropp K.E., Blankenship D.T., Tsay J.T., Smith P.L., Wierschke J.D., Subramaniam A., and Birkenberger L.A., Gene *190*, 55–62, 1997.
30. Tong G.Z., Qiu H.J., and Zhou Y.J., Adv Nat Sci *2*, 147–153, 2000.
31. Tschopp J.F., Brust P.F., Cregg J.M., Stillman C.A., and Gingeras T.R., Nucleic Acids Res *15*, 3859–3876, 1987.
32. Yang S.X., Kwang J.K., and Laegreid W., Arch Virol *143*, 601–612, 1998.
33. Zeman D., Neiger R., Yaeger M., Nelson E., Benfield D., Leslie-Steen P., Thomson J., Miskimins D., Daly R., and Minehart M., J Vet Diagn Invest *5*, 522–528, 1993.