Analysis of transglycosylation product

For each 10-μL reaction, purified recombinant Endo-Om (4.2 μU) or heat-inactivated enzyme was incubated with 100 mM sodium acetate buffer (pH 6.0) containing 2 mM NGA2-Asn-Fmoc with or without 50 mM pNP-Glc. After incubation at 30°C for 3 h, the reaction was terminated by heating to 95°C for 5 min. The reaction mixture was subjected to HPLC with a Shodex Asahipak NH2P-50 4E column using the same method as described above. Elution was monitored at 274 nm and transglycosylation product was recovered. The fractions were lyophilized and dissolved in Milli-Q water to yield a concentration of ~10 μM. The mass spectra of each sample were acquired with a MALDI-QIT-TOF mass spectrometer (AXIMA-QIT; Shimadzu Corp., Kyoto, Japan), using a matrix of 2,5-dihydroxybenzoic acid dissolved to 10 mg/mL in 40% ethanol.

Digestion of N-glycan on denaturing or nondenaturing glycoproteins

Four micrograms of RNase B (New England Biolabs, Ipswich, MA) or 1 μg of human holo-transferrin (Sigma-Aldrich, St. Louis, MO) was used as a substrate. For each 10-μL reaction (denaturing conditions), substrate protein was denatured at 95°C for 10 min with 25 mM Tris–HCl buffer (pH 8.5) containing 0.5% SDS, 0.75% 2-mercaptoethanol, 1% Triton X-100 and 10 mM EDTA. For nondenaturing conditions, substrate protein was adjusted to a total volume of 10 μL with 1% Triton X-100 and 10 mM EDTA. The N-glycan digestion was performed as follows: the 10-μL substrate protein solution was combined with 1 μL of purified recombinant Endo-Om (4.49 μU/μL), 4 μL of 250 mM sodium acetate buffer (pH 5.3) containing 1.25 mM NaCl, and (if applicable) 1 μL of sialidase (10 mU/μL; Roche Diagnostics K.K.), and incubated at 30°C. Commercially obtained Endo-H (100 U/reaction with 1 × G5 reaction buffer; New England Biolabs) or PNGase F (0.2 U/reaction; Roche Diagnostics K.K.) was used as positive control and incubated at 37°C. Following overnight incubation (up to 26 h), reactions were terminated by the addition of 5 μL of 4 × SDS–PAGE loading buffer (Wako Pure Chemical Industries, Ltd., Osaka, Japan) and subsequent heating to 95°C for 5 min. Samples were subjected to SDS–PAGE and stained with Coomassie brilliant blue G-250.

Acknowledgements

Abbreviations

CAZy, carbohydrate-active enzymes; EDTA, ethylenediaminetetraacetic acid; EGTA, ethylene glycol tetraacetic acid; ENGase, endo-β-N-acetylglucosaminidase; ER, endoplasmic reticulum; Fmoc, 9-fluorenylmethyloxycarbonyl; fOSs, free oligosaccharides; GH, glycoside hydrolase; GlcNAc, N-acetylglucosamine; HPLC, high-performance liquid chromatography; MALDI-QIT-TOF MS, matrix-assisted laser desorption ionization-quadrupole ion trap-time-of-flight mass

spectrometry; Mr, molecular mass; PA, pyridylamino; PCR, polymerase chain reaction; PMSF, phenylmethylsulfonyl fluoride; PNGase, peptide:N-glycanase; pNP-Glc, 4-nitrophenyl-β-D-glucopyranoside; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis.

References

Abbott DW, Macauley MS, Vocadlo DJ, Boraston AB. 2009. *Streptococcus pneumoniae* endohexosaminidase D, structural and mechanistic insight into substrate-assisted catalysis in family 85 glycoside hydrolases. *J Biol Chem.* 284:11676–11689.

Chantret I, Frénoy JP, Moore SE. 2003. Free-oligosaccharide control in the yeast *Saccharomyces cerevisiae*: Roles for peptide:N-glycanase (Png1p) and vacuolar mannosidase (Ams1p). *Biochem J.* 373:901–908.

Chantret I, Moore SE. 2008. Free oligosaccharide regulation during mammalian protein N-glycosylation. *Glycobiology.* 18:210–224.

Collin M, Olsén A. 2001. EndoS, a novel secreted protein from *Streptococcus pyogenes* with endoglycosidase activity on human IgG. *EMBO J.* 20:3046–3055.

Fischl RM, Stadlmann J, Grass J, Altmann F, Léonard R. 2011. The two endo-β-N-acetylglucosaminidase genes from *Arabidopsis thaliana* encode cytoplasmic enzymes controlling free N-glycan levels. *Plant Mol Biol.* 77:275–284.

Fu D, van Halbeek H. 1992. N-glycosylation site mapping of human serotransferrin by serial lectin affinity chromatography, fast atom bombardment-mass spectrometry, and ^{1}H nuclear magnetic resonance spectroscopy. *Anal Biochem.* 206:53–63.

Fujita K, Kobayashi K, Iwamatsu A, Takeuchi M, Kumagai H, Yamamoto K. 2004. Molecular cloning of *Mucor hiemalis* endo-β-N-acetylglucosaminidase and some properties of the recombinant enzyme. *Arch Biochem Biophys.* 432:41–49.

Fujita K, Takami H, Yamamoto K, Takegawa K. 2004. Characterization of endo-β-N-acetylglucosaminidase from alkaliphilic *Bacillus halodurans* C-125. *Biosci Biotechnol Biochem.* 68:1059–1066.

Haga Y, Totani K, Ito Y, Suzuki T. 2009. Establishment of a real-time analytical method for free oligosaccharide transport from the ER to the cytosol. *Glycobiology.* 19:987–994.

Hamaguchi T, Ito T, Inoue Y, Limpaseni T, Pongsawasdi P, Ito K. 2010. Purification, characterization and molecular cloning of a novel endo-β-N-acetylglucosaminidase from the basidiomycete, *Flammulina velutipes*. *Glycobiology.* 20:420–432.

Hirayama H, Seino J, Kitajima T, Jigami Y, Suzuki T. 2010. Free oligosaccharides to monitor glycoprotein endoplasmic reticulum-associated degradation in *Saccharomyces cerevisiae*. *J Biol Chem.* 285:12390–12404.

Huang W, Giddens J, Fan SQ, Toonstra C, Wang LX. 2012. Chemoenzymatic glycoengineering of intact IgG antibodies for gain of functions. *J Am Chem Soc.* 134:12308–12318.

Kato T, Fujita K, Takeuchi M, Kobayashi K, Natsuka S, Ikura K, Kumagai H, Yamamoto K. 2002. Identification of an endo-β-N-acetylglucosaminidase gene in *Caenorhabditis elegans* and its expression in *Escherichia coli*. *Glycobiology.* 12:581–587.

Kato T, Kitamura K, Maeda M, Kimura Y, Katayama T, Ashida H, Yamamoto K. 2007. Free oligosaccharides in the cytosol of *Caenorhabditis elegans* are generated through endoplasmic reticulum-Golgi trafficking. *J Biol Chem.* 282:22080–22088.

Kuroda K, Kitagawa Y, Kobayashi K, Tsumura H, Komeda T, Mori E, Motoki K, Kataoka S, Chiba Y, Jigami Y. 2007. Antibody expression in protease-deficient strains of the methylotrophic yeast *Ogataea minuta*. *FEMS Yeast Res.* 7:1307–1316.

Kuroda K, Kobayashi K, Kitagawa Y, Nakagawa T, Tsumura H, Komeda T, Shinmi D, Mori E, Motoki K, Fuju K, et al. 2008. Efficient antibody production upon suppression of O mannosylation in the yeast *Ogataea minuta*. *Appl Environ Microbiol.* 74:446–453.

Kuroda K, Kobayashi K, Tsumura H, Komeda T, Chiba Y, Jigami Y. 2006. Production of Man$_5$GlcNAc$_2$-type sugar chain by the methylotrophic yeast *Ogataea minuta*. *FEMS Yeast Res.* 6:1052–1062.

Ling Z, Suits MD, Bingham RJ, Bruce NC, Davies GJ, Fairbanks AJ, Moir JW, Taylor EJ. 2009. The X-ray crystal structure of an *Arthrobacter protophormiae* endo-β -N-acetylglucosaminidase reveals a (β/α)$_8$ catalytic

domain, two ancillary domains and active site residues key for transglyco-sylation activity. *J Mol Biol.* 389:1–9.

Moore SE, Bauvy C, Codogno P. 1995. Endoplasmic reticulum-to-cytosol transport of free polymannose oligosaccharides in permeabilized HepG2 cells. *EMBO J.* 14:6034–6042.

Muramatsu H, Tachikui H, Ushida H, Song X, Qiu Y, Yamamoto S, Muramatsu T. 2001. Molecular cloning and expression of endo-β-N-acetylglucosaminidase D, which acts on the core structure of complex type asparagine-linked oligosaccharides. *J Biochem.* 129:923–928.

Renzi F, Manfredi P, Mally M, Moes S, Jenö P, Cornelis GR. 2011. The N-glycan glycoprotein deglycosylation complex (Gpd) from *Capnocytophaga canimorsus* deglycosylates human IgG. *PLoS Pathog.* 7: e1002118.

Spiro RG. 2004. Role of N-linked polymannose oligosaccharides in targeting glycoproteins for endoplasmic reticulum-associated degradation. *Cell Mol Life Sci.* 61:1025–1041.

Stals I, Samyn B, Sergeant K, White T, Hoorelbeke K, Coorevits A, Devreese B, Claeyssens M, Piens K. 2010. Identification of a gene coding for a deglycosylating enzyme in *Hypocrea jecorina*. *FEMS Microbiol Lett.* 303:9–17.

Suzuki T, Funakoshi Y. 2006. Free N-linked oligosaccharide chains: Formation and degradation. *Glycoconj J.* 23:291–302.

Suzuki T, Park H, Kitajima K, Lennarz WJ. 1998. Peptides glycosylated in the endoplasmic reticulum of *yeast* are subsequently deglycosylated by a soluble peptide: N-glycanase activity. *J Biol Chem.* 273:21526–21530.

Suzuki T, Yano K, Sugimoto S, Kitajima K, Lennarz WJ, Inoue S, Inoue Y, Emori Y. 2002. Endo-β-N-acetylglucosaminidase, an enzyme involved in processing of free oligosaccharides in the cytosol. *Proc Natl Acad Sci USA.* 99:9691–9696.

Takegawa K, Nakoshi M, Iwahara S, Yamamoto K, Tochikura T. 1989. Induction and purification of endo-β-N-acetylglucosaminidase from *Arthrobacter protophormiae* grown in ovalbumin. *Appl Environ Microbiol.* 55:3107–3112.

Takegawa K, Yamabe K, Fujita K, Tabuchi M, Mita M, Izu H, Watanabe A, Asada Y, Sano M, Kondo A, et al. 1997. Cloning, sequencing, and expression of *Arthrobacter protophormiae* endo-β-N-acetylglucosaminidase in *Escherichia coli*. *Arch Biochem Biophys.* 338:22–28.

Tarentino AL, Maley F. 1974. Purification and properties of an endo-β-N-acetylglucosaminidase from *Streptomyces griseus*. *J Biol Chem.* 249:811–817.

Trimble RB, Tarentino AL. 1991. Identification of distinct endoglycosidase (endo) activities in *Flavobacterium meningosepticum*: Endo F1, endo F2, and endo F3. Endo F1 and endo H hydrolyze only high mannose and hybrid glycans. *J Biol Chem.* 266:1646–1651.

Tzelepis GD, Melin P, Jensen DF, Stenlid J, Karlsson M. 2012. Functional analysis of glycoside hydrolase family 18 and 20 genes in *Neurospora crassa*. *Fungal Genet Biol.* 49:717–730.

Umekawa M, Higashiyama T, Koga Y, Tanaka T, Noguchi M, Kobayashi A, Shoda S, Huang W, Wang LX, Ashida H, et al. 2010. Efficient transfer of sialo-oligosaccharide onto proteins by combined use of a glycosynthase-like mutant of *Mucor hiemalis* endoglycosidase and synthetic sialo-complex-type sugar oxazoline. *Biochim Biophys Acta.* 1800:1203–1209.

Umekawa M, Huang W, Li B, Fujita K, Ashida H, Wang LX, Yamamoto K. 2008. Mutants of Mucor hiemalis endo-beta-N-acetylglucosaminidase show enhanced transglycosylation and glycosynthase-like activities. *J Biol Chem.* 283:4469–4479.

Umekawa M, Li C, Higashiyama T, Huang W, Ashida H, Yamamoto K, Wang LX. 2010. Efficient glycosynthase mutant derived from *Mucor hiemalis* endo-β-N-acetylglucosaminidase capable of transferring oligosaccharide from both sugar oxazoline and natural N-glycan. *J Biol Chem.* 285:511–521.

Yin J, Li L, Shaw N, Li Y, Song JK, Zhang W, Xia C, Zhang R, Joachimiak A, Zhang HC, et al. 2009. Structural basis and catalytic mechanism for the dual functional endo-β-N-acetylglucosaminidase A. *PLoS One.* 4: e4658.

Yoshida Y, Tanaka K. 2010. Lectin-like ERAD players in ER and cytosol. *Biochim Biophys Acta.* 1800:172–180.

JSBA

Protease-Deficient *Saccharomyces cerevisiae* Strains for the Synthesis of Human-Compatible Glycoproteins

Kazuya Tomimoto,[1] Yasuko Fujita,[1] Tomoko Iwaki,[1] Yasunori Chiba,[2] Yoshifumi Jigami,[3] Ken-ichi Nakayama,[4] Yoshihiro Nakajima,[1] and Hiroko Abe[1,†]

[1]*Health Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 2217-14 Hayashi, Takamatsu, Kagawa 761-0395, Japan*
[2]*Bioproduction Research Institute, AIST, Central 6, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan*
[3]*Research Center for Medical Glycoscience, AIST, Central 2, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan*
[4]*Bioproduction Research Institute, AIST, 2-17-2-1 Tsukisamu-higashi, Toyohira-ku, Sapporo 062-8517, Japan*

Saccharomyces cerevisiae **strains engineered previously to produce proteins with mammalian high mannose structures showed severe growth defects and decreased protein productivity. In strain YAB101, derived from one of these strains by a mutagenesis technique based on the disparity theory of evolution, these undesirable phenotypes were alleviated. Here we describe further engineering of YAB101 with the aim of synthesizing heterologous glycoproteins with** $Man_5GlcNAc_2$, **an intermediate for the mammalian hybrid and complex type oligosaccharides. About 60% conversion of** $Man_8GlcNAc_2$ **to** $Man_5GlcNAc_2$ **was observed after integration of** *Aspergillus saitoi* **α-1,2-mannosidase fused to the transmembrane domain of** *S. cerevisiae* **Och1. To obtain a higher yield of the target protein, a protease-deficient version of this strain was generated by disruption of** *PEP4* **and** *PRB1*, **resulting in YAB101-4. Inactivation of these vacuolar proteases enhanced the secretion of human interferon-β by approximately 10-fold.**

Key words: glucan engineering; protease-deficient strain; mutagenesis; yeast

Mammalian cell culture has emerged as the most reliable technique for the production of therapeutic glycoproteins. It has, however, several drawbacks, including difficulty in genetic modification, long fermentation times, and vulnerability to viral contamination. The use of yeast to produce therapeutic glycoproteins has significant advantages, including the use of simple genetic modification and expression systems and a high rate of cell proliferation. In yeast, however, glycosylation of proteins differs from that in mammalian cells, which can trigger an immune response in humans. To overcome this problem, yeast cells have been engineered to produce heterologous glycoproteins with mammalian *N*-linked oligosaccharides.[1–4] Although *Pichia pastoris* strains have been developed that produce the mammalian complex type oligosaccharide,[5–7] we chose *Saccharomyces cerevisiae* for glycoengineering in

view of the wealth of information about genes, metabolism, and secretory mechanisms that is available for this organism. Moreover, the genetic tools and biochemical methods developed for *S. cerevisiae* allow for convenient analysis and elucidation of biochemical the features of the oligosaccharides and glycoproteins. Furthermore, *S. cerevisiae* is commonly used in a wide range of industrial processes. Hence we believe that glycoengineering in *S. cerevisiae* is important.

S. cerevisiae strains in which 3 genes (*och1Δ mnn1Δ mnn4Δ*) responsible for the biosynthesis of the outer mannan chain were disrupted[8,9] produced glycoproteins with a mammalian type high mannose structure instead of the mannan type of *S. cerevisiae*, $Man_8GlcNAc_2$ (also referred to as M8 in conjunction with a strain name). By further trimming of the mannose residues of $Man_8GlcNAc_2$ with α-1,2-mannosidase, $Man_5GlcNAc_2$ (abbreviated M5) as an intermediate for the mammalian hybrid and complex type oligosaccharide was synthesized. Since this trimming is the first step in the formation of hybrid and complex type oligosaccharides, increasing its efficiency is critical. Within this strain, however, the conversion of $Man_8GlcNAc_2$ was partial, so that overall 10% of the cell-wall mannoprotein contained $Man_5GlcNAc_2$.[8] In addition, the strain was temperature-sensitive and had a growth defect at 30 °C, which resulted in a low yield of heterologously produced protein. This decrease in protein productivity forms an obstacle to the engineering of an efficient glycoprotein production system using yeast. Therefore, the development of a *S. cerevisiae* strain that efficiently produces a protein with $Man_5GlcNAc_2$ oligosaccharide is highly desirable.

As reported previously, another *och1Δ mnn1Δ mnn4Δ* triple disruptant, strain TIY20 (M8), also had a growth defect, was temperature-sensitive, and gave a poor yield of heterologously produced protein.[10] To overcome these deleterious phenotypes, TIY20 (M8) was changed by a mutagenesis technique based on the disparity theory of evolution.[10] This technique employs an error-prone DNA polymerase δ that lacks the normal

---

proof-reading function and thus allows the introduction of numerous non-lethal mutations into the genome.[11,12] The resulting mutant strain, YAB101 (M8 Mut, "Mut" indicating use of a mutagenesis technique based on the disparity theory of evolution) did not have the growth defect or temperature-sensitive phenotype of the parental TIY20 strain (M8). The efficiency with which a heterologous protein was produced in YAB101 (M8 Mut) also improved in comparison to TIY20 (M8).[10] These properties make YAB101 (M8 Mut) a suitable starting strain for the synthesis of heterologous glycoproteins with mammalian N-linked oligosaccharides.

During heterologous protein production in *S. cerevisiae*, the target products are often degraded by proteases. Major intracellular proteolytic activities are attributed to vacuolar proteases, proteinases A and B, encoded by the *PEP4* and *PRB1* genes respectively.[13] Disruption of both genes creates a strain that can be considered protease deficient, providing a higher yield of heterologously produced proteins.[14,15]

In this study, we isolated a useful strain able efficiently to produce proteins glycosylated with Man5GlcNAc2 oligosaccharides by modification of YAB101 (M8 Mut). To trim Man8GlcNAc2 synthesized in YAB101 (M8 Mut) to Man5GlcNAc2, α-1,2-mannosidase encoded by the *msdS* gene of *Aspergillus saitoi* was fused to the transmembrane domain of a *S. cerevisiae* type-II membrane protein and expressed in YAB101 (M8 Mut). Further, to improve protein productivity, we isolated a protease-deficient version of this strain.

Materials and Methods

Yeast strains, media, and genetic methods. The strains used in this study are listed in Table 1. YPAD and SD media[16] were used to culture yeast cells and to select yeast transformants respectively. Transformation of *S. cerevisiae* was done by the lithium-acetate method.[17]

Integration of an msdS-expression cassette. The oligonucleotides used in this study are listed in Table 2. To construct strains TIY20-1 (M5) and YAB101-1 (M5 Mut) by integration of an *OCH1*TMD-*msdS* fusion gene, a DNA-fragment coding for aa 37 to 513 of *A. saitoi* *msdS*, lacking the signal peptide, was synthesized according to the codon usage of *S. cerevisiae* and fused to the transmembrane domain (TMD) of *S. cerevisiae* OCH1 (aa 1 to 60) (Fig. 1A). This *OCH1*TMD-*msdS* fusion gene was inserted into the *Eco*RI and *Sal*I sites of YEp352GAP-II, yielding plasmid YEp352GAP-II (*OCH1*TMD-*msdS*).[18] From this plasmid, the fusion gene, including

---

Table 1. Strains Used in This Study

Strain	Genotype	Source
W303-1B	*MATα leu2-3, 112 his3-11, 15 ade2-1 ura3-1 trp1-1 can1-100*	28
TIY20	W303-1B *mnn1Δ::hisG mnn4Δ::hisG och1Δ::hisG*	Provided by T. Ishii
YAB101	W303-1B *mnn1Δ::hisG mnn4Δ::hisG och1Δ::hisG other, unmapped mutations*	10
TIY20-1	W303-1B *mnn1Δ::hisG mnn4Δ::hisG och1Δ::hisG trp1Δ::msdS-TRP1*	This study
YAB101-1	W303-1B *mnn1Δ::hisG mnn4Δ::hisG och1Δ::hisG other, unmapped mutations trp1Δ::msdS-TRP1*	This study
YAB101-2	W303-1B *mnn1Δ::hisG mnn4Δ::hisG och1Δ::hisG other, unmapped mutations trp1Δ::msdS-TRP1 pep4Δ::hisG*	This study
YAB101-3	W303-1B *mnn1Δ::hisG mnn4Δ::hisG och1Δ::hisG other, unmapped mutations trp1Δ::msdS-TRP1 prb1Δ::hisG*	This study
YAB101-4	W303-1B *mnn1Δ::hisG mnn4Δ::hisG och1Δ::hisG other, unmapped mutations trp1Δ::msdS-TRP1 pep4Δ::hisG prb1Δ::hisG*	This study

---

Table 2. Primers Used in This Study

Primer	Sequence
GAPF-NotI	5'-CCCCCGCGGCCGCGGAACAACAAGAAGTTTAATGACGCGGAGGCC-3'
GAPR-KpnI	5'-GGGGGGGTACCGAATCGAAAATGTCATTAAAATAGTATATAAATTG-3'
TRP1-900F	5'-GAGCTCCTCGACAGCAGATCTGATGACTGG-3'
Man1-SalI	5'-GGGCCCGTCGACTTATGTACTACTCACCCGCACTGGATGTGCCTCGG-3'
Man1-F	5'-CGAAGAACCTCGCCG-3'
Man1-R	5'-GTCAAGTGTTGCGAGCTC-3'
PEP4-DF	5'-CAAAACTAACATGTTCAGCTTGAAAGCATCGACGGTATCGATAAGCTTG-3'
PEP4-DR	5'-GCCAAACCAACCGCATTGTTGCCCAAATCGCTCTAGAACTAGTGGATCC-3'
Man1-F	5'-CGAAGAACCTCGCCG-3'
Man1-R	5'-GTCAAGTGTTGCGAGCTC-3'
PEP4-ELF	5'-ATTTAATCCAAATAAAATTCAAACAAAAACCAAAACTAACATGTTCAGC-3'
PEP4-ELR	5'-AGTAAGAAAAGTTTAGCTCAAATTGCTTTGGCCAAACCAACCGCATTGT-3'
PEP4-F	5'-GAGAAGCCTACCACGTAAGGGAAGAATAAC-3'
PEP4-R	5'-CCCGCATATAATGACATTATGGGCAGCAGC-3'
PRB1-DF	5'-CTAATTCTAACAAGCAAAGATGAAGTTAGCGACGGTATCGATAAGCTTG-3'
PRB1-DR	5'-CTCTCACTTGATCAAAGATTAAATCGGTCGCTCTAGAACTAGTGGATCC-3'
PRB1-ELF	5'-CTTCATCGCCAATAAAAAAACAAACTAAACCTAATTCTAACAAGCAAAG-3'
PRB1-F	5'-GTTACGTCCCGTTATATTGGAGTTCTTCCC-3'
PRB1-R	5'-AGGGACTCCGACTTGTAACCTCGAGACGCC-3'
AFsig-NRI	5'-GTTTGAATTCATGAGATTTCCTTC-3'
AFsig-CSac	5'-GTTTGAGCTCTCTTTTATCTAGAG-3'

the GAPDH-promoter and terminator, was amplified with Expand High Fidelity Enzyme (Roche, Basel, Switzerland) and primers GAPF-NotI and GAPR-KpnI, which created *NotI* and *KpnI* restriction sites flanking the expression cassette. The PCR product was TA-cloned into pCR2.1TOPO (Life Technologies, Carlsbad, CA), and from the resulting subclone, the *OCH1*TMD-*msdS* expression cassette was cloned as a *NotI-KpnI* fragment into pRS304. Before yeast transformation, the resulting plasmid was linearized with *EcoRV* to facilitate integration into the *TRP1* locus of TIY20 (M8) or YAB101 (M8 Mut). Correct integration of the *OCH1*TMD-*msdS* DNA fragment was confirmed by PCR with primers TRP1-900F and Man1-SalI annealing about 900 bp upstream of the *TRP1* promoter and within the *msdS* gene respectively. Transcription of *OCH1*TMD-*msdS* gene was confirmed by RT-PCR by the SuperScript II One-Step RT-PCR System with Platinum Taq DNA Polymerase (Life Technologies) and primers Man1-F and Man1-R.

Gene disruption of PEP4 *and* PRB1. Strain YAB101-2 (M5 Mut *pep4Δ*) was obtained by displacing *PEP4* in YAB101-1 (M5 Mut) with a *PEP4*-HUH-cassette consisting of *Salmonella hisG* genes flanking both ends of the *URA3* gene (HUH cassette) and extended with regions homologous to the 5′ and 3′ UTRs of *PEP4* by PCR. For this, the HUH-cassette, cloned as an *EcoRI-PvuII* fragment from pSP73-HUH[19] between the *EcoRI* and *SmaI* sites of pBSIISK (+) (Agilent Technologies, Inc., Santa Clara, CA), was amplified by PCR in 2 steps, first with primers PEP4-DF and PEP4-DR and then, to extend the overlap with the *PEP4* UTRs, with primers PEP4-ELF and PEP4-ELR. The resulting *PEP4*-HUH PCR product was integrated into YAB101-1 (M5 Mut) by selection of transformants on SD-plates lacking uracil. Correct integration was confirmed by PCR with primers PEP4-F and PEP4-R. To pop out the *URA3* marker, 5FOA-selection was done as described by Takamatsu *et al.*[9] The resulting YAB101-2 (M5 Mut *pep4Δ*) strain was a *pep4* disruptant that had lost the *URA3* gene, as confirmed by PCR with primers PEP4-F and PEP4-R.

Strains YAB101-3 (M5 Mut *prb1Δ*) and YAB101-4 (M5 Mut *pep4Δ prb1Δ*) were constructed in the same manner by disruption of *PRB1* in YAB101-1 (M5 Mut) and YAB101-2 (M5 Mut *pep4Δ*) respectively, using PCR-amplified HUH-cassettes with overhangs homologous to the 5′ and 3′ UTRs of *PRB1*. The primers for primary PCR were PRB1-DF and PRB1-DR, and for the second PCR, PRB1-ELF and PRB1-ELR. Replacement of *PBR1* by the HUH-cassette was confirmed by PCR with primers PRB1-F and PRB1-R, and subsequent removal of the *URA3* marker was confirmed by PCR with primers PRB1-F and PBR1-R.

Construction of human interferon-β expression plasmids. To construct plasmid YEp352GAPII-MFhIFN-β, the prepro-sequence of the α-factor gene was amplified from YEp352GAPII-α-factor-MUC1 (kindly provided by Dr. Koh Amano of National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan) and cloned as an *EcoRI-SacI* fragment into YEp352GAP-II[18] to create pGAP2α. PCR was done with primers AFsig-NRI and AFsig-CSac. The gene for histidine-tagged human interferon-β (hIFN-β), containing the enterokinase cleavage site, optimized for *S. cerevisiae* codon usage, was synthesized by Hokkaido System Science (Sapporo, Japan) and inserted in-frame with the α-factor in pGAP2α to yield YEp352GA-PII-MFhIFN-β.

Growth analysis. To construct growth curves, YPAD medium was inoculated from fresh pre-cultures, grown overnight in YPAD at 30 °C, and incubated at 30 °C. To monitor cell growth, the optical density of a culture was determined spectrophotometrically at 600 nm (OD600) (at the indicated time-points in the case of the growth curves). For reliable measurement by OD, cultures were diluted to make the OD600 less than 1. At the indicated time points, OD600 measurements were made to monitor cell growth. For plate-assays, 10-fold serial dilutions of the preculture were spotted onto YPAD plates and then incubated for 2 d at the temperatures indicated.

Analysis of the N-*linked oligosaccharides of the mannoproteins.* All *S. cerevisiae* strains were grown in YPAD medium containing 300 mM KCl at 30 °C for 72 h, and 60% (w/v) glucose was added every 12 h to

maintain a final glucose concentration of 2% (w/v). Extraction of mannoproteins, digestion with glycopeptidase F (Takara Bio, Shiga, Japan), pyridylamination, and structural analysis of the oligosaccharides released were performed as described previously.[8] Purified pyridylaminated (PA) oligosaccharide was analyzed by HPLC (Prominence, Shimadzu, Kyoto, Japan) with a size-fractionation column (TSKgel Amide-80, 3 μm column, 4.6 × 150 mm, Tosoh, Tokyo) and 2 eluents: 70% acetonitrile and 30% 200 mM triethylamine acetate (eluent A); and 30% acetonitrile and 70% 200 mM triethylamine acetate (eluent B). The column was equilibrated with eluent A at a flow rate of 1 mL/min. After injection of the sample, the concentration of eluent B was increased from 0% to 50% over 50 min. PA-oligosaccharides were detected by measuring the fluorescence (excitation wavelength, 310 nm; emission wavelength 380 nm).

Cytokine ELISA. Interferon-β levels in the culture supernatant were measured with a Human IFN-β ELISA kit (PBL Interferon Source, Piscataway, NJ). Yeast cells carrying YEp352GAPII-MFhIFN-β were grown in 3 mL of YPAD containing 0.3 M sorbitol for 2–3 d at 30 °C with aeration. The starting OD600 was 0.1, and the cells grew to OD600 6 to 7 in 2 d. They were harvested by centrifugation of the cultures at 3,000 rpm for 5 min, and then the supernatants were analyzed by ELISA. At the same time, the OD600 of a suitably-diluted culture was measured. Routinely, for each strain for which interferon-β secretion was assessed, 5 transformants were analyzed in parallel.

Results and Discussion

Isolation of S. cerevisiae *TIY20-1 (M5) and YAB101-1 (M5 Mut) producing* Man₅GlcNAc₂ *oligosaccharide*

Man₈GlcNAc₂ *N*-linked oligosaccharides must be converted by α-1,2-mannosidase to obtain glycoproteins with a Man₅GlcNAc₂ structure, an intermediate of human hybrid and complex type oligosaccharides. A previous our attempt to introduce α-1,2-mannosidase activity into yeast demonstrated its feasibility, but the engineered strain had growth defects that resulted in low yields of heterologously produced protein. In this strain, α-1,2-mannosidase was expressed from a plasmid and tagged with the His-Asp-Glu-Lue (HDEL) peptide,[8] which acts as a retention/retrieval signal for the endoplasmic reticulum (ER) in yeast cells,[20] but of the Man₈GlcNAc₂ *N*-linked oligosaccharides, not more than 10% was converted to Man₅GlcNAc₂.[8] To enhance the formation of Man₅GlcNAc₂, α-1,2-mannosidase (*msdS*) was fused to the transmembrane domain (TMD) of *S. cerevisiae* Och1, an integral type-II membrane protein that is part of the Golgi apparatus.[21,22] The expression cassette, containing *OCH1*-TMD-*msdS* under the control of the GAPDH promoter (Fig. 1A), was integrated at the *TRP1* locus into 2 strains with a W303 background, YAB101 (M8 Mut) and TIY20 (M8), that form Man₈GlcNAc₂ glycoproteins after disruption of yeast-specific glycosyl transferase genes *OCH1*, *MNN1*, and *MNN4*. YAB101 (M8 Mut) has been derived from TIY20 (M8) to overcome the growth defect and the temperature-sensitive phenotype of the TIY20 strain (M8).[10] The resulting strains, in which the expression of the *OCH1*TMD-*msdS* fusion was confirmed by RT-PCR (data not shown), were designated YAB101-1 (M5 Mut) and TIY20-1 (M5) respectively (Table 1).

Growth analysis of these strains at 30 °C in liquid medium (Fig. 1B) and by serial dilution on solid media (Fig. 1C) showed a higher growth rate for YAB101-1 (M5 Mut) than for TIY20-1 (M5), while the growth of YAB101-1 (M5 Mut) did not differ from that of parental
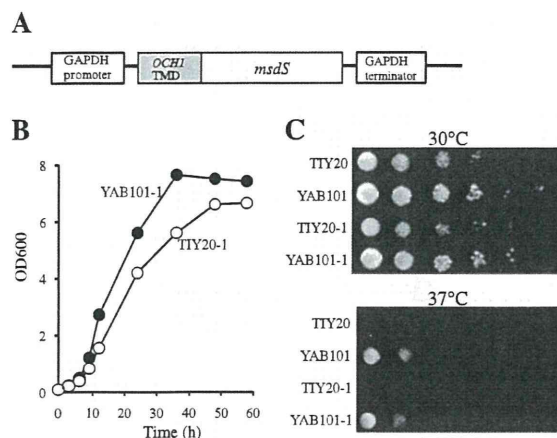
A

B   **C**

Fig. 1. Integration of the OCH1TMD-msdS Fusion Gene into Chromosomes of YAB101 (M8 Mut) and TIY20 (M8), and Phenotypes of YAB101-1 (M5 Mut) and TIY20-1 (M5).

A, Schematic overview of the OCH1TMD-msdS expression cassette with the GAPDH-promoter and terminator. B, Growth analysis of YAB101-1 (solid circles) and TIY20-1 (hollow circles) in liquid culture. At the indicated time points, OD600 measurements were made to monitor cell growth. C, Growth analysis by serial dilution on plates.
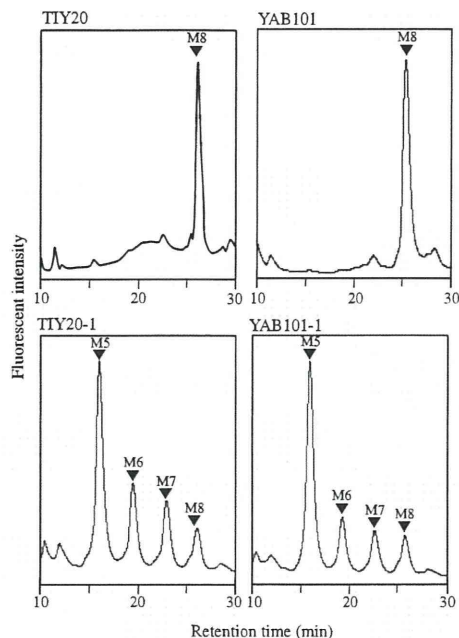

Fig. 2. Chromatograms of N-Linked Oligosaccharides Released from the Cell-Wall Proteins of TIY20 (M8), YAB101 (M8 Mut), TIY20-1 (M5), and YAB101-1 (M5 Mut).

N-Linked oligosaccharides were separated by HPLC with a TSKgel Amide-80 column (see "Materials and Methods"). Elution times of authentic pyridylaminated (PA) oligosaccharides $Man_5GlcNAc_2$-PA (M5), $Man_6GlcNAc_2$-PA (M6), $Man_7GlcNAc_2$-PA (M7), and $Man_8GlcNAc_2$-PA (M8) are indicated by arrowheads.

YAB101 (M8 Mut), indicating that the expression of OCH1TMD-msdS did not have an impact on the growth of this strain. That the presence of OCH1TMD-msdS did not affect growth was underlined by the finding that temperature sensitivity of TIY20-like cells (M8) remained unchanged, whereas the YAB101-1 cells (M5 Mut) grew as well as their parental strain at 37 °C (Fig. 1C).

Analysis of the structures of the N-linked oligosaccharides

In order to determine the effect of expression of the α-1,2-mannosidase fused to the transmembrane domain of Och1, we compared the structure of N-linked oligosaccharides in the cell-wall mannoprotein from strains YAB101 (M8 Mut), TIY20 (M8), YAB101-1 (M5 Mut), and TIY20-1 (M5) by HPLC after digestion with glycopeptidase F and pyridylamination (see "Materials and Methods" above) (Fig. 2). Confirming previous analysis,[10] the oligosaccharide structure from TIY20 (M8) and YAB101 (M8 Mut) was $Man_8GlcNAc_2$, while the expression of α-1,2-mannosidase in either TIY20-1 (M5) or YAB101-1 (M5 Mut) resulted in a mixture of $Man_{5-8}GlcNAc_2$, and more than 60% of the oligosaccharides of the cell-wall mannoprotein were converted to $Man_5GlcNAc_2$. This ratio is much higher than the 10% that has been reported previously for a strain in which α-1,2-mannosidase was tagged with the HDEL peptide and expressed from a plasmid.[8] Proteins with the HDEL peptide enter a retrograde transport pathway for return to the ER from the Golgi apparatus.[23,24] Choi et al. achieved high efficiency in the conversion to $Man_5GlcNAc_2$ by using a transmembrane domain of the type-II membrane protein of yeast to localize α-1,2-mannosidase to the secretory pathway of P. pastoris.[5] Thus, the cellular localization of α-1,2-mannosidase using a type-II transmembrane domain appears to contribute to the efficiency with which $Man_8GlcNAc_2$ is converted to $Man_5GlcNAc_2$. This conversion might even be enhanced by increasing the expression level of α-1,2-mannosidase. Thus the more efficient conversion to $Man_5GlcNAc_2$ in our system have been the result of anchoring α-1,2-mannosidase with a type-II transmembrane domain, while genomic integration of the OCHTMD-msdS cassette might lead to more efficient expression of the gene.

Production of human interferon-β

To determine whether the engineered strains are suitable for producing heterologous glycoproteins, we analyzed the efficiency of secretion of human interferon-β (hIFN-β) from TIY20-1 (M5) and YAB101-1 (M5 Mut). hIFN-β is a glycoprotein secreted by fibroblasts in response to viral infection or exposure to double-stranded RNA.[25] It functions as an anti-viral and anti-tumor agent. The gene for hIFN-β was fused to the prepro-sequence of the α-factor gene to facilitate secretion, and was expressed by the GAPDH promoter on plasmid YEp352GAPII-MFhIFN-β (Fig. 3A), which was transformed into TIY20-1 (M5) and YAB101-1 (M5 Mut). Five transformants from each strain were cultured in YPAD for 3 d, and then the levels of secreted hIFN-β were measured by ELISA. As shown in Fig. 3B, compared to the TIY20-1 strain (M5), the YAB101-1 strain (M5 Mut) secreted about 14-fold more hIFN-β into the growth media. This result suggests that the increased fitness obtained by the mutagenesis procedure based on the disparity theory of evolution also had a positive effect on hIFN-β productivity.
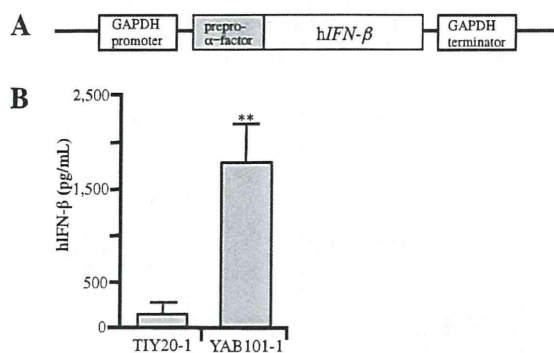
A

B

Fig. 3. Amounts of hIFN-β Secreted by YAB101-1 (M5 Mut) and by TIY20-1 (M5).

A, Schematic overview of α-factor-hIFN-β fusion expressed by YEp352GAPII-MFhIFN-β. B, The level of hIFN-β in each culture was determined by ELISA. Results shown are mean ± standard error for 5 transformants of each strain. Differences between groups were evaluated by t test. **$p < 0.01$.

A vacuolar protease-deficient version of YAB101-1 (M5 Mut)

A factor that limits the yield of heterologously produced proteins is degradation by host-specific proteases released into the medium due to fermentation stress or cell-lysis.[14,26] In order to determine whether the levels of secreted hIFN-β can be enhanced by disruption of vacuolar proteases, the genes for *PEP4* and *PRB1* were disrupted in YAB101-1 (M5 Mut). *PEP4* and *PRB1* are essential for the maturation and activation of other vacuolar proteases in that their disruption effectively creates a protease-deficient strain.[14] The *PEP4* gene was disrupted in YAB101-1 (M5 Mut), yielding YAB101-2 (M5 Mut *pep4Δ*) (Table 1), while disruption of *PRB1* in YAB101-1 (M5 Mut) and YAB101-2 (M5 Mut *pep4Δ*) resulted in YAB101-3 (M5 Mut *prb1Δ*) and YAB101-4 (M5 Mut *pep4Δ prb1Δ*) respectively. Analyzed at 30 °C, the growth rates of these strains showed no differences, indicating that the deletion of vacuolar proteases in our glyco-engineered strains did not affect cell growth (Fig. 4A).

The secretion of hIFN-β by YAB101-1 (M5 Mut), YAB101-2 (M5 Mut *pep4Δ*), YAB101-3 (M5 Mut *prb1Δ*), and YAB101-4 (M5 Mut *pep4Δ prb1Δ*) was assessed after transformation with hIFN-β expression plasmid YEp352GAPII-MFhIFN-β (see "Material and Methods" above). Cells were grown for 2 d and then the amount of hIFN-β in the culture medium was measured by ELISA (Fig. 4B). Secretion of hIFN-β by YAB101-4 (M5 Mut *pep4Δ prb1Δ*) increased approximately 10-fold as compared to YAB101-1 (M5 Mut), suggesting that decreased protease activity in the double deletion mutant provides the optimal genetic background to produce secreted hIFN-β.

Still, the secretion levels were too poor to obtain sufficient amounts of hIFN-β to analyze its oligosaccharide structures. Demolder *et al.* have reported that induction of hIFN-β in their strain, which was also deficient in vacuolar proteinases, was toxic and inhibited cell proliferation.[27] This phenomenon, however, was not observed in our strains (data not shown), perhaps due to the inclusion of sorbitol in the growth medium (see "Materials and Methods" above). Demolder *et al.*
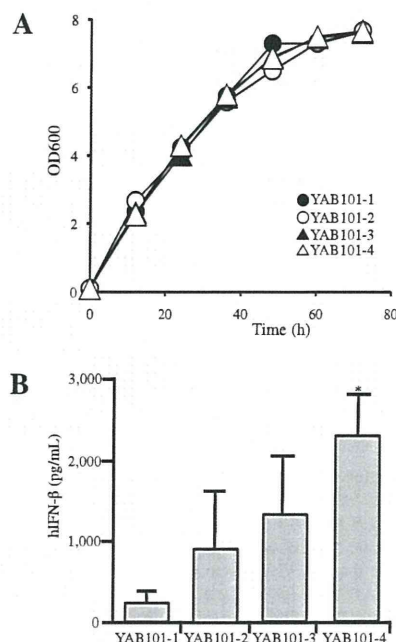
A

B

Fig. 4. Growth Analysis of YAB Strains and Secretion of hIFN-β Produced in YAB Strains Carrying YEp352GAPII-MFhIFN-β.

A, Growth analysis of YAB101-1 (M5 Mut) (solid circles), YAB101-2 (M5 Mut *pep4Δ*) (hollow circles), YAB101-3 (M5 Mut *prb1Δ*) (solid triangles), and YAB101-4 (M5 Mut *pep4Δ prb1Δ*) (hollow triangles) in liquid YPAD medium at 30 °C. B, hIFN-β levels were determined by ELISA after 2 d of growth. Results for 5 transformants of each strain are shown as mean ± standard error. *$p < 0.05$ as compared to YAB101-1 (M5 Mut) (one-way ANOVA followed by Dunnett's test).

also found that they could increase the secretion level of hIFN-β by overexpression of *HSP70*.[27] Thus, it is possible that overexpression of this chaperone in our strains also helps to raise the secretion of hIFN-β. Despite the big difference in strain background that Demolder et al. used an integrated hIFN-β expression-cassette under the control of the inducible *GAL1* promoter,[27] our set-up with an episomal plasmid and the constitutive *GAPDH*-promoter in a glycoengineered strain gave similarly poor secretion of hIFN-β suggesting that other, less well-known, variables affect the heterologous production of this glycoprotein. Still, the present study indicates that disruption of *PEP4* and *PRB1* improved hIFN-β secretion efficiency in a yeast strain that synthesized glycoprotein with $Man_5GlcNAc_2$ oligosaccharide. Hence we believe that YAB101-4 (M5 Mut *pep4Δ prb1Δ*) is favorable to use as a starting strain for more complex glycoengineering to make possible the production of proteins with mammalian N-linked oligosaccharides. To our knowledge, YAB101-4 (M5 Mut *pep4Δ prb1Δ*), which was isolated in this study, is the most useful strain for the synthesis of glycoproteins with $Man_5GlcNAc_2$ oligosaccharide in *S. cerevisiae*.

Acknowledgments

References

1) Chiba Y, Sakuraba H, Kotani M, Kase R, Kobayashi K *et al.*, *Glycobiology*, **12**, 821–828 (2002).

2) Hamilton SR, Davidson RC, Sethuraman N, Nett JH, Jiang Y *et al.*, *Science*, **313**, 1441–1443 (2006).

3) Hamilton SR and Gerngross TU, *Curr. Opin. Biotechnol.*, **18**, 387–392 (2007).

4) Vervecken W, Kaigorodov V, Callewaert N, Geysens S, De Vusser K, and Contreras R, *Appl. Environ. Microbiol.*, **70**, 2639–2646 (2004).

5) Choi BK, Bobrowicz P, Davidson RC, Hamilton SR, Kung DH *et al.*, *Proc. Natl. Acad. Sci. USA*, **100**, 5022–5027 (2003).

6) Bobrowicz P, Davidson RC, Li H, Potgieter TI, Nett JH *et al.*, *Glycobiology*, **14**, 757–766 (2004).

7) Wildt S and Gerngross TU, *Nat. Rev. Microbiol.*, **3**, 119–128 (2005).

8) Chiba Y, Suzuki M, Yoshida S, Yoshida A, Ikenaga H, Takeuchi M, Jigami Y, and Ichishima E, *J. Biol. Chem.*, **273**, 26298–26304 (1998).

9) Takamatsu S, Chiba Y, Ishii T, Nakayama K, Yokomatsu-Kubota T, Makino T, Fujibayashi Y, and Jigami Y, *Glycoconj. J.*, **20**, 385–397 (2004).

10) Abe H, Takaoka Y, Chiba Y, Sato N, Ohgiya S *et al.*, *Glycobiology*, **19**, 428–436 (2009).

11) Wagner J, Gruz P, Kim SR, Yamada M, Matsui K, Fuchs RP, and Nohmi T, *Mol. Cell*, **4**, 281–286 (1999).

12) Goodman MF, *Annu. Rev. Biochem.*, **71**, 17–50 (2002).

13) Jones EW, *Methods Enzymol.*, **194**, 428–453 (1991).

14) Gleeson MA, White CE, Meininger DP, and Komives EA, *Methods Mol. Biol.*, **103**, 81–94 (1998).

15) Komeda T, Sakai Y, Kato N, and Kondo K, *Biosci. Biotechnol. Biochem.*, **66**, 628–631 (2002).

16) Sherman F, *Methods Enzymol.*, **194**, 3–21 (1991).

17) Klebe RJ, Harriss JV, Sharp ZD, and Douglas MG, *Gene*, **25**, 333–341 (1983).

18) Nakayama K, Maeda Y, and Jigami Y, *Glycobiology*, **13**, 673–680 (2003).

19) Alani E, Cao L, and Kleckner N, *Genetics*, **116**, 541–545 (1987).

20) Pelham HR, Hardwick KG, and Lewis MJ, *EMBO J.*, **7**, 1757–1762 (1988).

21) Nakayama K, Nagasu T, Shimma Y, Kuromitsu J, and Jigami Y, *EMBO J.*, **11**, 2511–2519 (1992).

22) Gaynor EC, te Heesen S, Graham TR, Aebi M, and Emr SD, *J. Cell Biol.*, **127**, 653–665 (1994).

23) Lewis MJ and Pelham HR, *Nature*, **348**, 162–163 (1990).

24) Semenza JC, Hardwick KG, Dean N, and Pelham HR, *Cell*, **61**, 1349–1357 (1990).

25) Pestka S, Langer JA, Zoon KC, and Samuel CE, *Annu. Rev. Biochem.*, **56**, 727–777 (1987).

26) Idiris A, Tohda H, Kumagai H, and Takegawa K, *Appl. Microbiol. Biotechnol.*, **86**, 403–417 (2010).

27) Demolder J, Fiers W, and Contreras R, *J. Biotechnol.*, **32**, 179–189 (1994).

28) Thomas BJ and Rothstein R, *Cell*, **56**, 619–630 (1989).

Molecular Epidemiology and Genetic History of Hepatitis C Virus Subtype 3a Infection in Thailand

Srunthron Akkarathamrongsin[a, b] Pitchaya Hacharoen[b] Pisit Tangkijvanich[c]
Apiradee Theamboonlers[b] Yasuhito Tanaka[d] Masashi Mizokami[e]
Yong Poovorawan[b]

[a]Inter-Department of Biomedical Sciences, Faculty of Graduate School, [b]Center of Excellence in Clinical Virology, and [c]Department of Biochemistry, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand; [d]Department of Virology and Liver Unit, Nagoya City University Graduate School of Medical Science, Nagoya, and [e]Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, Ichikawa, Chiba, Japan

Abstract

Objective: Among all hepatitis C virus (HCV) infections, subtype 3a is the most common genotype in Thailand. This study investigates the molecular epidemiology and epidemic history of HCV subtype 3a in Thailand. *Methods:* Three hundred and fifty-six serum samples were collected from HCV-infected Thai patients. The virus was isolated, after which the core and NS5B regions were sequenced. Subsequently, the HCV genotype was classified by phylogenetic analysis based on the core and NS5B regions. Molecular evolution analysis of HCV subtype 3a was estimated using BEAST (Bayesian Evolutionary Analysis by Sampling Trees) v.1.5.4. *Results:* Based on our phylogenetic analyses, subtype 3a (38.5%) was the most prevalent, followed by 1a (21%), 1b (13.8%), genotype 6 (19.9%) [comprised of subtypes 6e (0.3%), 6f (11%), 6i (1.9%), 6j (1.9%) and 6n (4.8%)] and 3b (5.6%). Our phylogenetic tree indicates the existence of a specific group of HCV subtype 3a strains in the Thai popula-

tion. Molecular evolutionary analysis dated the most recent common ancestor of the Thai HCV subtype 3a strains as existing approximately 200 ago, and a Bayesian skyline plot showed that this particular strain spread to Thailand during the mid-1970s and early 1980s. This period overlaps with the Vietnam War (1955–1975) and the widespread use of injection stimulants introduced by the US Army during this time. *Conclusion:* The estimated history of HCV subtype 3a infection in Thailand may help to predict the future burden of HCV-related diseases and facilitate better public health control and surveillance.

© 2013 S. Karger AG, Basel

Introduction

Hepatitis C virus (HCV) is a major public health problem affecting approximately 170 million people worldwide [1], and persistent HCV infection may eventually lead to cirrhosis and culminate in hepatocellular carcinoma (HCC). In addition, serious HCV complications often occur without clinical symptoms until after two or three decades of viral incubation [2, 3]. More than 20

Prof. Yong Poovorawan, MD
Center of Excellence in Clinical Virology
Department of Pediatrics, Faculty of Medicine
Chulalongkorn University, Bangkok 10330 (Thailand)
E-Mail Yong.P@chula.ac.th

years of epidemiological surveillance have proven that HCV-related HCC is increasing in developed countries [4]. Although in Europe and the United States an average of 3.7% of cirrhosis cases per year contribute to HCC incidence, the growing incidence in Japan is expected to reach a plateau in the near future [5, 6].

HCV is a member of Flaviviridae which can be classified into six major genotypes and many subtypes [7, 8]. Genotype distribution differs according to geographic regions, mode of transmission and treatment responses [2, 9, 10]. Genotypes 1, 2 and 3 are distributed worldwide and contribute to the majority of HCV infection. Subtypes 1a and 1b are most common in Europe, the USA and Japan, and subtypes 2a and 2b are predominant in Japan, North America and Europe, while subtype 3a is most prevalent in the Indian subcontinent and Thailand [11–13]. On the other hand, the less common genotypes 4, 5 and 6 are restricted to certain geographic areas. Genotype 4 is prevalent in the Middle East, genotype 5 in South Africa and genotype 6 in South China and Southeast Asia [13–16].

The mode of HCV transmission is associated with genotype distribution; for example, evolutionary analysis has suggested that HCV subtype 1a may be spread via blood transfusion and unsafe medical practices [17]. In addition, subtypes 1a and 3a are prevalent in young individuals and particularly in injecting drug users (IDUs) [18] in the USA, UK, Uzbekistan, Russia and Thailand [10, 19–21]. Past and present knowledge of HCV prevalence is important in order to be able to predict the future burden of HCV-related liver diseases. However, due to its asymptomatic nature and long-term complications, it is difficult to estimate the past prevalence of HCV infection from medical records. Thus, other methods are required. Coalescence theory is based on the hypothesis that all individuals originated from one common ancestor, and that populations which have evolved independently tend to genetically diverge over the time. Applying the coalescence theory and a molecular clock to phylogenetic reconstruction of present day HCV sequences allows HCV population dynamics to be estimated prior to the virus's discovery and investigation [9, 12], and can be used to estimate the modification of viral population size and viral epidemic behavior [22, 23]. A fine example of this method's efficacy comes from the elucidation that the high prevalence of HCV subtype 4a in Egypt was due to unsterile needle injection during mass treatment campaigns against schistosomiasis from the 1930s to the 1950s [9].

The distribution of HCV genotypes in Thailand has been investigated in this study and previously, and it has been found that subtype 3a is the dominant strain (39–51%) in all regions and populations. In addition, subtype 3a has a very high prevalence in IDUs (75–83%) and our phylogenetic analysis shows a specific clustering of Thai HCV, which indicates a specific transmission route. Other subtypes are also present in Thailand and the Thai population such as the 1a, 1b, 2a and 6 variants, but at a lower prevalence in comparison to subtype 3a [12, 24]. In Myanmar, Laos and Cambodia, Thailand's neighbors, HCV genotype 6 variants are predominant rather than 3a [14, 25, 26]. Thus, due to its high prevalence in the Thai population and geographic specificity, this study focused on the analysis of subtype 3a from Thai isolates along with reference strains from other countries.

Population-based surveys have shown that the HCV seroprevalence is 2.2% in Thailand [12]; however, HCC incidence caused by HCV infection is low in this country. Previous studies have demonstrated that the initial spread time of HCV infection is associated with the prevailing socioeconomic conditions, reflecting the present incidence as well as predicting the future development of HCC [11, 27–30]. Because HCC is associated directly with the incubation period of HCV infection and HCV subtype 3a is the predominant genotype [11, 30], the initial spread of HCV subtype 3a among the Thai population may occur later than in other countries.

Due to the epidemic history of HCV transmission, which reflects the present viral population and disease burden, the objective of the study is to estimate the time of HCV subtype 3a introduction and spread in the Thai population. The suspected mode of transmission history of HCV infection has also been described. This estimation may be useful for predicting the incidence of HCV-related diseases in the country and also used as data for HCV prevention and public health control.

Materials and Methods

Sample and Clinical Data Collection
The study protocol was approved by the Ethics Committee of the Faculty of Medicine, Chulalongkorn University before using the specimens. Serum samples were collected from HCV-infected patients at Chulalongkorn Hospital (Bangkok), the hospitals in Phetchabun Province (Phetchabun General Hospital, Lomsak General Hospital and Lomkao Crown Prince Hospital), as well as the National Blood Center, Thailand. Patients who were tested for anti-HCV and had positive results were sent to the Center of Excellence in Clinical Virology in order to investigate the viral RNA and genotype for treatment. Only samples with HCV RNA positive by RT-PCR were recruited for this study. Coded anonymous samples were collected from 2003 to 2009. All specimens