of hyperglycemia during decompensation.[30] Decreased npRQ was correlated with glucose intolerance and fibrosis stage, suggesting that the clinicopathogenesis of NAFLD is closely associated with glucose intolerance, and that early intervention for glucose intolerance is important in clinical practice.

To examine the utility of npRQ as a marker of disease progression in NAFLD, we compared AUROC for npRQ with those for various other parameters and scoring systems in three patterns to discriminate NASH from NAFLD (stage 0 vs stages 1–3), significant fibrosis (stages 0–1 vs stages 2–3) and advanced fibrosis (stages 0–2 vs stage 3). As the decrease in npRQ became significant at stage 2 (Fig. 2a), the AUROC for npRQ for differentiation of stages 2–3 from stages 0–1 was superior to other parameters. Therefore, our results indicate that decreased npRQ can be used to detect NASH, including relatively early stage NASH in many NAFLD patients. In addition to npRQ, AUROC for HOMA-IR, AUC glucose, fasting insulin and NAFIC score were each approximately 0.850 and also showed differences for each of the three differentiation patterns. Therefore, these parameters also have the ability to detect NASH from the early stages to the development of severe fibrosis. NAFIC score, the scoring system for fibrosis proposed by Sumida *et al.*,[14] comprises three measurements (serum ferritin, insulin and type IV collagen 7S) and is easy to calculate. A validation study by the Japan Study Group of NAFLD (JSG-NAFLD) reported that NAFIC score was superior to other several previously established scoring systems in detecting NASH with fibrosis among Japanese NAFLD patients, and also for predicting severe fibrosis.[14] In the present study, AUROC for NAFIC score for differentiation of stage 3 from stages 0–2 was 0.9200 and was the highest among the various parameters, supporting the conclusions of the JSG-NAFLD report.[14] NAFLD fibrosis score,[15] which consists of six variables (age, BMI, hyperglycemia, platelet count, albumin and AST/ALT ratio) has been reported to reliably predict advanced fibrosis. In a meta-analysis by Angulo *et al.*,[15] NAFLD fibrosis score had an AUROC of 0.85 for predicting advanced fibrosis (stages 3–4). The fact that subjects with stage 4 were excluded from the present study but were included as "advanced stage" in previously reported studies,[14,15,34] may also contribute to the lower value of AUROC in the present study. In addition, it is uncertain whether a scoring system established using data from Caucasian populations is applicable to Asian patients, because Asian patients tend to develop NASH and other metabolic complications at a lower BMI than Caucasians.[3] The FIB-4 index was developed as a scoring system for estimation of liver fibrosis in subjects with HIV and hepatitis C virus co-infection.[16] It relies on patient age, AST, ALT and platelet count. The advantage of FIB-4 index is that it is easy to calculate and does not require the use of BMI. In a validation study of JSG-NAFLD,[34] FIB-4 index was superior to other fibrosis scoring systems in Japanese NAFLD patients for excluding advanced fibrosis. In our study, the AUROC for FIB-4 index for discrimination of stage 3 from stages 0–2 was not satisfied (0.7143) and was inferior to NAFIC score. As described above, the lower AUROC for FIB-4 index than in previous reports[34] may be due to the fact that stage 4 patients were excluded from the present study, but were included in previous reports. As the number of subjects in the present study was small, we cannot conclude which parameters and scoring system is best for the estimation of severity of NAFLD, nor we can definitively state that npRQ is the best method for differentiation of NASH from NAFLD. However, Table 4 shows that npRQ, HOMA-IR, AUC glucose, fasting glucose, insulin and NAFIC scores all could provide useful information for the detection of NASH, including patients in the early fibrosis stage, whereas NAFLD fibrosis score and type IV collagen were useful for identification of advanced fibrosis. It is important to select these parameters and scoring systems according to the purpose: to differentiate early-stage NASH from NAFLD, or to detect advanced fibrosis.

Although we believed that npRQ is useful for the estimation of disease severity in NAFLD patients, unfortunately, npRQ measurement is not always possible in clinical practice. This is because indirect calorimetry for measurement of npRQ was primarily used in inpatient and research settings[7] and not all clinicians are familiar with the equipment used. In addition, patients must be tested in the early morning after fasting, and 24-h urine specimens must be collected for calculation of npRQ. Okumura *et al.*[35] investigated the use of serum biochemistry to predict npRQ in patients with viral cirrhosis. That report concluded that serum NEFA can be used to predict npRQ. However, in the present study, npRQ was not correlated with NEFA. This may be because the subjects in the present study were NAFLD patients without cirrhosis, and not viral cirrhosis patients with decreased glycogen storage.[35] In NAFLD patients with glucose intolerance, HOMA-IR may be useful for prediction of npRQ without calorimetry, because these two parameters were negatively correlated.

In conclusion, npRQ is useful for the estimation of disease severity in NAFLD patients with glucose intolerance. It enables the detection of NASH with relatively

early-stage fibrosis among NAFLD patients. As this measurement can provide useful information without the burden of blood collection, it should be included in clinical practice in addition to anthropometry and blood analysis. When an NAFLD patient exhibits low npRQ, the patient should undergo further examination such as a liver biopsy to allow for early intervention.

REFERENCES

1 Okanoue T, Umemura A, Yasui K, Itoh Y. Nonalcoholic fatty liver disease and nonalcoholic steatohepatitis in Japan. *J Gastroenterol Hepatol* 2011; **26** (Suppl 1): 153–62.
2 Hashimoto E, Tokushige K. Prevalence, gender, ethnic variations, and prognosis of NASH. *J Gastroenterol* 2011; **46**: 63–6.
3 Sumida Y, Eguchi Y, Ono M. Current status and agenda in the diagnosis of nonalcoholic steatohepatitis in Japan. *World J Hepatol* 2010; **2**: 374–83.
4 Chalasani N, Younossi Z, Lavine JE *et al.* The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. *Hepatology* 2012; **55**: 2005–23.
5 Ratziu V, Bellentani S, Cortez-Pinto H, Day C, Marchesini G. A position statement on NAFLD/NASH based on EASL 2009 special conference. *J Hepatol* 2010; **53**: 372–84.
6 Chitturi S, Farrell GC, Hashimoto E, Saibara T, Lau GK, Sollano JD. Guidlines for the assessment and management of non-alcoholic fatty liver disease in the Asia-Pasific region: executive summary. *J Gastroenterol Hepatol* 2007; **22**: 775–77.
7 Haugen HA, Chan LN, Li F. Indirect calorimetry: a practical guide for clinicians. *Nutr Clin Pract* 2007; **22**: 377–88.
8 Tajika M, Kato M, Mohri H *et al.* Prognostic value of energy metabolism in patients with vial liver cirrhosis. *Nutrition* 2002; **18**: 229–34.
9 Angulo P, Keach JC, Batts KP, Lindor KD. Independent predictors of liver fibrosis in patients with nonalcoholic steatohepatitis. *Hepatology* 1999; **30**: 1356–62.
10 Miyaaki H, Ichikawa T, Nakao K *et al.* Clinicopathological study of nonalcoholic fatty liver disease in Japan: the risk factors for fibrosis. *Liver Int* 2008; **28**: 519–24.
11 Kimura Y, Hyogo H, Ishitobi T, Nabeshima Y, Arihiro K, Chayama K. Postprandial insulin secretion pattern is associated with histological severity in non-alcoholic fatty liver disease patients without prior known diabetes mellitus. *J Gastroenterol Hepatol* 2011; **26**: 517–22.
12 Manchanayake J, Chitturi S, Nolan C, Farell GC. Postparndial hyperinsulinemia is universal in non-diabetic patients with nonalcoholic fatty liver disease. *J Gastroenterol Hepatol* 2011; **26**: 510–6.
13 Eguchi Y, Mizuta T, Sumida Y *et al.* The pathological role of visceral fat accumulation in steatosis, inflammation, and progression of nonalcoholic fatty liver disease. *J Gastroenterol* 2011; **46**: 70–8.
14 Sumida Y, Yoneda M, Hyogo H *et al.* A simple clinical scoring system using ferritin, fasting insulin, and type IV collagen 7S for predicting steatohepatitis in nonalcoholic fatty liver disease. *J Gastroenterol* 2011; **46**: 257–68.
15 Angulo P, Hui JM, Marchesini G *et al.* The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology* 2007; **45**: 846–54.
16 Sterling RK, Lissen E, Clumeck N *et al.* Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 2006; **43**: 1317–25.
17 Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus. Provisional report of a WHO Consultation. *Diabet Med* 1998; **15**: 539–53.
18 Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentration in man. *Diabetologia* 1985; **28**: 412–9.
19 Korenaga K, Korenaga M, Uchida K, Yamasaki T, Sakaida I. Effects of a late evening snack combined with alpha-glucosidase inhibitor on liver cirrhosis. *Hepatol Res* 2008; **38**: 1087–97.
20 Harris JA, Benedict FG. A biometric study of basal metabolism in man. *Carnegie Inst* 1919; **279** (3): 48–9.
21 Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Non-alcoholic steatohepatitis: a proposal for grading and staging the histological lesions. *Am J Gastroenterol* 1999; **94**: 2467–74.
22 Kleiner DE, Brunt EM, Van Natta M *et al.* Design and validation of a histological scoring system for nonalchoholic fatty liver disease. *Hepatology* 2005; **41**: 1313–21.
23 Matsuda M, DeFronzo RA. Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. *Diabetes Care* 1999; **22**: 1462–70.
24 Shiga T, Moriyoshi Y, Nagahara H, Shiratori K. Nonalcoholic fatty liver is a risk factor for postprandial hyperglycemia, but not for impaired fasting glucose. *J Gastroenterol* 2009; **44**: 757–64.
25 Haukeland JW, Konopski Z, Linnestad P *et al.* Abnormal glucose tolerance is a predictor of steatohepatitis and fibrosis in patients with non-alcoholic fatty liver disease. *Scand J Gastroenterol* 2005; **40**: 1469–77.
26 Wong VW, Hui AY, Tsang SW *et al.* Prevalence of undiagnosed diabetes and postchallenge hyperglycemia in Chinese patients with non-alcoholic fatty liver disease. *Aliment Pharmacol Ther* 2006; **24**: 1215–22.

27 Urata Y, Okita K, Korenaga K, Uchida K, Yamasaki T, Sakaida I. The effect of supplementation with branched-chain amino acids in patients with liver cirrhosis. *Hepatol Res* 2007; **37**: 510–6.
28 Tsuchiya M, Sakaida I, Okamoto M, Okita K. The effect of a late evening snack in patients with liver cirrhosis. *Hepatol Res* 2005; **31**: 95–103.
29 Okamoto M, Sakaida I, Tsuchiya M, Suzuki C, Okita K. Effect of late evening snack on the blood glucose level and energy metabolism in patients with cirrhosis. *Hepatol Res* 2003; **27**: 45–50.
30 Vaag A, Alford F, Henriksen FL, Christopher M, Beck-Nielsen H. Multiple defects of both hepatic and peripheral intracellular glucose processing contribute to the hyperglycemia of NIDDM. *Diabetologia* 1995; **38**: 326–36.
31 Galgani JE, Heilbronn LK, Azuma K *et al.* Metabolic flexibility in response to glucose is not impaired in people with type 2 diabetes after controlling for glucose disposal rate. *Diabetes* 2008; **57**: 841–45.
32 Yokoyama H, Mori K, Emoto M *et al.* Non-oxidative glucose disposal is reduced in type 2 diabetes, but can be restored by aerobic exercise. *Diabetes Obes Metab* 2008; **10**: 400–7.
33 Kabadai UM, Eisenstein AB, Tucchi J, Pellicone J. Hyperglucagonemia in hepatic cirrhosis: its relation to hepatocellular dysfunction and normalization on recovery. *Am J Gastroenterol* 1984; **79**: 143–9.
34 Sumida Y, Yoneda M, Hyogo H *et al.* Validation of the FIB4 index in a Japanese nonalcoholic fatty liver disease population. *BMC Gastroenterol* 2012; **12**: 2.
35 Okumura H, Yamauchi R, Urano E *et al.* Examination of blood biochemistry index for tailor-made nutritional management of cirrhotic patients. *Jap J Nutr Assesment* 2011; **28**: 18–20 (in Japanese).

Glycobiology vol. 23 no. 6 pp. 736–744, 2013
doi:10.1093/glycob/cwt012
Advance Access publication on February 22, 2013

Identification and characterization of endo-β-*N*-acetylglucosaminidase from methylotrophic yeast *Ogataea minuta*†

Satoshi Murakami[2], Yuki Takaoka[2], Hisashi Ashida[3], Kenji Yamamoto[4], Hisashi Narimatsu[5], and Yasunori Chiba[1,2]

[2]Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba Central 6, 1-1-1 Higashi, Tsukuba 305-8566, Japan; [3]Department of Science and Technology on Food Safety, Faculty of Biology-Oriented Science and Technology, Kinki University, Kinokawa 649-6493, Japan; [4]Research Institute for Bioresources and Biotechnology, Ishikawa Prefectural University, Nonoichi 921-8836, Japan; and [5]Research Center for Medical Glycoscience, AIST, Tsukuba 305-8568, Japan

Received on December 12, 2012; revised on February 14, 2013; accepted on February 15, 2013

In four yeast strains, *Ogataea minuta*, *Candida parapolymorpha*, *Pichia anomala* and *Zygosaccharomyces rouxii*, we identified endo-β-*N*-acetylglucosaminidase (ENGase) homologous sequences by database searches; in each of the four species, a corresponding enzyme activity was also confirmed in crude cell extract obtained from each strain. The *O. minuta* ENGase (Endo-Om)-encoding gene was directly amplified from *O. minuta* genomic DNA and sequenced. The Endo-Om-encoding gene contained a 2319-bp open-reading frame; the deduced amino acid sequence indicated that the putative protein belonged to glycoside hydrolase family 85. The gene was introduced into *O. minuta*, and the recombinant Endo-Om was overexpressed and purified. When the enzyme assay was performed using an agalacto-biantennary oligosaccharide as a substrate, Endo-Om exhibited both hydrolysis and transglycosylation activities. Endo-Om exhibited hydrolytic activity for high-mannose, hybrid, biantennary and (2,6)-branched triantennary N-linked oligosaccharides, but not for tetraantennary, (2,4)-branched triantennary, bisecting *N*-acetylglucosamine structure and core-fucosylated biantennary N-linked oligosaccharides. Endo-Om also was able to hydrolyze N-glycans attached to RNase B and human transferrin under both denaturing and nondenaturing conditions. Thus, the present study reports the detection and characterization of a novel yeast ENGase.

[1]To whom correspondence should be addressed: Tel: +81-29-861-6083; Fax: +81-29-861-6083; e-mail: y-chiba@aist.go.jp
†We dedicate this article to Dr. Yoshifumi Jigami (1948–2011) with our gratitude.

Keywords: Endo-β-*N*-acetylglucosaminidase / Endo-Om / *Ogataea minuta* / transglycosylation / yeast

Introduction

Endo-β-*N*-acetylglucosaminidase (ENGase, EC 3.2.1.96) hydrolyses the *N*,*N*′-diacetylchitobiose core of asparagine-linked oligosaccharides on glycopeptides and glycoproteins. The ENGase's are detected among various organisms, where the enzymes are classified into two groups, glycoside hydrolase (GH) families 18 and 85, based on their amino acid sequence similarity to homologues found within the carbohydrate-active enzymes (CAZy) database (http://www.cazy.org). GH family 18 enzymes are known to comprise bacterial and fungal enzymes, such as Endo-H from *Streptomyces plicatus* (Tarentino and Maley 1974), Endo-F_1, -F_2 and -F_3 from *Elizabethkingia meningoseptica* (Trimble and Tarentino 1991), Endo-S from *Streptococcus pyogenes* (Collin and Olsén 2001), Endo-T from *Trichoderma reesei* (Stals et al. 2010) and Endo-FV from *Flammulina velutipes* (Hamaguchi et al. 2010). These enzymes are thought to contribute to the acquisition of nutrients while providing defense against the host anti-infective immune response (Collin and Olsén 2001; Renzi et al. 2011). GH family 85 enzymes are widely distributed from bacteria to mammals, and the genes of some these enzymes have been cloned and characterized, including those encoding Endo-A from *Arthrobacter protophormiae* (Takegawa et al. 1997), Endo-D from *Streptococcus pneumoniae* (Muramatsu et al. 2001), Endo-BH from *Bacillus halodurans* C-125 (Fujita, Takami et al. 2004), Endo-M from *Mucor hiemalis* (Fujita, Kobayashi et al. 2004), *Arabidopsis thaliana* At5g05460 and At3g11040 (Fischl et al. 2011), Endo-CE from *Caenorhabditis elegans* (Kato et al. 2002) and HsENGase from *Homo sapiens* (Suzuki et al. 2002).

Recent structural and mutational studies of the ENGases have identified key residues implicated in the enzyme activity and reaction mechanism (Abbott et al. 2009; Ling et al. 2009; Yin et al. 2009). Moreover, many application studies of transglycosylation, which is an alternative reaction of ENGase transferring a glycan to an alcoholic acceptor, have been performed for the purposes of the chemoenzymatic synthesis of glycopeptides and glycoproteins carrying homogeneous N-glycan decoration (Umekawa et al. 2008, Umekawa, Higashiyama et al. 2010; Umekawa, Li et al. 2010; Huang

et al. 2012). The physiological significance of GH family 85 ENGases in eukaryotes is poorly understood, whereas the bacterial paralogues are regarded as having same roles as GH family 18 enzymes *in vivo*. Additionally, although the both biological and enzymatic knowledge of ENGases has been increasing gradually, the existence of ENGase in yeasts (e.g. *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, which are model organisms for use in eukaryotic glycobiological studies) has not (to our knowledge) been previously reported. Hence, little information appears to be available concerning the role of ENGases in yeasts.

In a preliminary study, we detected an ENGase activity toward complex oligosaccharides when assaying a crude cell extract prepared from the yeast *Ogataea minuta*. This organism is a methylotrophic yeast that utilizes methanol as a sole carbon source; as we proposed in previous reports (Kuroda et al. 2006, 2007, 2008), *O. minuta* might serve as a useful host for exogenous glycoprotein production. In the present study, we identified ENGase homologous sequences from four yeast strains including *O. minuta* by database searches; in each of the four species, a corresponding enzyme activity also was confirmed in crude cell extract obtained from each strain. Moreover, we describe the cloning, expression and functional characterization of the *O. minuta* ENGase (Endo-Om)-encoding gene.

Results

Measurement of endogenous ENGase activity from O. minuta

We attempted to measure endogenous ENGase activity in crude cell extracts prepared from *O. minuta* by using 9-fluorenylmethyloxycarbonyl (Fmoc)-labeled asparagine-linked agalactobiantennary oligosaccharide (NGA2-Asn-Fmoc, Figure 1A) as an artificial substrate. The commercially available Endo-M from *M. hiemalis* was also assayed as a positive control. The reaction mixture was analyzed by high-performance liquid chromatography (HPLC). On a chromatogram of an Endo-M reaction mixture, a peak corresponding to NGA2-Asn-Fmoc (10.9 min) decreased concomitantly with the appearance of a hydrolysis product (*N*-acetylglucosamine (GlcNAc)-Asn-Fmoc) at 6.7 min (Figure 1B). Correspondingly, the product was also detected in the *O. minuta* reaction mixture (Figure 1B). For *O. minuta*, mean specific activity (measured in triplicate) was 7.3 μU/mg.

Cloning of an Endo-Om gene from O. minuta

To identify the ENGase homologous gene from *O. minuta*, a BLAST search against an in-house database of a sequence collection of genomic DNA from *O. minuta* (unpublished data) was carried out by using the amino acid sequence of Endo-M (accession no. BAB43869). A sequence encoding a putative ENGase from *O. minuta* was identified; the predicted protein shared 33% amino acid sequence identity and 45% amino acid sequence homology (*E*-value = 6e−050) with the conserved N-terminal domain of Endo-M. Based on this sequence, the full-length gene was amplified from *O. minuta* genomic DNA by polymerase chain reaction (PCR), and then subcloned and sequenced. The Endo-Om gene was intronless and contained an open-reading frame of 2319 bp encoding 772 amino acids (Figure 2). We deposited this sequence in

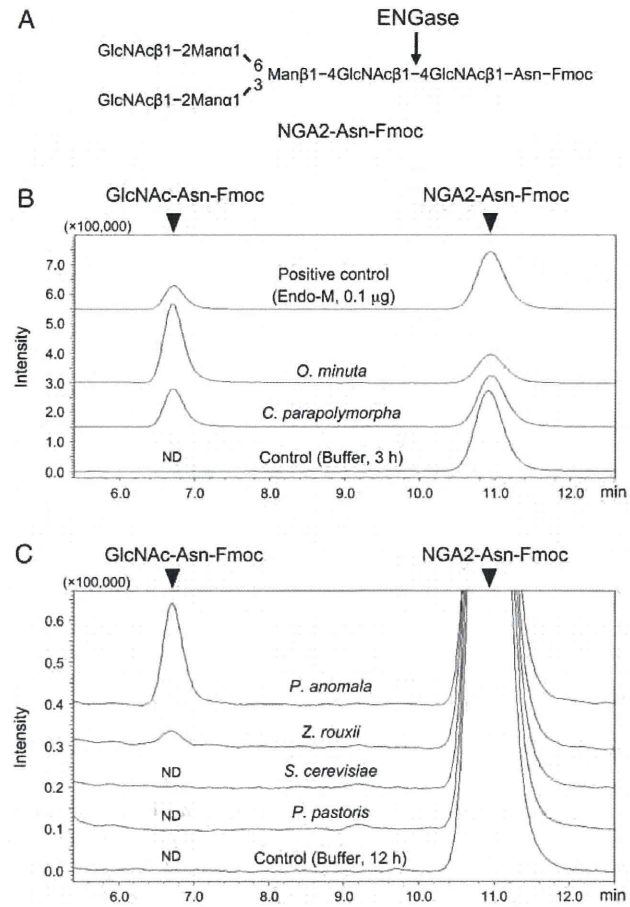

Fig. 1. Detection of endogenous ENGase activity from yeasts. Crude extracts (50 μg/reaction) prepared from yeast cells were incubated with 100 pmol of NGA2-Asn-Fmoc (**A**) in 100 mM sodium acetate buffer (pH 5.3) containing 0.5 M NaCl at 30°C for 3 h (**B**) or 12 h (**C**). Endo-M (0.1 μg/reaction) was used as positive control. Reaction mixtures were subjected to HPLC as described under Materials and methods. ND, not detectable.

DNA Data Bank of Japan (accession no. AB762085). The deduced amino acid sequence contained the region conserved among GH family 85 ENGases, including putative N-terminal active sites. Notably, the predicted protein included conserved residues Glu196 and Asn194, corresponding to the catalytic general acid/base and the essential residue for oxazolinium ion intermediate formation, respectively, as seen in Endo-M and other ENGases (Fujita, Kobayashi et al. 2004; Umekawa et al. 2008). The predicted molecular mass (Mr) and predicted isoelectric point based on amino acid sequence were 87,398 and 5.59, respectively.

Database searching and measurement of endogenous activity of ENGases from yeasts

The National Center for Biotechnology Information sequence database (http://www.ncbi.nlm.nih.gov) was surveyed to find other yeast ENGase homologues using the amino acid sequence of Endo-Om. The protein sequences encoding putative ENGase homologues were detected for three yeasts: *Candida*

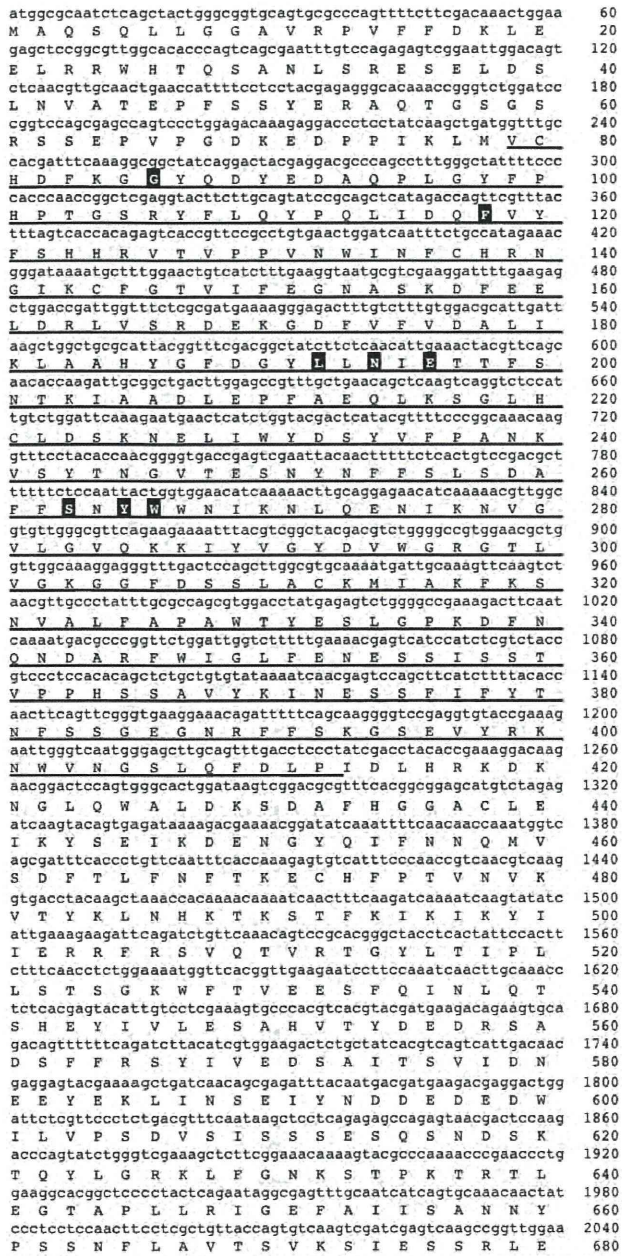

Fig. 2. Nucleotide sequence of the Endo-Om-encoding gene and the deduced amino acid sequence. Underlining indicates regions conserved among GH family 85 ENGases. The putative active sites are highlighted in black.

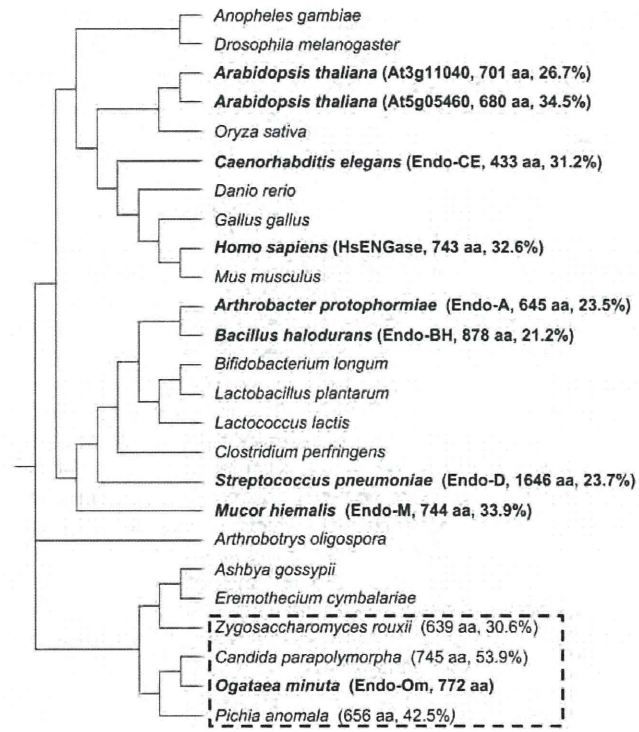

Fig. 3. Phylogenetic analysis of GH family 85 ENGases from yeasts and other organisms. The phylogenetic tree is based on an alignment of total length amino acid sequences and was constructed by the Clustal W web tool (http://www.genome.jp/tools/clustalw) with the neighbor-joining method. The dashed box indicates a group of yeasts. For yeast homologues and representative ENGases, amino acid numbers and homologies against Endo-Om are shown in parentheses. Accession numbers are as follows: Insects, *Anopheles gambiae* str. PEST (XP_310876) and *Drosophila melanogaster* (NP_573215). Plants, *Arabidopsis thaliana* ORF At3g11040 (NP_187715), ORF At5g05460 (NP_196165) and *Oryza sativa* Japonica Group (NP_001055263). Nematode, *Caenorhabditis elegans* (BAB84821). Fish, *Danio rerio* (NP_001074047). Bird, *Gallus gallus* (XP_427660). Mammals, *Homo sapiens* (AAM80487) and *Mus musculus* (NP_766161). Bacteria, *Arthrobacter protophormiae* (AAD10851), *Bacillus halodurans* C-125 (NP_241651), *Bifidobacterium longum* NCC2705 (NP_696499), *Lactobacillus plantarum* WCFS1 (NP_784014), *Lactococcus lactis* subsp. *lactis* IL1403 (NP_267648), *Clostridium perfringens* str. 13 (NP_561734) and *Streptococcus pneumoniae* (BAB62042). Fungi, *Mucor hiemalis* (BAB43869), *Arthrobotrys oligospora* (EGX51101), *Ashbya gossypii* (NP_986144) and *Eremothecium cymbalariae* (XP_003645294).

parapolymorpha (synonymous with *Hansenula polymorpha*, a methylotrophic yeast; accession no. EFW94296), *Pichia anomala* (accession no. CAC69142), and *Zygosaccharomyces rouxii* (accession no. XP_002495262). The amino acid sequence homologies to Endo-Om were 53.9, 42.5 and 30.6%, respectively (Figure 3). According to our search, no other sequence homologous to GH family 85 ENGase was detectable among published nucleotide and protein sequences for yeasts. A rooted phylogenetic tree of the GH family 85 ENGases from yeasts and various organisms is represented in Figure 3. The yeast ENGase homologues, including Endo-Om, clustered with putative fungal enzymes.

We also confirmed the presence of endogenous ENGase activity in crude cell extracts prepared from *C. parapolymorpha*, *P. anomala* and *Z. rouxii* (Figure 1B and C). The mean specific activities (measured in triplicate) were 3.8, 0.2 and 0.02 μU/mg, respectively. In contrast, the hydrolysis product was not detected even after 12 h in extracts of *S. cerevisiae* or

Pichia pastoris, two yeasts that do not harbor apparent ENGase-encoding sequences within their genomes (Figure 1C). Additional control experiments (data not shown) demonstrated that the enzyme activity was not detected in culture supernatants from any of these yeasts, including *O. minuta*.

Overproduction and purification of recombinant Endo-Om

The Endo-Om-encoding open-reading frame was amplified and expressed in *O. minuta* as described in our previous studies (Kuroda et al. 2007, 2008). The recombinant Endo-Om with N-terminal His-tag and 3 × FLAG epitope tag (Sigma-Aldrich, St. Louis, MO) was produced under the control of a methanol-inducible alcohol oxidase promoter in *O. minuta* protease-deficient strain TK10-1-2. The recombinant Endo-Om was purified using a HisTrap HP affinity column, and separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS–PAGE). A primary band was observed at a position corresponding to 91 kDa (Figure 4), which is consistent with the calculated Mr of Endo-Om with His-tag and 3 × FLAG-tag. The purified enzyme exhibited 9.92 mU/mg of specific activity under the standard assay conditions. Almost 100% of the activity was retained after 5 months' storage at −20°C in a solution of 10% glycerol solution at pH 7.4.

Various properties of recombinant Endo-Om

We analyzed the kinetic parameters of the purified Endo-Om. K_m and k_{cat} values of the recombinant Endo-Om toward NGA2-Asn-Fmoc were 5.54 mM and 5.91 s^{-1}, respectively. The enzyme exhibited optimum activity at 50°C and pH 5.5, and the activity was decreased <20% at pH 4.5 or 8.0. The effects of various divalent cations (10 mM) and chelators (5 mM) on the activity also were examined. The enzyme was inactive in the presence of Fe^{2+}, Cu^{2+} or Zn^{2+} ions; activity was unchanged in the presence of ethylenediaminetetraacetic acid (EDTA) or ethylene glycol tetraacetic acid (EGTA) (Figure 5).

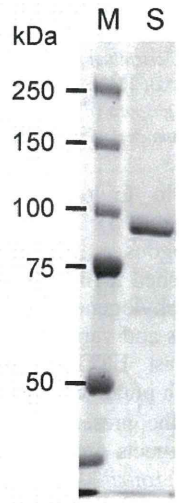

Fig. 4. SDS–PAGE of recombinant Endo-Om. The purified Endo-Om (lane S, total 2 μg) was separated using a 7.5% gel. M, molecular mass markers.

Transglycosylation activity of recombinant Endo-Om

The transglycosylation activity of recombinant Endo-Om was assayed using NGA2-Asn-Fmoc as a donor and 4-nitrophenyl-β-D-glucopyranoside (pNP-Glc) as an acceptor. The reaction mixture (with or without acceptor) was incubated at 30°C for 3 h, and the reaction products were subjected to HPLC (Figure 6A). A peak of transglycosylation product appeared at 4.3 min only in the sample with acceptor. The transglycosylation product was fractionated and subjected to matrix-assisted laser desorption ionization-quadrupole ion trap-time-of-flight mass spectrometry (MALDI-QIT-TOF MS). In MALDI-QIT-TOF MS, all mass spectra are obtained for Na^+ adduct ions, and sometimes K^+ adduct ions appear in mass spectra. The *m/z* values of molecular ions from the transglycosylation product are shown in Figure 6B. These values corresponded to the calculated value for $GlcNAc_2Man_3GlcNAcGlc$-pNP (Mr = 1397.44), indicating that $GlcNAc_2Man_3GlcNAc$ was transferred to pNP-Glc. Likewise, the transglycosylation activity of recombinant Endo-Om was also confirmed using GlcNAc as an acceptor (data not shown).

Substrate specificity of recombinant Endo-Om

Substrate specificity was examined with various pyridylamino (PA)-sugar chains as summarized in Table I. Relative activity was calculated by defining trimannosyl core structure (M3B) as 100%. The recombinant Endo-Om exhibited hydrolysis activity toward high-mannose, hybrid, biantennary and (2,6)-branched triantennary oligosaccharides, but not toward tetraantennary, (2,4)-branched triantennary, bisecting GlcNAc structure and core-fucosylated biantennary oligosaccharides. The enzyme exhibited substrate preference for high-mannose oligosaccharides (with the exception of M5A). The hydrolysis rate for biantennary oligosaccharides increased as the length of the sugar chain shortened. The recombinant Endo-Om showed almost the same relative activity for both biantennary (37%) and (2,6)-triantennary oligosaccharides (35%), and also for both hybrid (10.6%) and asialo biantennary (8.3%) oligosaccharides.

We also evaluated the hydrolytic activity under denaturing and nondenaturing conditions for N-glycoproteins (Figure 7). With RNase B, which carries a single high-mannose oligosaccharide composed of $Man_{5-9}GlcNAc_2$, the recombinant Endo-Om was able to almost completely hydrolyze the

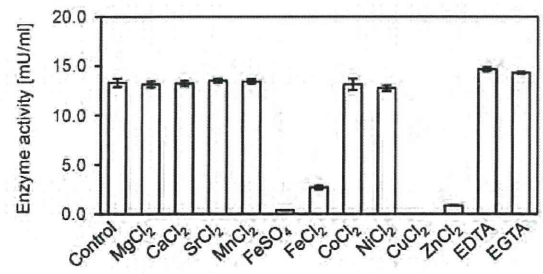

Fig. 5. Effect of divalent cations and chelators on enzyme activity. The recombinant Endo-Om was incubated with various divalent cations (10 mM), EDTA (5 mM), or EGTA (5 mM) at 30°C for 10 min, and then residual activity was assayed. Values are expressed as means ± SD ($n = 3$).

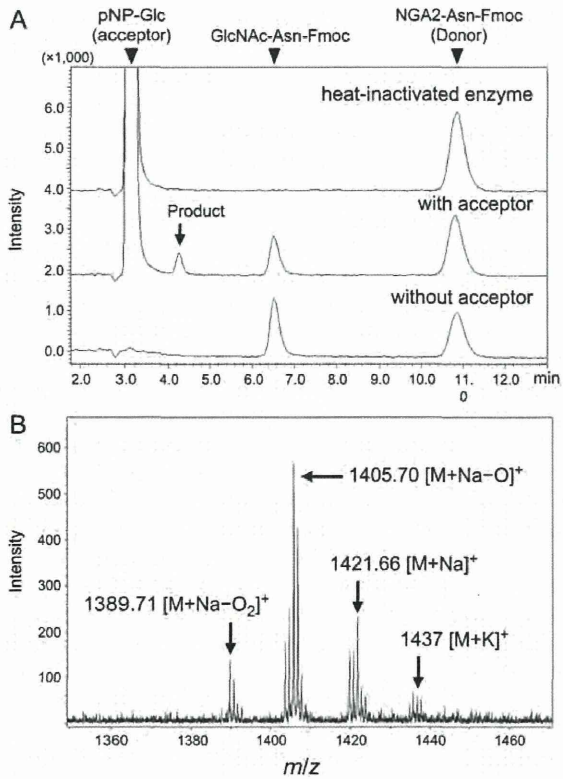

Fig. 6. Analysis of the transglycosylation activity of recombinant Endo-Om. (**A**) HPLC analysis of the transglycosylation reaction. The recombinant Endo-Om (4.2 μU) and 2 mM NGA2-Asn-Fmoc was incubated with or without 50 mM pNP-Glc as acceptor. (**B**) MALDI-QIT-TOF MS analysis of the transglycosylation product.

Table I. Relative activity of Endo-Om toward various PA-oligosaccharides

PA-oligosaccharide	Structure	Relative activity (%)[a]
PA-trimannosyl core (M3B)		100
PA-high-mannose oligosaccharide (M5A)		24.5
PA-high-mannose oligosaccharide (M6B)		184
PA-high-mannose oligosaccharide (M8A)		179
PA-high-mannose oligosaccharide (M9A)		146
PA-hybrid oligosaccharide		10.6
PA-hybrid oligosaccharide (bisecting GlcNAc)		ND[b]
PA-sialyl biantennary oligosaccharide		2.9
PA-asialo biantennary oligosaccharide		8.3
PA-core-fucosylated biantennary oligosaccharide		ND[b]
PA-agalacto-biantennary oligosaccharide		37
PA-agalacto-biantennary oligosaccharide (bisecting GlcNAc)		ND[b]
PA-agalacto (2,6)-branched triantennary oligosaccharide		35
PA-agalacto (2,4)-branched triantennary oligosaccharide		ND[b]
PA-agalacto tetraantennary oligosaccharide		ND[b]

● Mannose; ■ GlcNAc; ○ Galactose; ▲ Fucose; ◆ Neu5Ac

[a]The value obtained with M3B taken as 100.
[b]ND, not detectable.

oligosaccharide under both conditions (Figure 7A). Human transferrin harbors two N-glycosylation sites and contains mainly sialylated biantennary and triantennary oligosaccharides in the ratio 85:15 (Fu and van Halbeek 1992). With human transferrin, the recombinant Endo-Om partially hydrolyzed the oligosaccharide under nondenaturing conditions (Figure 7B, left). Additionally, removal from the substrate of terminal sialic acids (by sialidase treatment) enhanced the hydrolytic activity (lane labeled "Endo-Om + Sialidase"). The signal corresponding to the position of unhydrolyzed transferrin still remained under denaturing conditions even in the sialidase treatment sample with long incubation up to 26 h (Figure 7B, right). Fu and van Halbeek (1992) have reported that the human transferrin carries both (2,4)- and (2,6)-triantennary oligosaccharides. Based on the results provided in Table I, the signal is likely to be due to a transferrin molecule carrying (2,4)-triantennary oligosaccharides resistant to the hydrolysis by Endo-Om.

Discussion

Recent studies have revealed the distribution of members of GH family 85 ENGase in various organisms from bacteria to mammals, although ENGase homologues have not been reported in yeasts. In the present study, we identified the Endo-Om-encoding gene of the yeast *O. minuta*, and used the cloned gene to produce, purify and characterize the recombinant enzyme. To the best of our knowledge, this is the first report of the cloning and characterization of an ENGase gene isolated from yeast. The Endo-Om exhibited high activity toward an agalacto-biantennary oligosaccharide, NGA2-Asn-Fmoc, whereas the bacterial ENGases (e.g. Endo-A, -D and -BH) do not hydrolyze complex oligosaccharides (Takegawa et al. 1989; Muramatsu et al. 2001; Fujita, Takami et al. 2004). Moreover, Endo-Om also exhibited transglycosylation activity. The substrate specificity of Endo-Om was closely similar to that reported for Endo-M (Fujita, Kobayashi et al. 2004) and Endo-CE (Kato et al. 2002). Endo-Om preferred high-mannose oligosaccharides than hybrid and complex oligosaccharides as substrates. The core fucose and bisecting GlcNAc structures are not suitable for substrate. Furthermore, Endo-Om showed almost same relative activity for both biantennary and (2,6)-triantennary oligosaccharides, whereas no activity for (2,4)-triantennary and tetraantennary oligosaccharides. The relatve activity is also similar between the substrates of hybrid and asialo biantennary structures.

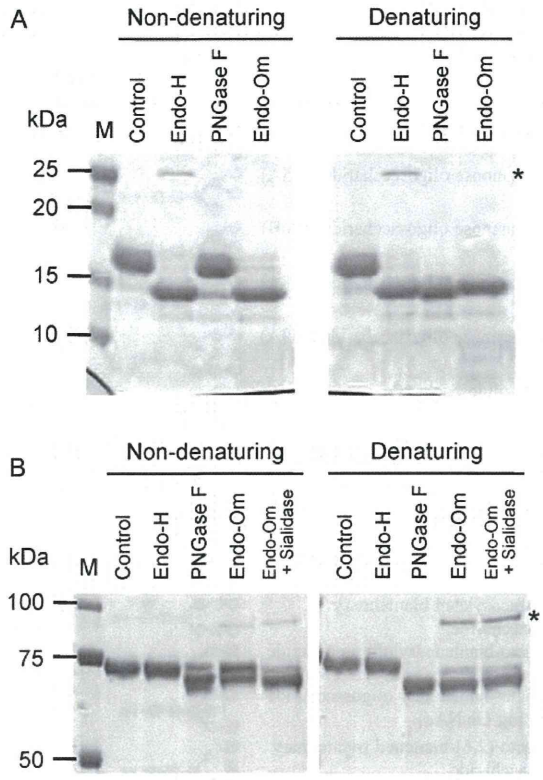

Fig. 7. Digestion of N-glycan on the glycoproteins by recombinant Endo-Om. RNase B (**A**) and human transferrin (**B**) were incubated with recombinant Endo-Om (4.49 μU) under denaturing or nondenaturing conditions as described under Materials and methods ". Appropriate negative (lacking enzyme) and positive (Endo-H and PNGase F) control reactions are provided. The positions of Endo-H (**A**) and Endo-Om (**B**) are indicated by asterisks.

These results suggest that the substrate recognition of Endo-Om depend on the structure attached to α-1,3-linked mannose residue of the trimannosyl core on the N-glycan; the Endo-Om activity is reduced by the addition of galactose and sialic acid residues and inhibited by β-1,4-branching structure.

In our study, sequences encoding putative ENGases of GH family 85 were identified from *O. minuta*, *C. parapolymorpha*, *P. anomala* and *Z. rouxii* by searching sequence databases, and corresponding enzyme activities were detected in these yeasts. Interestingly, genes encoding GH family 85 ENGase homologues were not identified among any other published yeast sequences, including those of *S. cerevisiae*, *S. pombe*, *P. pastoris* and *Candida boidinii*. Consistent with these *in silico* results, no ENGase enzyme activity was detected in crude cell extracts of *S. cerevisiae* or *P. pastoris*. Previously, Suzuki et al. (1998, 2002) made the same observation regarding the absence of ENGases in *S. cerevisiae*. These facts suggest that yeasts can be classified into two groups based on the presence of ENGase activity. In other work (data not shown), we constructed and characterized the growth phenotype of *O. minuta Endo-OmΔ* cells (i.e. a strain deleted for the Endo-Om-encoding gene). We did not detect any difference in growth phenotype in comparison with the parent wild-type strain (data not shown), suggesting that the ENGase-encoding gene is not essential in *O. minuta*. Similar lack of apparent phenotypes were reported for an *A. thaliana At5g05460* and *At3g11040* double-knockout line (Fischl et al. 2011) and for a *C. elegans eng-1* (Endo-CE) gene-deletion mutant (Kato et al. 2007).

Previous studies have shown that prokaryotic ENGases both GH families 18 and 85 contain an N-terminal signal sequence and are secreted into growth media and/or retained on cell surfaces (Fujita, Takami et al. 2004; Renzi et al. 2011). It is also known that some eukaryotic GH family 18 ENGase homologues have a signal sequence (Stals et al. 2010; Tzelepis et al. 2012). In contrast, the GH family 85 enzymes from eukaryotes such as HsENGase, Endo-CE and *A. thaliana* ENGases do not contain a signal sequence and are cytosolic proteins (Kato et al. 2002; Suzuki et al. 2002; Fischl et al. 2011). The yeast ENGase homologues reported here appear not to contain a signal sequence by hydropathy plot analysis and some web tool for the prediction of protein localization sites (data not shown). Consistent with these analyses, ENGase enzyme activity was not detected in culture supernatants, confirming that yeast ENGases are likely to be localized to the cytosol.

Our results suggest a possible function for yeast ENGases. We note that recombinant Endo-Om preferred high-mannose oligosaccharides as substrates and that the enzyme also was able to hydrolyze oligomannosyl N-glycan on RNaseB efficiently under both denaturing and nondenaturing conditions. In multicellular organisms, it has been reported that the various kinds of high-mannose-type free oligosaccharides (fOSs) that bear *N,N'*-diacetylchitobiose (GN2) or single GlcNAc (GN1) at the reducing end accumulate in the cytosol (Spiro 2004; Suzuki and Funakoshi 2006; Chantret and Moore 2008). Recent studies have demonstrated that the GN2-form of fOSs is generated through two distinct pathways. In the first route, fOSs are generated from dolicholpyrophosphate-linked oligosaccharides by an unknown mechanism in the endoplasmic reticulum (ER) lumen, and are exported to the cytosol in an adenosine triphosphate-dependent manner (Moore et al. 1995; Haga et al. 2009). In the second route, fOSs released by cytosolic peptide:*N*-glycanase (PNGase) from misfolded glycoprotein are retrotranslocated from the ER to the cytosol by an "ER-associated degradation" pathway (Spiro 2004; Chantret and Moore 2008; Yoshida and Tanaka 2010), where the GN2-form of fOSs is sequentially converted into GN1-form by cytosolic ENGase and then further de-mannosylated by α-mannosidase (Chantret et al. (2003); Chantret and Moore 2008). In *S. cerevisiae*, it has been reported that almost all GN2-form fOSs are generated through the latter pathway and that no GN1-form fOS is observed, a result that is in agreement with the absence of an ENGase-encoding gene in this yeast (Hirayama et al. 2010). Additionally, an increase of GN2-form fOSs and marked decrease of the GN1-form have been observed in ENGase-encoding gene-deficient lines of *A. thaliana* (Fischl et al. 2011) and *C. elegans* (Kato et al. 2007). Based on these observations, we propose that Endo-Om acts *in vivo* to trim GN2-form oligosaccharides to GN1-form oligosaccharides

and/or to directly digest oligosaccharides off of misfolded glycoproteins.

In summary, we identified the GH85 family ENGase genes from several yeast strains. The *O. minuta* gene was cloned and overexpressed. This ENGase enzyme, Endo-Om, showed hydrolyzing activity toward high-mannose, hybrid, biantennary and (2,6)-branched triantennary N-glycans, while also demonstrating transglycosylation activity. The physiological role of Endo-Om remains unclear, and further experiments are in progress to evaluate the difference between the *Endo-OmΔ* and wild-type *O. minuta* cells.

Materials and methods

Yeast strains and preparation of crude extract from yeast cells

Ogataea minuta IFO10746, *Candida parapolymorpha* DL-1 ATCC26012, *Pichia anomala* ATCC36904 and *Zygosaccharomyces rouxii* ATCC2623 were obtained from American Type Culture Collection (Manassas, VA). *Saccharomyces cerevisiae* W303-1A and *Pichia pastoris* SMD1168 were obtained from commercial vendors. Yeast cells were cultured at 26°C for 48 h in YMB medium (0.3% yeast extract, 0.3% malt extract, 0.5% peptone, 1% glucose) and then harvested and re-suspended in protein extraction buffer (50 mM sodium phosphate buffer (pH 7.4) containing 1.25 M NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.5% glycerol and 1 × complete protease inhibitor cocktail (Roche Diagnostics K.K., Tokyo, Japan)). The cells were disrupted with glass beads and the supernatant was assayed as a crude extract.

Cloning and expression plasmid construction of Endo-Om-encoding gene

Genomic DNA from *O. minuta* IFO10746 was prepared using Dr. GenTLE (Takara Bio, Shiga, Japan). Based on an in-house sequence database (unpublished data), the Endo-Om-encoding gene was amplified by sense primer (5′-CGATGACAAGG GATCATGGCGCAATCTCAGCTACTGG-3′) and antisense primer (5′-GCACCGTCTCGGATCTCACACCCAAACCTCA CTCC-3′), and then subcloned into pCR-Blunt II-TOPO vector using Zero Blunt TOPO PCR Cloning Kit (Invitrogen, Carlsbad, CA). After confirmation of the sequence, the PCR product was inserted into the *Bam*HI site of pOMEA1-6H3F using In-Fusion Advantage PCR Cloning Kit (Clontech Laboratories, Inc., Mountain View, CA) and recovered in *E. coli* DH5α cells. pOMEA1-6H3F is an expression vector for *O. minuta* that places the inserted gene under control of a methanol-inducible promoter while encoding an N-terminal His-tag and 3 × FLAG epitope tag. The pOMEA1-6H3F-Endo-Om was prepared by using a QIAprep Spin miniprep kit (Qiagen, Hilden, Germany), and then linearized with *Not*I and used for transformation of *O. minuta* protease-deficient strain TK10-1-2 (*och1Δ*, *pep4Δ*, *prb1Δ*, *ura3Δ*, *ade1Δ*; kindly provided by Kyowa Hakko Kirin Co., Ltd.). *Ogataea minuta* competent cells were prepared as described by Kuroda et al. (2006). Transformants were selected at 30°C for 48 h on plates of SDC−Ade medium (0.67% yeast nitrogen base without amino acids, 2% glucose, 0.5% casamino acid, 0.1 mg/mL uracil).

Expression and purification of recombinant Endo-Om

Transformed *O. minuta* was precultured at 30°C in 250 mL of YPD medium (1% yeast extract, 2% peptone and 2% glucose). To induce protein expression, cells were harvested and re-suspended in 200 mL of BMMYC medium (1% yeast extract, 2% peptone and 1.34% yeast nitrogen base without amino acids, 1% methanol, 2% casamino acid, 0.2 mg/mL adenine, 0.1 mg/mL uracil, 0.1 M potassium phosphate buffer (pH 6.0)) and cultured at 20°C for 48 h. Cells then were harvested and re-suspended in 20 mL of the protein extraction buffer described above. Cells were disrupted with glass beads and centrifuged. The supernatant was dialyzed against equilibrating buffer (20 mM sodium phosphate buffer (pH 7.4) containing 0.5 M NaCl and 0.5 mM PMSF). The dialyzed protein solution was centrifuged (27,000 × *g*, 4°C, 20 min), adjusted to 50 mM imidazole, and then filtered using a GD/X syringe filter (GE Healthcare Bio-science Corp., Piscataway, NJ). The protein solution was applied to a HisTrap HP column (GE Healthcare Bio-science Corp.) previously equilibrated with equilibrating buffer containing 50 mM imidazole. After the column had been washed, proteins were eluted by stepwise elution with equilibrating buffer containing 50, 100 and 200 mM imidazole. Endo-Om eluted with the 200 mM imidazole solution. The enzyme fraction was concentrated and exchanged with equilibrating buffer by using Amicon Ultra centrifugal filter units (Merck Millipore, Billerica, MA), and glycerol was added to a final concentration of 10%.

Enzyme assay

ENGase activity was assayed using NGA2-Asn-Fmoc (kindly provided by Otsuka Chemical Co., Ltd., Osaka, Japan) as a substrate. Commercially obtained Endo-M (0.1 μg/reaction; Tokyo Chemical Industry Co., Ltd., Tokyo, Japan) was used as a positive control. For each 10-μL reaction, 2 μL of enzyme solution was incubated with 100 pmol of NGA2-Asn-Fmoc in 100 mM sodium acetate buffer (pH 5.3) containing 0.5 M NaCl to enhance the stability of the enzyme. After incubation at 50°C for 1 h, the reaction was terminated by heating to 95°C for 5 min. The reaction mixture was subjected to HPLC with a hydrophilic interaction chromatography column, Shodex Asahipak NH2P-50 4E (4.6 × 250 mm; Showa Denko K.K., Tokyo, Japan). Elution was performed by isocratic elution with acetonitrile/0.2 M triethylamine acetate (pH 7.0) (at 57/43) at a flow rate of 1.0 mL/min at 40°C. Fmoc fluorescence was monitored at 315 nm with excitation at 265 nm. One unit was defined as the activity yielding 1 μmol of GlcNAc-Asn-Fmoc per minute under the assay conditions.

Substrate specificity for PA-sugar chains (Takara Bio, or Masuda Chemical Industry Co., Ltd., Kagawa, Japan) was determined using the same conditions as described above, except that 10 pmol of substrate was used for one reaction. The resulting reaction mixture was subjected to reverse-phase HPLC with a Cosmosil $5C_{18}$-AR-II (2.0 × 150 mm; Nacalai Tesque, Kyoto, Japan). Elution was performed using a linear gradient of 0.025–0.25% 1-butanol in 0.1 M ammonium acetate buffer (pH 4.0) at a flow rate of 0.5 mL/min at 40°C over 24 min. Elution was monitored at 400 nm with excitation at 320 nm.