Analysis of Glycosylation Site Occupancy Reveals a Role for Ost3p and Ost6p in Site-specific N-Glycosylation Efficiency*

Benjamin L. Schulz‡ and Markus Aebi§

Asparagine-linked glycosylation is the most common post-translational modification of proteins catalyzed in eukaryotes by the multiprotein complex oligosaccharyltransferase. Apart from the catalytic Stt3p, the roles of the subunits are ill defined. Here we describe functional investigations of the Ost3/6p components of the yeast enzyme. We developed novel analytical tools to quantify glycosylation site occupancy by enriching glycoproteins bound to the yeast polysaccharide cell wall, tagging glycosylated asparagines using endoglycosidase H glycan release, and detecting peptides and glycopeptides with LC-ESI-MS/MS. We found that the paralogues Ost3p and Ost6p were required for efficient glycosylation of distinct defined glycosylation sites. Our results describe a novel method for relative quantification of glycosylation occupancy in the genetically tractable yeast system and show that eukaryotic oligosaccharyltransferase isoforms have different activities toward protein substrates at the level of individual glycosylation sites. Molecular & Cellular Proteomics 8:357–364, 2009.

N-Glycosylation is important in protein folding and in modulating protein interactions, stability, and activity (1–4). Oligosaccharyltransferase (OTase) transfers glycans from a lipid pyrophosphate donor to selected asparagines in N-glycosylation sequons (NX(S/T); X ≠ P) in polypeptides in the lumen of the endoplasmic reticulum (ER) (5) or in the periplasm of some bacteria (6). Most OTases require a highly defined lipid-linked oligosaccharide substrate (7) while accepting a large number of glycosylation sites located on many different proteins. Glycosylation is proposed to be coupled to translocation into the ER in eukaryotes and occurs before protein folding (8, 9), whereas bacterial OTase can glycosylate acceptor sites located in flexible domains of folded proteins in vivo and in vitro (10). In higher eukaryotes, OTase is a multiprotein complex consisting of eight different subunits (Fig. 1), but in some protozoan, archaeal, and bacterial species N-glycosylation is catalyzed by a single protein OTase, homologous with Stt3p, the catalytic protein subunit of the eukaryotic OTase (5, 11). The roles of the additional subunits in the eukaryotic OTase are largely unknown but may allow regulation of activity toward preferred glycan and protein substrates. Alternatively as N-glycosylation requires flexible segments in the acceptor protein, the additional subunits may couple OTase to the translocation machinery or inhibit substrate protein folding to extend OTase substrate range or increase glycosylation efficiency.

In yeast, the presence of either Ost3p (UniProtKB number P48439) or Ost6p (UniProtKB number Q03723) results in two alternative OTase complexes (12, 13), which differ in their protein substrate-specific activities (12, 14). Ost3p and Ost6p share the same predicted topology of four transmembrane helices with an N-terminal hydrophilic domain located in the lumen of the ER. This domain is predicted to have a thioredoxin-like fold (15).

As we suspected that Ost3/6p affect glycosylation of a defined but unknown subset of glycosylation sites, we developed an MS-based analytical method that could concurrently measure the glycan occupancy of many sites in different proteins. This method involved enrichment of glycoproteins covalently linked to the yeast polysaccharide cell wall, release of glycans from these glycoproteins with endoglycosidase H leaving previously glycosylated asparagines residues “tagged” with a single N-acetylglucosamine (GlcNAc), and LC-ESI-MS/MS detection of peptides and glycopeptides after protease digestion. This method builds on previous studies that used glycosidase digests to identify previously glycosylated asparagine residues (16, 17). We used this method together with yeast genetics approaches to reveal that the Ost3/6p subunits of the yeast OTase influenced the glycosylation efficiency of specific sites. Our results show that in yeast the presence of multiple OTase isoforms increases the range of efficiently glycosylated protein substrates.

EXPERIMENTAL PROCEDURES

Chemicals were obtained from Sigma-Aldrich unless specified otherwise.

Yeast Strains—Yeast strains used were YG889 (MATa ade2–101 ura3–52 his3Δ200 tyr1Δost3::HIS3 Δost6::kanMX4) (12) and SS330 (MATa ade2–101 ura3–52 his3Δ2200 tyr1) (18). Cells were grown at permissive temperature in either YPD (2% Bacto peptone, 1% yeast extract, and 2% glucose) or, for maintaining plasmids, in minimal

From the Institute of Microbiology, Department of Biology, Eidgenössische Technische Hochschule (ETH) Zurich, 8093 Zurich, Switzerland

Received, May 16, 2008, and in revised form, August 12, 2008
Published, MCP Papers in Press, October 14, 2008, DOI 10.1074/mcp.M800219-MCP200

1 The abbreviations used are: OTase, oligosaccharyltransferase; ER, endoplasmic reticulum; GlcNAc, N-acetylglucosamine; rcf, relative centrifugal force; HexNAc, N-acetylhexosamine.

□ S
mass spectra were acquired with LTQ-ICR-FT in the mass range of proteins were digested with either trypsin (4°C, 16 h) and subjected to SDS-PAGE analysis. The proteins were excised from the gel bands, digested with trypsin (1 h, 37°C), and analyzed by LC-ESI-MS/MS with an LTQ-FT-ICR-MS instrument. Cysteines were reduced/alkylated by addition of dithiothreitol to 10 mM final concentration, and 2 mM iodoacetamide to 25 mM final concentration, and additional incubation with iodoacetamide to 25 mM and additional incubation with agitation at 30°C for 30 min followed by addition of iodoacetamide to 25 mM and additional incubation with agitation at 30°C for 1 h. Covalently linked proteins were then removed by washing the 16,000 rcf pellet five times with 50 mM Tris-HCl, pH 8, 2% SDS, 7 M urea, and 2 M thiourea. The pellets were resuspended in 2% acetonitrile and 0.2% formic acid and analyzed by LC-ESI-MS/MS.

Cell Wall Protein Sample Preparation—Yeast cells were grown to midlog phase in minimal medium at 23°C. Cells were harvested and lysed at 4°C using glass beads in 50 mM Tris-HCl, pH 7.5, 1 mM Complete protease inhibitor mixture (Roche Diagnostics), and 2 mM PMSF. Based on previously reported methods (19), covalently linked cell wall material was pelleted by centrifugation at 16,000 rcf for 1 min; washed three times with 50 mM Tris-HCl, pH 7.5; and resuspended in 50 mM Tris-HCl, pH 8, 2% SDS, 7 M urea, and 2 M thiourea. Cysteines were reduced/alkylated by addition of dithiothreitol to 10 mM and incubation with agitation at 30°C for 30 min followed by addition of iodoacetamide to 25 mM and additional incubation with agitation at 30°C for 1 h. Covalently linked proteins were then removed by washing the 16,000 rcf pellet five times with 50 mM Tris-HCl, pH 8, 2% SDS, 7 M urea, and 2 M thiourea followed by an additional five washes with 2% SDS. The pellet was resuspended in 2% SDS, 1× G5 buffer (New England Biolabs, Ipswich, MA), and 500 units of endoglycosidase H/100 μl (New England Biolabs) and incubated at 37°C with agitation for 16 h. Endoglycosidase H and SDS were removed by washing the 16,000 rcf pellet six times with 50 mM NH4HCO3, the pellet was resuspended in 50 mM NH4HCO3, and proteins were digested with either trypsin (4 μg/ml) or Asp-N (1 μg/ml) at 37°C with agitation for 16 h. Insoluble material was pelleted at 16,000 rcf for 1 min, and soluble peptides were desalted using C18 ZipTips (Millipore) and dried. Desalted peptides were resuspended in 2% acetonitrile and 0.2% formic acid and analyzed by LC-ESI-MS/MS.

Mass Spectrometry—Cell wall peptides and glycopeptides were analyzed by LC-ESI-MS/MS with an LTQ-FT-ICR-MS instrument (Thermo Scientific, Waltham, MA). Samples were injected into an Eksigent nano-HPLC system (Eksigent Technologies, Dublin, CA) with an autosampler and separated on a home-made reverse-phase column (75 μm × 80 mm) packed with C18 material (AQ, 3 μm, 200 Å; Bischoff GmbH, Leonberg, Germany). The column was equilibrated with solvent A (A: 3% acetonitrile and 0.2% formic acid in water). Peptides were eluted using the following gradient: 0–50 min, 0–60% B; 50–53 min, 60–97% B; 53–60 min, 97% B (B: 80% acetonitrile and 0.2% formic acid in water) at a flow rate of 0.2 μl/min. High accuracy mass spectra were acquired with LTQ-ICR-FT in the mass range of 300–2000 m/z and a target value of 5 × 10^5 ions. Up to four data-dependent MS/MS spectra were recorded in parallel at the ion trap of the most intense ions with charge state 2+ or higher using collision-induced dissociation. Target ions already selected for MS/MS were dynamically excluded for 60 s. General mass spectrometric conditions were as follows: normalized collision energy, 32% for MS/MS; ion selection threshold, 500 counts for MS/MS; activation, q = 0.25; and activation time of 30 ms for MS/MS acquisitions.

Data Analysis—Peak lists were extracted from raw data using Mascot Distiller (Version 2.1, Matrix Science, London, UK) with default parameters. Peptide identities based on MS/MS data were assigned using an in-house installation of Mascot (Version 2.1., Matrix Science) searching the Saccharomyces cerevisiae protein database (downloaded from European Molecular Biology Laboratory-European Bioinformatics Institute: fgcz_4921 4932_yeast_contaminants_20070811; 6068 sequences; 2,968,192 residues) with the following parameters: fixed modification of carbamidomethylated cysteines; variable modification of oxidized methionines, deamidated asparagines, and HexNAc-asparagines; no enzyme specified; 0.01 Da peptide tolerance; 0.6-Da fragment ion tolerance; and detected ion-specific charge state. Peptides with scores below the Mascot threshold for reliable identifications were excluded. Peptide abundances were determined manually from LC-ESI-MS data using Xcalibur (Version 2.0, Thermo Scientific) by summing the base peak chromatogram intensity for the entire isotopic distribution of each selected peptide ion over the elution peak. The glycosylation occupancy at a given sequon was determined by the abundance of the GlcNAc-modified peptide as a fraction of the sum of the GlcNAc-modified and unmodified versions of the same peptide. Each cell type was analyzed with five or more independent biological preparations, and each data point was determined as mean ± range. Data were compared using a two-sided Mann-Whitney test, a non-parametric statistical test appropriate to the data characteristics and low sample sizes. Two peptides detected contained two glycosylation sequons, both of which were occupied in wild type cells. In some cells these glycopeptides were also detected in a singly glycosylated form. However, as it was not possible to define the affected sequon, occupancy of these sites was combined in a single measurement (Asn^233 and Asn^237 of Crh2p as CRH2_233 and Asn^266 and Asn^268 of Ecm33p as ECM33_268).

Protein Sequence Analysis—Secondary structural elements and surface exposure of detected glycosylation sequons were predicted with JPred (20). Multiple sequence alignment based on amino acid sequence was performed with MUSCLE (multiple sequence comparison by log-expectation) (21), and a phylogenetic tree was built from this alignment with SCY-PHY (subfamily classification in phylogenomics) (22).

RESULTS

Quantitative Glycomics Method Development—We investigated the roles of the thioredoxin-like components Ost3p/6p in OTase function in vivo. As yeast OTase contains either Ost3p or Ost6p (12, 13), we generated strains where both the OST3 and OST6 loci were deleted (Δost3/Δost6). Expression of plasmid-encoded Ost3p or Ost6p then led to normal levels of uniform OTase (12). This allowed us to compare the phenotype of strains expressing either Ost3p- or Ost6p-containing OTase complexes. We then asked whether glycosylation of specific acceptor sites was affected by Ost3p/6p function. To detect such alterations in substrate recognition, we developed a novel analytical method to provide site-specific relative quantification of glycosylation occupancy. We took advantage of yeast cell wall architecture and prepared glycoproteins covalently attached to the polysaccharide cell wall matrix via glycosylphosphatidylinositol anchor remnants or alkali-sensitive linkages (19). After endoglycosidase H digested leaving glycosylated asparagines with a single GlcNAc,
protease digestion, and MS analysis, the N-glycan occupancy at many sites could be determined (Figs. 2 and 3).

Glycan release with endoglycosidase H provided a clear distinction between previously glycosylated and unglycosylated versions of the same peptide with a \( \Delta m/z \) of 203.08 Da and a difference in retention time of 5–90 s with the GlcNAc-modified peptide eluting before the unmodified peptide (Fig. 2, a and b). This allowed unambiguous identification and relative quantification of GlcNAc-peptide/peptide pairs. MS/MS spectra of GlcNAc-modified (c) and unmodified peptide (d) are shown.

![Fig. 2. LC-ESI-MS/MS identification and relative quantification of glycosylation occupancy at Asn177 of Crh1p (CHR1_177) from cells with wild type levels of purely Ost6p-containing OTase (Δost3/Δost6 pOST6). Single ion chromatograms corresponding to the GlcNAc-modified (a) and unmodified (b) peptide Phe175-Lys186 from Crh1p ([M + 3H] \(^3\) ions of \( m/z \) 587.26 and 519.57, respectively) are shown. The ratio of the area under the GlcNAc-peptide to the total of GlcNAc-modified and unmodified peptide gave a relative occupancy of 0.77 in this single measurement. MS/MS spectra and Mascot matched fragmentation of GlcNAc-modified (c) and unmodified peptide (d) are shown.](image-url)
ing to loss of asparagine-GlcNAc in a b or y fragment ion series (Fig. 2c and supplemental Figs. 1–31). Additionally deamidation of a non-glycosylated asparagine residue preceding a glycine residue was observed in four peptides containing glycosylation sequons CRH2_210, CWP1_45, ECM33_197, and GAS5_344 (Table I). NG sequences have been reported to be especially prone to deamidation during the course of standard proteomic sample preparation as the reduced steric hindrance of the small glycine residue favors the cyclic intermediate in the asparagine to aspartic acid conversion (23, 24).

Although no protease was specified in the search parameters, all identified glycopeptides corresponded to trypsin or Asp-N cleavage events with the exception of the glycopeptides containing the CRH2_28 and GAS5_24 glycosylation sites (Crh2p: Ala24–Glu25 and Gas5p: Ala20–Lys33, respectively) (Table I). These peptides are most likely the result of signal peptide cleavage as predicted between Ala24–Ala25 in Crh2p but between Ala19–Ala20 rather than at the predicted Ser22–Ser23 cleavage site in Gas5p (UniProt).

**OTase Site-specific N-Glycosylation Activity**—Analysis of cell wall glycoproteins in wild type yeast using our method (Fig. 3 and Tables I and II) robustly detected 26 N\(\times\)(S/T) sequons, 85% (22 of 26) of which were occupied with more N\(\times\)T than N\(\times\)S sequons glycosylated. Although all detected sequons were predicted to be surface-exposed, sequons in predicted secondary structural elements were less likely to be glycosylated (data not shown).

We next used our method to ask whether the two OTase isoforms defined by the presence of either Ost3p or Ost6p showed site-specific glycosylation activity. For this, we analyzed yeast cells with wild type levels of OTase but with exclusively Ost3p or Ost6p (Fig. 3 and Table II). Efficient glycosylation of 30% (6 of 20) of normally occupied sites required the presence of a particular OTase isoform with 25% (5 of 20) improved in the presence of Ost3p and 5% (1 of 20) with Ost6p. In the absence of both Ost3p and Ost6p, 50% (10 of 20) of normally occupied sites were underglycosylated (Fig. 3 and Table II).

**DISCUSSION**

N-Glycosylation is a general co- or post-translational modification affecting many glycosylation sites on numerous proteins. However, only ~70% of N\(\times\)(S/T) sequons in proteins translocated into the ER are actually glycosylated, and the fundamental processes controlling site-specific control of N-glycosylation are poorly understood. The Ost3p and Ost6p subunits of yeast OTase are predicted to contain a thioredoxin-like fold, and the two isoforms of OTase containing either one or the other of these two proteins show different protein-specific glycosylation activities. We were unable to further dissect this activity with standard immunoblot-based methods as these can only measure the average glycosylation state of multiple sites in a single protein and only then when appropriate antisera are available. Such specific methods are inherently inadequate for detailed investigation of the general process of N-glycosylation. To increase the resolution of measurement of OTase function, we therefore developed an MS-based method to concurrently measure N-glycan occupancy at many glycosylation sites in different proteins. Efficient MS detection of numerous glycosylation sites required enrichment of glycoproteins. We chose an enrichment based
on the natural covalent linkage of some proteins to the yeast polysaccharide cell wall matrix as this did not rely on N-glycans and therefore allowed detection of both glycosylated and non-glycosylated versions of the same sequon. Glycosylation occupancy measured by this method was relative as GlcNAc-asparagine modification undoubtedly affected MS ionization efficiency. In addition, only mature, successfully secreted proteins were analyzed. This biased the analysis against any sites that must be glycosylated for correct protein folding and ER export or sites where glycosylation would make correct folding impossible. Our analysis therefore allowed concurrent relative quantification of in vivo OTase activity on a wide variety of polypeptide substrates. Our method is a complementary analytical approach to the recently reported method allowing quantification of glycosylation occupancy at a lower number of predefined glycosylation sites in glycoproteins in clinical samples from patients with congenital disorders of glycosylation by multiple reaction monitoring LC-MS/MS (25).

Although designed for use in the yeast model system, it would be possible to modify our approach to allow analysis of the glycosylation occupancy of mammalian proteins. Nascent polypeptides still in the lumen of the ER could be released with endoglycosidase H before N-glycan modification in the Golgi rendered glycans endoglycosidase H-resistant. To allow more general analysis, peptide-N-glycosidase F could be used to release high mannose, complex, and hybrid glycans. However, upon peptide-N-glycosidase F glycan release, previously glycosylated asparagine residues are converted to aspartic acid with a resulting Δmass of only 1 Da between previously glycosylated and unglycosylated versions of the same peptide (potentially leading to confusion with the natural peptide isotopic distribution) and differences in retention

---

**Analysis of Site-specific N-Glycosylation Occupancy**

Peptides containing N-glycosylation sequons identified by mass spectrometry

Proteins covalently linked to the yeast polysaccharide cell wall were prepared, glycans were released by endoglycosidase H, and proteins were digested with trypsin or Asp-N, detected by LC-ESI-MS/MS, and identified using Mascot. All methionines are oxidized, and all cysteines are alkylated.

| Glycosylation site | Protein | Observed (m/z) | ΔMass | Expect | Sequence |
|--------------------|---------|----------------|-------|--------|----------|
| BGL2_284           | P15703  | 668.7856       | -0.0012 | 2.20e-04 | F | DEDWKPNTSGTS ↓ D |
| CCW14_87           | O13547  | 996.9353       | 0.0030 | 4.80e-08 | A | DAAYSAFKSSCSEQNASLG ↓ D |
| CRH1_177           | P53301  | 587.2603       | 0.0000 | 2.30e-02 | K | FHN3TYLTDAMDK ↓ T |
| CRH1_177           | P53301  | 519.5700       | 0.0000 | 8.40e-04 | K | FHN3TYLTDAMDK ↓ T |
| CRH2_28*           | P32623  | 786.8190       | 0.0003 | 2.90e-02 | A | ATFCNAQTACPE ↓ D |
| CRH2_28            | P32623  | 685.2800       | 0.0000 | 8.20e-04 | A | ATFCNAQTACPE ↓ D |
| CRH2_96*           | P32623  | 1095.5200      | 0.0000 | 7.90e-05 | K | DYSKSKLGNANTFLGNSVSEAG ↓ D |
| CRH2_96            | P32623  | 993.9700       | 0.0000 | 3.90e-06 | K | DYSKSKLGNANTFLGNSVSEAG ↓ D |
| CRH2_233*          | P32623  | 694.3300       | 0.0000 | 4.00e-02 | R | TLKPNYATNLTG ↓ Y |
| CRH2_310*          | P32623  | 1014.4774      | 0.0032 | 3.20e-06 | K | NGTSAYVYTTSEFLAK ↓ D |
| CRH2_310           | P32623  | 913.4300       | 0.0000 | 1.30e-08 | K | N9GTSAYVYTTSEFLAK ↓ D |
| CWP1_45            | P28319  | 974.4617       | -0.0021 | 4.50e-09 | R | SGSDLQYLSVSYEHDNLGTLK ↓ L |
| ECM33_197a         | P38248  | 960.9200       | 0.0000 | 2.00e-03 | S | DSQLGSSSLNDNLTTLAF ↓ G |
| ECM33_210*         | P38248  | 816.4192       | 0.0076 | 6.20e-04 | F | DNLWANNFLTLR ↓ D |
| ECM33_210          | P38248  | 714.8800       | 0.0000 | 3.50e-05 | F | DNLWANNFLTLR ↓ D |
| ECM33_268a         | P38248  | 881.7988       | 0.0037 | 1.20e-05 | K | AAFSNLTLYTGIFIAANTOLK ↓ V |
| ECM33_305          | P38248  | 1270.1612      | 0.0013 | 1.80e-04 | K | VQTGGAAEVGTGNSSTDLSLLSLK ↓ S |
| ECM33_305          | P38248  | 1168.6172      | 0.0072 | 3.50e-11 | K | VQTGGAAEVGTGNSSTDLSLLSLK ↓ S |
| ECM33_329a         | P38248  | 1033.4400      | 0.0000 | 7.10e-10 | R | GGAANFSSSNFSNPSNALK ↓ K |
| ECM33_329          | P38248  | 931.9000       | 0.0000 | 4.40e-12 | R | GGAANFSSSNFSNPSNALK ↓ K |
| GAS1_40*           | P22146  | 923.4270       | 0.0026 | 5.10e-07 | K | FFYNSNGOFSYR ↓ G |
| GAS1_57*           | P22146  | 1094.1471      | 0.0086 | 2.40e-12 | R | GVAYQADTANETSSTVNDPLANYESCR ↓ D |
| GAS1_95*           | P22146  | 667.6386       | 0.0073 | 3.00e-04 | R | VYANINTLTDHSEMCK ↓ A |
| GAS1_253*          | P22146  | 1216.5800      | 0.0100 | 7.10e-05 | K | NLSPVFSFSEYGCLEVPTR ↓ L |
| GAS1_253           | P22146  | 1115.0400      | 0.0000 | 5.80e-07 | K | NLSPVFSFSEYGCLEVPTR ↓ L |
| GAS3_201*          | Q03655  | 852.4230       | -0.0012 | 1.50e-02 | R | DMKQYISKHDNPIGVGSAA ↓ D |
| GAS3_350*          | Q03655  | 995.1604       | 0.0036 | 5.60e-02 | K | DDFVNLNSLQKYNVLPTKESEISS ↓ D |
| GAS3_350           | Q03655  | 927.4700       | 0.0000 | 4.00e-02 | K | DDFVNLNSLQKYNVLPTKESEISS ↓ D |
| GASS_24*           | Q08193  | 805.8934       | 0.0040 | 1.20e-02 | A | ASSSNSTSPSIEK ↓ G |
| GASS_60*           | Q08193  | 899.7400       | 0.0000 | 4.30e-07 | R | GDYQPGSSSNLTPLADASVCDR ↓ D |
| GASS_344*          | Q08193  | 948.7200       | 0.0000 | 1.30e-06 | K | VSNPENGNYGSTNNYSTCPYDEK ↓ G |
| TOS1_417           | P38288  | 526.2700       | 0.0000 | 5.20e-06 | K | AAVIFNSSDK ↓ T |
| YGP1_118*          | P38616  | 624.8200       | 0.0000 | 7.10e-03 | R | VVNETIQDK ↓ S |

* GlcNAc-modified asparagines are bold.

* Deamidated asparagines are underlined.

---
times typically much less than peak width (data not shown). The benefits of both a large GlcNAc tag on glycosylated asparagines and general glycan release could be obtained by using a mixture of exoglycosidases and endoglycosidases D and H, an approach compatible with most N-glycans, as previously reported (16). It is also possible that O-GlcNAc modification of peptides could be confused with GlcNAc-asparagine tags in cases in which MS/MS is not of sufficient quality to determine the exact modified residue.

In wild type yeast, the glycosylation site usage pattern observed with our novel method was in general agreement with results compiled from database analysis (26, 27). 85% of sites were completely occupied, whereas the remainder were not modified. NXT sequons tended to be glycosylated more often than NXS sequons, which is possibly a reflection of the general preference of OTase for NXS sequons reported in various studies. The observed tendency for unoccupied sequons to be located in regions of predicted secondary structure also suggests that local conformational factors are likely to be important in defining the glycosylation state of a given sequon.

To compare the protein substrate specificities of OTase isoforms containing either Ost3p or Ost6p, we used our analytical method to examine yeast cells with wild type levels of OTase but expressing only one of Ost3p or Ost6p. These analyses showed site-specific partial underglycosylation with different sites affected in an OTase isoform-dependent manner. The observed effect on glycosylation was not caused by a general reduction in OTase activity as different sites were underglycosylated in the absence of Ost3p or Ost6p. ∆ost3/∆ost6 double mutant cells showed underglycosylation of sites requiring Ost3p and Ost6p as well as of additional sites, correlating with the lower level of OTase in these cells due to OTase complex instability (12). These results confirmed that Ost3p and Ost6p did impart protein substrate specificity to OTase activity at the level of individual glycosylation sites with 30% of sites requiring a specific OTase isoform for efficient glycosylation. Ost3p improved glycosylation at more sites than Ost6p, correlating with its roughly 4 times greater abundance in wild type yeast and the more severe phenotype in ∆ost3 mutant cells (28). It has been proposed that the presence of either Ost3p or Ost6p causes OTase to associate with different translocons (13). However, Ost3p and Ost6p had largely complementary functions as few glycosylation sites specifically required the presence of one or the other protein for efficient glycosylation.

### Table II

| Glycosylation site | Wild type | pOST3 | ∆ost3/∆ost6 | pOST6 | p |
|-------------------|-----------|-------|-------------|-------|---|
| GAS1_40           | 1.00 ± 0.00^a | 1.00 ± 0.00 | 1.00 ± 0.00 | 1.00 ± 0.00 | 1.00 ± 0.00 |
| GAS1_95           | 1.00 ± 0.00 | 0.99 ± 0.01 | 1.00 ± 0.00 | 0.99 ± 0.01 | 1.00 ± 0.00 |
| GAS3_201          | 1.00 ± 0.00 | 1.00 ± 0.00 | 0.96 ± 0.05 | 1.00 ± 0.00 | 1.00 ± 0.00 |
| YGP1_118          | 1.00 ± 0.00 | 1.00 ± 0.00 | 1.00 ± 0.00 | 0.96 ± 0.07 | 1.00 ± 0.00 |
| GAS5_60           | 1.00 ± 0.00 | 1.00 ± 0.00 | 1.00 ± 0.00 | 0.96 ± 0.07 | 1.00 ± 0.00 |
| ECM33_197         | 1.00 ± 0.00 | 1.00 ± 0.00 | 1.00 ± 0.00 | 0.96 ± 0.07 | 1.00 ± 0.00 |
| CRH2_310          | 1.00 ± 0.00 | 1.00 ± 0.00 | 0.99 ± 0.01 | 0.98 ± 0.16 | 0.96 ± 0.01^b |
| GAS6_24           | 1.00 ± 0.00 | 0.99 ± 0.01 | 1.00 ± 0.00 | 0.95 ± 0.05 | 0.94 ± 0.08 |
| ECM33_305         | 0.99 ± 0.02 | 1.00 ± 0.00 | 0.97 ± 0.05 | 0.94 ± 0.08 | 0.94 ± 0.08 |
| ECM33_329         | 1.00 ± 0.00 | 0.99 ± 0.02 | 0.86 ± 0.12 | 0.93 ± 0.14 | 0.93 ± 0.14 |
| CRH1_177          | 1.00 ± 0.00 | 1.00 ± 0.00 | 0.80 ± 0.12 | 0.92 ± 0.06 | 0.92 ± 0.06 |
| ECM33_268         | 1.00 ± 0.00 | 1.00 ± 0.00 | 0.99 ± 0.01 | 0.91 ± 0.08 | 0.91 ± 0.08 |
| ECM33_210         | 1.00 ± 0.00 | 1.00 ± 0.00 | 0.99 ± 0.01 | 0.83 ± 0.13 | 0.83 ± 0.13 |
| CRH2_28           | 1.00 ± 0.00 | 1.00 ± 0.00 | 0.97 ± 0.03 | 0.78 ± 0.04 | 0.78 ± 0.04 |
| GAS1_57           | 1.00 ± 0.00 | 1.00 ± 0.01 | 0.97 ± 0.03 | 0.67 ± 0.13 | 0.67 ± 0.13 |
| GAS5_344           | 1.00 ± 0.00 | 0.97 ± 0.02 | 0.67 ± 0.18 | 0.44 ± 0.16 | 0.44 ± 0.16 |
| CRH2_233          | 1.00 ± 0.00 | 1.00 ± 0.00 | 0.99 ± 0.02 | 0.78 ± 0.11 | 0.78 ± 0.11 |
| GAS3_350          | 1.00 ± 0.00 | 1.00 ± 0.00 | 0.99 ± 0.02 | 0.78 ± 0.11 | 0.78 ± 0.11 |
| GAS3_253          | 0.99 ± 0.01 | 0.99 ± 0.01 | 0.99 ± 0.02 | 0.99 ± 0.01 | 0.99 ± 0.01 |
| BGL2_294          | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 |
| CCW14_87          | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 |
| CWP1_45           | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 |
| TOS1_417          | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 |

^a Error is S.E.
^b Bold indicates p < 0.05 in a two-sided Mann-Whitney test comparing wild type cells with cells containing ost3 and ost6 deletions but expressing high copy number plasmid-encoded Ost3p (pOST3), Ost6p (pOST6), or neither Ost3p nor Ost6p (p).
Efficient glycosylation of a third of N-glycosylation sites required the specific presence of either the Ost3p or Ost6p OTase isoform. Yeast Ost3/6p is therefore involved in controlling OTase activity toward a substantial proportion of protein substrates even though the Ost3/6p subunits are both the evolutionarily most recent additions to the OTase complex (5) and the physically most peripheral (29). Only the genomes of vertebrates and some fungi code for two different OST3/6 orthologues, and phylogenetic analysis indicates that gene duplication in these organisms occurred independently for these orthologues (Fig. 4). Thus, in some organisms an evolutionary advantage may originate from refined control of substrate-specific N-glycosylation. Conceivably fine tuned glycosylation is of special importance not only for fungi, which rely on glycosylated secretory enzymes for nutrition, but also for vertebrates where development depends on complex networks of interaction among secreted glycoproteins. Recently mutations in genes encoding the human homologues of Ost3/6p (termed implantation-associated protein and N33) have been associated with non-syndromic mental retardation, a novel form of congenital disorder of glycosylation (30, 31). Together with our results, this suggests that refined activity of the human OTase is required for efficient glycosylation of one or more proteins needed for the development of normal cognitive functions.

Our novel method will be of utility for more detailed analyses of the protein substrate specificity of OTase in the genetically tractable yeast model system. Many previous studies have investigated the characteristics of sequons that contribute to efficient glycosylation using model glycoproteins in ex vivo or in vitro systems or bioinformatics comparisons of glycoproteins from many different systems (27, 32–36). The systemic analysis of glycosylation site occupancy we present here represents a novel tool for the phenotypic analysis of defined yeast mutant strains altered in N-glycosylation processes in the ER and hence with reduced N-glycan occupancy. Such analysis will provide a data set of natural glycosylation sites to formulate algorithms that predict the use of
Analysis of Site-specific N-Glycosylation Occupancy

N-glycosylation sequons based on the primary amino acid sequence of putative glycoproteins.

Acknowledgments—We acknowledge the technical assistance of the staff of the Functional Genomics Center Zurich. We thank F. Fritsch, J. P. A. Grimshaw, M. S. Brozzo, R. Glockshuber, C. U. Stirmann, G. Capitani, and M. G. Grütter for fruitful discussions.

* This work was supported by the Swiss National Science Foundation and ETH Zürich within the framework of the GlycoNIT program. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Present address: School of Molecular and Microbial Sciences, The University of Queensland, Brisbane, Queensland 4072, Australia.

§ To whom correspondence should be addressed. Tel: 41-44-832-64-13; Fax: 41-44-632-11-48; E-mail: aebi@microbiol.ethz.ch.

REFERENCES

1. Helenius, A., and Aebi, M. (2004) Roles of N-linked glycans in the endoplasmic reticulum. Annu. Rev. Biochem. 73, 1019–1049
2. Ohitsu, K., and Marth, J. D. (2006) Glycosylation in cellular mechanisms of health and disease. Cell 126, 855–867
3. Sharon, N., and Lis, H. (1995) Lectins—proteins with a sweet tooth: function, structure, and applications. Cell 82, 1148–1150

36. Shakin-Eshleman, S. H., Spitalnik, S. L., and Kasturi, L. (1996) The amino acid sequence of N-linked core-glycosylation efficiency.
35. Mellquist, J. L., Kasturi, L., Spitalnik, S. L., and Shakin-Eshleman, S. H. (1996) The amino acid sequence of N-linked core-glycosylation efficiency.
34. Caragea, C., Sinapov, J., Silvescu, A., Dobbs, D., and Honavar, V. (2007) Biases and complex patterns in the residues flanking protein N-glycosylation sites. Glycobiology 17, 95–101
33. Kelleher, D. J., Banerjee, S., Cura, A. J., Samuelson, J., and Gilmore, R. (2007) Dolichol-linked oligosaccharide selection by the oligosaccharyltransferase in prokaryotic organisms. J. Cell Biol. 177, 29–37
32. Chen, W., and Helenius, A. (2000) Role of ribosome and translocon complex during folding of influenza hemagglutinin in the endoplasmic reticulum of living cells. Mol. Biol. Cell 11, 765–772
31. Weerapan, E., and Imperiali, B. (2006) Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. Glycobiology 16, 919–1010
30. Sprig, J., Bodmer, D., Wacker, M., Burda, P., and Aebi, M. (2005) The 3.4-kDa Ost4 protein is required for the assembly of two distinct oligosaccharyltransferase complexes in yeast. Glycobiology 15, 1936–1946
29. Yan, A., and Lennarz, W. J. (2005) Two oligosaccharyl transferase complexes exist in yeast and associate with two different translocons. Glycobiology 15, 1407–1415
28. Schwarz, M., Knauer, R., and Lehle, L. (2005) Yeast oligosaccharyltransferase consists of two functionally distinct sub-complexes, specified by the Ost3p or Ost6p subunit. FEBS Lett. 579, 6564–6568
27. Petrov, J. S., Siew, N. D., Glimmer, J. A., Martinez-Yamout, M., Dyson, H. J., and Skolnick, J. (2001) Genomic-scale comparison of sequence and structure-based methods of function prediction: does structure provide additional insight? Protein Sci 10, 1005–1014
26. Högland, P., Matthiesen, R., Ebertz, F., Heijrup, P., Roepstorff, P., Jensen, O. N., and Bunkenborg, J. (2007) An enzymatic deglycosylation scheme enabling identification of core fucosylated N-glycans and O-glycosylation site mapping of human plasma proteins. J. Proteome Res. 6, 3021–3031
25. Wada, Y., Taji, M., and Yoshida, S. (2004) Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. Anal. Chem. 76, 6560–6565
24. Vijayraghavan, U., Company, M., and Abelson, J. H. (1999) Isolation and characterization of pre-mRNA splicing mutants of Saccharomyces cerevisiae. Genes Dev. 13, 2012–2016
23. Yin, Y. Q., de Groot, P. W., Dekker, H. L., de Jong, L., Klis, F. M., and de Koster, C. G. (2005) Comprehensive proteomic analysis of Saccharomyces cerevisiae cell walls: identification of proteins covalently attached via glycosylphosphatidylinositol remnants or mild alkali-sensitive linkages. J. Biol. Chem. 280, 20984–20991
22. Cuff, J. A., Clamp, M. E., Siddiqi, A. S., Finlay, M., and Barton, G. J. (1999) JPred: a consensus secondary structure prediction server. Bioinformatics 14, 892–893
21. Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797
20. Splander, K. (1998) Phylogenetic inference in protein superfamilies: analysis of SH2 domains. Proc. Int. Conf. Intell. Syst. Mol. Biol. 6, 165–174
19. Krokhin, O. V., Antonovic, M., Enn, W., Wildins, J. A., and Standing, K. G. (2006) Deamidation of Asn-Gly- sequences during sample preparation for proteomics: consequences for MALDI and HPLC-MALDI analysis. Anal. Chem. 78, 6645–6650
18. Robinson, N. E., Robinson, Z. W., Robinson, B. R., Robinson, A. L., Robinson, J. A., Robinson, M. L., and Robinson, A. B. (2004) Structure-dependent nonenzymatic deamidation of glutaminyl and asparaginyl peptideptide. J. Pept. Res. 63, 426–436
17. Hülsmeier, A. J., Paesold-Burda, P., and Hennet, T. (2007) N-Glycosylation site occupancy in serum glycoproteins using multiple reaction monitoring liquid chromatography mass spectrometry. Mol. Cell. Proteomics 6, 2132–2138
16. Ben-Dor, S., Esterman, N., Rubin, E., and Sharon, N. (2004) Biases and complex patterns in the residues flanking protein N-glycosylation sites. Glycobiology 14, 95–101
15. Petrescu, A.-J., Milac, A.-L., Petrescu, S. M., Dwek, R. A., and Wormald, M. R. (2004) Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. Glycobiology 14, 103–114
14. Karamguly, D., Kelleher, D. J., and Gilmore, R. (1995) Functional characterization of Ost3p. Loss of the 34-kDa subunit of the Saccharomyces cerevisiae oligosaccharyltransferase results in biased underglycosylation of acceptor substrates. J. Cell Biol. 130, 567–577
13. Kelleher, D. J., Karamguly, D., Mandon, E. C., and Gilmore, R. (2003) Oligosaccharyltransferase isoforms that contain different catalytic STT3 subunits have distinct enzymatic properties. Mol. Cell 12, 101–111
12. Garshasi, M., Hadavi, V., Habibi, H., Kahrizi, K., Kariminejad, R., Bejhati, F., Tzschach, A., Najmabadi, H., Ropers, H. H., and Kuss, A. W. (2008) A defect in the TUSC3 gene is associated with autosomal recessive mental retardation. Am. J. Hum. Genet. 82, 1158–1164
11. Molinari, F., Foulquier, F., Tarpey, P. S., Morelle, W., Boissel, S., Teague, J., Edkins, S., Futreal, P. A., Stratton, M. R., Turner, G., Matthijs, G., Gecz, J., Munnich, A., and Colleaux, L. (2008) Oligosaccharyltransferase-subunit mutations in non-syndromic mental retardation. Am. J. Hum. Genet. 82, 1150–1157
10. Bause, E., and Legler, G. (1981) The role of the hydroxy amino acid in the triplet sequence Asn-Xaa-Thr(Ser) for the N-glycosylation step during glycoprotein biosynthesis. Biochem. J. 195, 639–644
9. Caragea, C., Sinapov, J., Silvescu, A., Dobbs, D., and Honavar, V. (2007) Glycosylation site prediction using ensembles of Support Vector Machine classifiers. BMC Bioinformatics 8, 438
8. Kasturi, L., Chen, H., and Shakin-Eshleman, S. H. (1997) Regulation of N-linked core glycosylation: use of a site-directed mutagenesis approach to identify Asn-Xaa-Ser/Thr sequences that are poor oligosaccharyl acceptors. Biochemistry 36, 415–419
7. Meilqust, J. L., Kasturi, L., Sptalnich, S. L., and Shakin-Eshleman, S. H. (1998) The amino acid following an Asn-Xaa-Ser/Thr sequence is an important determinant of N-linked core-glycosylation efficiency. Biochemistry 37, 6833–6837
6. Shakin-Eshleman, S. H., Sptalnich, S. L., and Kasturi, L. (1996) The amino acid at the X position of an Asn-Xaa-Ser sequence is an important determinant of N-linked core-glycosylation efficiency. J. Biol. Chem. 271, 6363–6366