A Search for Hyperglycosylation Signals in Yeast Glycoproteins*

Received for publication, June 15, 2004, and in revised form, July 19, 2004
Published, JBC Papers in Press, July 26, 2004, DOI 10.1074/jbc.M406678200

Raul Conde†‡§, Rosario Cueva‡, Guadalupe Pablo¶§, Julio Polaina¶§, and Germán Larriba‡¶

From the †Universidad de Extremadura, Departamento de Microbiología, F. de Ciencias, 06071 Badajoz and the §Instituto de Agroquímica y Tecnología de Alimentos, Consejo Superior de Investigaciones Científicas, 46980 Paterna, Valencia, Spain

*N-oligosaccharides of Saccharomyces cerevisiae glycoproteins are classified as core and mannan types. The former contain 13–14 mannoses whereas mannann-type structures consist of an inner core extended with an outer chain of up to 200–300 mannoses, a process known as hyperglycosylation. The selection of substrates for hyperglycosylation poses a theoretical and practical question. To identify hyperglycosylation determinants, we have analyzed the influence of the second amino acid (Xaa) of the sequon in this process using the major exoglucanase as a model. Our results indicate that negatively charged amino acids inhibit hyperglycosylation, whereas positively charged counterparts promote it. On the basis of the tridimensional structure of Exg1, we propose that Xaa influences the orientation of the inner core making it accessible to mannan polymerase I in the appropriate position for the addition of α-1,6-mannoses. The presence of Glu in the Xaa of the second sequon of the native exoglucanase suggests that negative selection may drive evolution of these sites. However, a comparison of invertases secreted by S. cerevisiae and Pichia anomala suggests that hyperglycosylation signals are also subjected to positive selection.

Protein glycosylation in eukaryotic cells is thought to play an essential role in many processes such as protein folding and transport, maintenance of protein and cell structure, and cell recognition and adhesion, as well as other functions. From the several types of protein glycosylation, N-glycosylation has received a great deal of attention not only because of its high frequency but also because several biochemical steps involved in this biosynthetic process are shared by yeast and humans, an indication that they have been conserved throughout evolution. These conserved steps occur in the membrane (i) or the lumen (ii and iii) of the ER and belong to three groups: (i) assembly of the precursor oligosaccharide, GlcNAc₂-Man₆Glc, on a lipid carrier (dolichol-PP), (ii) transfer of the oligosaccharide to the nascent or recently synthesized protein acceptor, and (iii) trimming of the three sugars and one mannose (for recent reviews, see Refs. 1 and 2).

However, once the glycoprotein leaves the ER, biochemical modification by trimming and/or addition of new sugars varies enormously between species and even between individual proteins of the same cell. This suggests that individual proteins carry the precise information for the final carbohydrate composition. In Saccharomyces cerevisiae, some of the protein-attached oligosaccharides leaving the ER (GlcNAc₂-Man₃) are poorly elongated with up to 13–14 mannoses (core-type), whereas many others are further elongated by the addition of an outer chain of up to 200 mannose residues in the Golgi apparatus (mannan-type), a process commonly known as hyperglycosylation. The outer chain consists of a backbone of α-1,6-mannoses with α-1,2 branches that are decorated with terminal α-1,3-mannose residues (1, 3). The biosynthesis of this complex is carried out by the ordered addition of mannoses in at least five biochemically defined steps (4–7) (Scheme 1). In S. cerevisiae, intracellular glycoproteins carry core-type oligosaccharides, whereas most extracellular glycoproteins carry outer-chain-elongated structures. This suggests that hyperglycosylation may protect protein molecules from environmental constraints.

It is well established that oligosaccharides are transferred from a lipid donor to specific asparagines in the tripeptide Asn-Xaa-(Ser/Thr), where Xaa is any amino acid except proline (8–12). However, the structural principles that govern the frequency of glycosylation of the different sequons still are a matter of controversy. A survey of the N-linked sites has indicated that Thr functions better than Ser in this process (13). Similarly, studies on glycosylation of the rabies virus glycoprotein have indicated that Ser containing sequons were poorly glycosylated in vitro relative to a similar series of sequons containing Thr in the third position (14). Also, work with yeast invertase showed that of the two overlapping sequons 4 and 5 of the protein (Asn92-Thr94-Ser95), the first one (Thr) was almost completely glycosylated, whereas the second (Ser) was barely glycosylated, if at all (15). Furthermore, a change in the tetrapeptide from NNSS to NTNS enabled both sequons to be glycosylated (16). However, the 100% glycosylation frequency in vivo of the Ser containing sequons present in the major yeast exoglucanase (Exg1) indicates that other parameters may have caused the bias in favor of Thr in the above mentioned studies or may have influenced our results (17). Additional studies in vitro again using the rabies virus glycoprotein as a model have indicated that introduction of specific amino acids, such as Trp, Asp, Glu, or Leu, in the X position convert the sequon to a poor oligosaccharide acceptor (18, 19).

Whereas studies on the identification of the structural features that influence the degree of occupancy of a sequon are scarce and controversial, no attempts have been reported to specifically characterize hyperglycosylation. Gene fusions between either carboxypeptidase Y or proteinase A and invertase suggest that...
the proteases bear dominant signals that suppress hyperglycosylation of the invertase domain present in the fusion protein (20, 21). In an effort to identify determinants that regulate the extension of the N-oligosaccharide elongation, we have analyzed the effect of sequon composition on this process. Analysis of most extracellular glycoproteins is a difficult task. For this reason, we have used S. cerevisiae Exg1, an extracellular protein amenable to study that has been well characterized in our laboratory. Exg1 is classified in family 5 of glycosyl hydrolases (22).

In vivo Exg1 glycosylation yields several glycoforms. One of these, Exg1b, contains 12% carbohydrate distributed into two short oligosaccharides, each consisting of a regular inner core whose outer chain is reduced to two or three residues of mannose, indicating that the α-1,6-mannose added by Och1 is capped by a stop-signal α-1,2-mannose, which may be elongated with a terminal α-1,3-mannose (23). These oligosaccharides are attached to both potential glycosylation sites (Asn165-Asn166-Ser167 and Asn325-Glu326-Ser327) present in the polypeptide (17, 24). Exg325 and Exg165 carry a single oligosaccharide attached to the second (Asn325) and the first (Asn165) glycosylation sites, respectively (17). Exg1a contains 30–40% carbohydrate and forms smears in SDS-acrylamide gels as do other heavily glycosylated yeast glycoproteins (i.e. invertase, acid phosphatase), and its synthesis is prevented in mutant mnn9. Analysis of glycosylation mutants has indicated that only the second oligosaccharide of Exg1b can be elongated to generate Exg1a, an indication that it should be more accessible than the first one to the enzymes that elongate the outer chain (25).

In this article we describe the effect of sequon composition, in particular the influence of the second amino acid of the tripeptide sequence (X) in the hyperglycosylation of Exg1, and we provide a structure-based hypothesis to explain our results. For that purpose, we have constructed mutated versions of the EXG1 gene in which the two sequons of the protein have been systematically mutated.

**EXPERIMENTAL PROCEDURES**

Yeast Strains and Growth Conditions—Wild type S. cerevisiae TD28 (MATa ura3-52 ins1-11 can1) and its ∆exg1 derivative, CV55, have been described before (17, 24). S. cerevisiae YS57-5A (MATa, och1::LEU2 leu2 ura3 his1 his3) and S. cerevisiae BYF109-1C (MATa can1-100 ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 kex2a::HIS3A) were kindly provided by Drs. Y. Jigami and R. Fuller, respectively. Yeast cells were maintained in YEPD medium (2% glucose, 1% yeast extract, 2% Bacto-peptone). For the production of external exoglucanase, cells were grown at 28 °C in liquid minimal medium supplemented with amino acids (26) until the middle exponential phase of growth.

Plasmid Constructs—Centromeric plasmids carrying the EXG1 gene in pRS316 (27) (pRB1) and its derivatives in which the first (pRB2, EXG10E), the second (pRB3, EXG1N0), or both (pRB4, EXG100) sequons were eliminated by substituting the corresponding Asn by Gln have been described before (17). All possible combinations between the X amino acids of the first and second sequons of Exg1 were performed by site-directed mutagenesis (QuikChange kit, Stratagene). First, using pRB1 as a template, a SalI/XbaI mutated fragment that included the first site as NES was obtained. It was then used to replace its counterpart in pRB1 and pRB4 to generate EXGEE and EXGEO, respectively. Similarly, a Kmpl/IclI mutated fragment including the second site as NNS was used to generate EXGN and EXGON from pRB1 and pRB4, respectively. Finally, EXG0N was obtained by introducing the SalI/XbaI mutated fragment into EXG0N. In a second set of experiments, we constructed all the possible variants of the second glycosylation site by introducing each one of other 19 amino acids (except proline) in the X position. For that purpose, the Kmpl/Ndel fragment was amplified by PCR using suitable oligonucleotides that carry the appropriate mutation, and the resulting product was used to replace its counterpart in EXG0N. All of the mutant constructs were confirmed by DNA sequencing.

**TABLE I**

A summary of the nature and amount of the Exg1p glycoforms secreted by the indicated glycosylation mutants

<table>
<thead>
<tr>
<th>Exg1 variant</th>
<th>First sequon</th>
<th>Second sequon</th>
<th>Associated glycoforms</th>
</tr>
</thead>
<tbody>
<tr>
<td>ExgNE</td>
<td>NNS</td>
<td>NES</td>
<td>Exg1a (10%) + Exg1b</td>
</tr>
<tr>
<td>ExgEE</td>
<td>NES</td>
<td>NES</td>
<td>Exg1a (10%) + Exg1b</td>
</tr>
<tr>
<td>ExgNN</td>
<td>NNS</td>
<td>NNS</td>
<td>Exg1a (30%) + Exg1b</td>
</tr>
<tr>
<td>ExgEN</td>
<td>NES</td>
<td>NNS</td>
<td>Exg1a (30%) + Exg1b</td>
</tr>
<tr>
<td>ExgE0</td>
<td>NES</td>
<td>QES</td>
<td>Exg1a-like</td>
</tr>
<tr>
<td>Exg0N</td>
<td>NNS</td>
<td>NNS</td>
<td>Exg1a-like (30%) + Exg1b</td>
</tr>
<tr>
<td>ExgN0</td>
<td>NNS</td>
<td>QES</td>
<td>Exg1a-like</td>
</tr>
<tr>
<td>ExgE0</td>
<td>NNS</td>
<td>QES</td>
<td>Exg1a-like (10%) + Exg1b</td>
</tr>
<tr>
<td>Exg00</td>
<td>NNS</td>
<td>QES</td>
<td>None</td>
</tr>
</tbody>
</table>

**SCHEME 1. Biosynthesis of N-linked oligosaccharides in S. cerevisiae.** Upper branch, mannn-type structures. Lower branch, core-type structures. Man, mannose; P, phosphate.
obtained by centrifugation of cells were concentrated and dialyzed using Amicon PM10 membranes and/or Centricon filters. Purification of the different glycoforms of exoglucanase was carried out by ion exchange chromatography column (TSK gel DEAE-5PW, 7.5 mm × 7.5 cm, TosoHaas). Exoglucanase activity was determined using p-nitrophenol-β-D-glucopyranoside as a substrate (26). SDS-PAGE and Western blots were performed as reported (28) using peroxidase to develop color. Standard deglycosylation reactions using endo H were carried out as described (25). (Endo H was a generous gift of Dr. F. Maley.)

**Exg1 Structure Analysis**—Crystal structure of the *S. cerevisiae* exo-1,3-β-exoglucanase has been determined recently (29).
RESULTS

Analysis of the Exg1 Glycoforms Secreted by S. cerevisiae Expressing Mutated Versions of EXG1 with Altered Sequon Composition—Wild type EXG1 or each one of the EXG1 constructs carrying all possible combinations of the first and second sequons was cloned in the centromeric plasmid PRS316 (27). These clones were then used to transform strain CV55 (H9004 exg1), and the transformants were grown at 28 °C for 15 h. The supernatant fluids, dialyzed and concentrated, were fractionated by ion exchange chromatography (HPLC). A summary of the nature and amount of the several glycoforms secreted by ectopic wild type EXG1 and glycosylation mutants is shown in Table I. Wild type EXG1 showed the typical profile, a minor and heterogeneous peak (Exg1a, 10%) preceded a major and sharp one (Exg1b, 90%) (Fig. 1A, panel a). Glycosylation mutants in which one sequon had been eliminated also yielded the expected results. Thus, mutant N0 only generated Exg165, whereas mutant 0E yielded an Exg1a-like glycoform and Exg325. Western blot analysis of Exg1b, Exg165, and Exg325 confirmed the nature of these forms indicating that the length of the short oligosaccharides attached to the modified sites was not significantly altered (Fig. 2). When the second sequon was eliminated (N325Q) and the first one was constructed NES (notice that this corresponds to the second sequon in native Exg1) (form E0), an Exg165-like enzyme was secreted indicating that the change introduced in the amino acid X of the first sequon (Asn to Glu) does not modify the elongation properties of the attached oligosaccharide. However, when the first sequon was eliminated (N165Q) and the second one was NNS (notice that Asn corresponds to the first sequon in native Exg1) (form 0N), only about 70% of the activity eluted as an Exg325-like form, whereas the residual 30% eluted as two peaks in the Exg1a region (Table I and Fig. 1A, panel c). This result suggested that the presence of Asn instead of Glu in the X residue of the second sequon increases the probabilities for elongation of the attached oligosaccharide. We should emphasize that because purified Exg1a and Exg1b have the same specific activity (32), the enzymatic activity exhibited by each glycoform can be taken as a good estimation of the amount of the associated Exg1p (see also below).

Mutant constructs with two glycosylation sites confirmed these observations and added new data (Table I). Thus, the EE construct (both sequons are identical to the second sequon of the wild type exoglucanase) generated an exoglucanase complement indistinguishable from the wild type counterpart by both HPLC (Fig. 1A) and SDS-PAGE (Fig. 2), indicating that the presence of Glu instead Asn at the X position of the first sequon does not alter the glycosylation pattern of wild type Exg1. On the other hand, the EN construct (both sequons exchange their positions) yielded two peaks in the Exg1a region (30%) and one eluting as Exg1b (70%) (Table I and Fig. 1A, panel b), the latter being further characterized by Western blotting as a form carrying two short oligosaccharides (Fig. 2). Finally, mutant NN (both sequons are identical to the first sequon of wild type exoglucanase) behaved as its EN counterpart (Table I and Fig. 2). Therefore, in all three constructs carrying Asn at the X position of the second sequon (0N, EN,
and NN), there is a significant increase in the amount of exoglucanase activity in the Exg1a region (30%) as compared with wild type (10%) (Fig. 1A). The absence of subglycosylated and/or non-glycosylated forms in transformants expressing wild type EXG1 or mutant constructs with two glycosylation sites as well as the absence of non-glycosylated Exg in transformants carrying constructs N0, 0E, E0, and 0N indicates that the transfer of oligosaccharides is very efficient and that there is enough lipid-linked oligosaccharide available to occupy all the sites offered by the nascent exoglucanase during its translocation into the lumen of the ER.

It should be noticed that regardless of their position in the molecule the QES sequon is always efficiently core-glycosylated. This contrasts with the results obtained with a variant of the rabies virus glycoprotein, which have indicated that the presence of Glu at the X position is associated with inefficient

Fig. 3. A, relative amounts of hyperglycosylated Exg produced by the constructs carrying the indicated amino acid at position X of the second sequon and lacking the first glycosylation site. B, elution profiles in HPLC of exoglucanase produced by wild type EXG1 (a) and constructs 0His (b) and 0Trp (c).
core glycosylation (18) and suggests that other protein signals must also control this process.

Characterization of the Hyperglycosylated Forms from ON, EN, and NN Constructs—Although mutant exoglucanases eluting in the Exg1a region seem to correspond to hyperglycosylated forms, the fact that an immature form of exoglucanase (form A) also elutes in this region prompted us to distinguish these possibilities. Form A is an endoplasmic reticulum form of Exg1 carrying a 21-amino acid propeptide; it is converted to mature form in the Golgi apparatus by elimination of the propeptide by the Kex2 protease (33, 34). Although Exg1a and form A co-elute in HPLC, their deglycosylated products have quite different retention times.

As shown in Fig. 1A (top of graph), following treatment with endo H, authentic form A is converted into a deglycosylated product eluting in HPLC as native Exg1b (fraction 19), whereas authentic form B yields a much more acidic compound (fraction 30). Similarly, after treatment with endo H, Exg1a from construct EN was quantitatively converted into a form undistinguishable from deglycosylated Exg1b. The absence of deglycosylated form A indicates that all of the exoglucanase eluting in the Exg1a region is indeed mature and hyperglycosylated. The same was true for the Exg1a-like counterparts generated by the Exg1 region is indeed mature and hyperglycosylated. The absence of deglycosylated form A indicates that all of the exoglucanase eluting in the Exg1a region indeed mature and hyperglycosylated. The same was true for the Exg1a-like counterparts generated by constructs EE and NN. As expected, deglycosylation of the Exg1a-like form generated by construct ON yielded a product that co-eluted with the deglycosylated product of the Exg325-like exoglucanase generated by the same construct.

These biochemical data were also supported by genetic evidence. Oligosaccharides from glycoproteins secreted by mutant Δoch1 are unable to elongate the inner core, but the protein portion of susceptible substrates (a factor or Exg1) are normally processed; accordingly, this mutant only secretes an Exg1b-like form with no traces of Exg1a. YS57–5A cells (Δoch1) transformed with the EN construct exclusively secreted Exg1b. Because the Δoch1 mutation does not prevent secretion of form A in a Δkex2 mutant (data not shown), we conclude that the Exg1a-like enzymes under study do not correspond to form A, but instead they behave as hyperglycosylated forms of mature exoglucanase.

In agreement with its hyperglycosylated nature, all of the Exg1a-like forms yielded by constructs NN, EN, and ON smeared when analyzed in SDS-PAGE where they exhibited a similar molecular size (Fig. 1C, panel a). Obviously, the absence of the first oligosaccharide (M, 3000) in the ON constructs is not enough to introduce detectable differences in the upper size limit. This observation indicates that the absence of the first oligosaccharide does not influence the degree of elongation of the second. It should also be noted that hyperglycosylated forms derived from constructs EN, NN, and ON elute clearly into two peaks, a property that we have extended now to wild type Exg1a where it was not as evident because of the low levels of this glycoform. Interestingly, as shown for the construct ON (Fig. 1C, panel b), the Exg1 molecules included in peak 1 have an average size larger than their counterparts from peak 2, indicating that in the former the oligosaccharide attached to the second site carries a more elongated outer chain. Although the precise origin and nature of these differences is under study, they likely are derived from different rounds of action of mannan polymerase I, mannan polymerase II, or both.

To further investigate the nature of the Exg1a-like enzyme secreted by NN and EN transformants, we purified the whole hyperglycosylated fraction by preparative HPLC and subjected it to treatment with endo H. The results were highly reproducible and almost identical for both transformants; accordingly, we will present only those results from construct NN. Time-course deglycosylation was followed by both ion exchange chromatography (Fig. 2A) and Western blots (Fig. 2B). The deglycosylation kinetics in terms of the enzymatic activity associated with the several glycoforms (Exg1a, peak 1 and peak 2; Exg1b, subglycosylated and Unglycosylated Exg1) is shown in Fig. 2C. The following conclusions can be derived. (i) As expected, the hyperglycosylated exoglucanase was quantitatively transformed into deglycosylated Exg1, which carries one GlcNAc attached to each sequon. The absence of more acidic products indicates that all the sequons of the hyperglycosylated molecules carried oligosaccharides. The final product of the reaction, deglycosylated Exg1, was also detected in Western blots as a 47-kDa band. (ii) A deglycosylation intermediate (fraction 27) is produced during the first stages of the deglycosylation reaction. Then, its levels decreased, and it was no longer visible after the levels of the substrate dropped to one-half. This intermediate eluted in the same position as the glycoform with a short oligosaccharide in the first position and a single GlcNAc in the second one and was also identified in the same samples by Western blots as a 50-kDa band; these features make it indistinguishable from the second intermediate detected during the deglycosylation of Exg1b (17). The most likely explanation is that the intermediate arises from the hyperglycosylated forms by elimination of the elongated residue. The levels of this intermediate are always very low because the endoglycosidase has more affinity for the oligosaccharide attached to the first position in such a way that not only it is produced at low levels, but it is also immediately transformed into deglycosylated enzyme (25). (iii) Both peaks of the hyperglycosylated region behaved similarly, although the deglycosylation kinetics of peak 1 was slightly faster. The preference of endo H for the first (short) oligosaccharide suggests the generation of an abundant intermediate carrying the elongated oligosaccharide attached to the second sequon. This intermediate could not be detected in the chromatograms indicating that its elution time is similar that of its precursor. Deglycosylation of the Exg1a-like enzyme from the ON transformant yielded similar results, except that there were no traces of the small intermediate (see Fig. 2). These results unambiguously demonstrate that the mutant glycoforms eluting in the Exg1a area in the HPLC column truly correspond to hyperglycosylated exoglucanase. Therefore, we conclude that substitution of Glu by Asn at the X position of the
second sequon significantly increases the hyperglycosylation efficiency.

Analysis of the Hyperglycosylation of Exg1 Variants with Amino Acid Substitution at the X Position of Sequon 2—To further analyze how hyperglycosylation is affected by the nature of the second amino acid of the sequon, we systematically changed the amino acid at position X of the second glycosylation site. To facilitate the analysis, these constructs were placed in a context in which the first glycosylation site had been eliminated (N165Q). As shown in Fig. 3A, analysis of the secreted Exg1 glycoforms derived from the new constructs indicated the presence of a substantial amount (up to 57%) of hyperglycosylated Exg1 when the amino acid X of the sequon was positively charged (Arg, His, Lys). By contrast, substitution of Glu by another negatively charged amino acid (Asp) scarcely improved the hyperglycosylation efficiency. Bulky lateral chains like those present in an aromatic amino acid (Trp, Tyr) also favor hyperglycosylation, although the elimination of OH group from Tyr (Phe) is deleterious. The rest of the amino acids improved (although to a lesser and variable degree) the amount of hyperglycosylated molecules. In these cases, it was noted that for amino acids belonging to the same class, the length of the lateral chain significantly influenced the percentage of hyperglycosylated molecules. For instance, Asp (18%) and Asn (32%) were better than Glu (7%) and Gln (17%), respectively. Similarly, Gly was better than Ala. To illustrate the differences in the amount of Exg1a, a comparison of the elution profiles of wild type Exg1 and construct 0H is shown in Fig. 3B (panels a and b). Finally, in agreement with previous studies (18, 19), the presence of Trp at the X position significantly decreased the efficiency of core glycosylation (Fig 3B, panel c). However, in contrast to these studies, Asn-Asp-Ser, Asn-Glu-Ser, and Asn-Leu-Ser were fully glycosylated in the context of the second sequon of Exg1.

All transformants grew at a similar rate and secreted a similar amount of exoglucanase. These observations make it unlikely that processes such as misfolding or a delay in the intracellular transport of some Exg1 variants could be responsible for the differences in the amount of the hyperglycosylated enzyme. In addition, when identical amounts of hyperglycosyl-

![Fig. 5. Representation of permitted orientations of Asn\textsuperscript{325} (sequon 2) resulting from the free rotation of its side chain.](http://www.jbc.org/)

Downloaded from http://www.jbc.org/ by guest on January 4, 2018
ated exoglucanase from each new construct (as determined by units of activity) were subjected to Western blot, all constructs yielded spots of similar intensity. The same was also true for the bands of the Exg325-like forms derived from each construct. This indicates that, at least within each group (Exg1a or Exg325-like), all variants have a similar specific activity, eliminating the remote possibility that changes in the X amino acid of the second sequon influence the activity of the enzyme. Overall, we conclude that the differences in hyperglycosylation among the several Exg1p variants should be ascribed to the ability of each one to be a substrate of mannosyltransferases and not to indirect effects in glycosylation.

**Structural Analysis of Exg1: Sequon 2 Environment**—The *S. cerevisiae* exoglucanase encoded by the EXG1 gene has been crystallized as a pseudohomodimer glycosylated at positions Asn

(A) strongly favors hyperglycosylation at Asn

barrel conformation characteristic of family 5 glycosidases (35) very similar to a closely related exoglucanase from *Candida albicans* (36). We have used the atomic coordinates of the protein to analyze the environment of sequon 2 at which hyperglycosylation occurs. The most striking result of the analysis reported in the previous section was that the presence of a basic residue at position 326 (position X) strongly favors hyperglycosylation at Asn

which has the opposite effect. Fig. 4 shows amino acid residues located in the vicinity of position 326. Likely, interactions between these residues lead to a local conformation more or less prone to hyperglycosylation. The extension of the asparagine-linked glycan core by successive addition of mannose residues that produce hyperglycosylation reaches a very large size and is subjected to steric restrictions. Therefore, positioning and orientation of Asn

must be critical to allow the action of the glycosylating enzyme machinery. Fig. 5 illustrates a hypothesis to account for the differences observed by the presence of either an acid or a basic residue at position 326. The disposition of residues neighbors to Glu

in wild-type Exg1 as determined by crystallographic analysis of the protein is shown in Fig. 5A. A similar disposition is assumed for the substitution E326D (Fig. 5B). However, replacement of Glu

by a basic residue (E326H, E326K, and E326R) allows an alternative positioning of Asn

(5C) that directs the growth of the glycosidic chain toward a region less restricted by steric constrains.

**DISCUSSION**

Previous studies in our laboratory have indicated that (i) both sequons of Exg1 and their immediate neighbors were more hydrophilic than the surrounding sequences, (ii) the first sequon was more hydrophilic than the second, and (iii) both sequons are located in turns of the polypeptide chain. Moreover, the first oligosaccharide is a better substrate for endo H but is never elongated, whereas the second is more resistant to endo H but has some probability (10%) of being extended with an outer chain (25).

Our new data indicate that exchange of sequons 1 and 2 does not modify their frequency of core glycosylation, which remained at 100%. This agrees with recent studies (37) indicating that, with the exception of proline, many amino acids at the Xaa position may support N-glycosylation; this process rather seems to depend to a great extent on the local regions surrounding the N-glycosylation sites, which were not modified in our experiments. In the case of rabies virus glycoprotein, the presence of a negatively charged amino acid (Glu or Asp) at X position led to inefficient core glycosylation (18, 19). This was not the case with Exg1 because the second sequon (NES) was 100% core glycosylated regardless of whether it was in the first or the second site. However, we found that (as described for rabies virus glycoprotein) the presence of Trp at the X position significantly inhibited N-glycosylation. Because both systems appear to be only distantly related, this coincidence suggest that NW(S/T) sites are not efficiently recognized by oligosaccharyl transferase, and accordingly when present they may give rise to protein isoforms carrying an occupied or unoccupied site. It is likely that the amino acids in the close vicinity of the NW(S/T) sequons also modulate their degree of occupancy.

On the other hand, our results indicate that the environment of a sequon in the folded protein may dramatically affect the frequency of hyperglycosylation of the attached core oligosaccharide. Thus, a displacement of the first sequon (NNS) of Exg1, which never elongates, to the second glycosylation site not only allowed its elongation but increased by 3-fold the hyperglycosylation frequency of the corresponding core oligosaccharide as compared with the original sequon (NES). This indicates that the environment of a sequon in the folded protein is crucial for its hyperglycosylation potential and that once this parameter is fixed, the nature of the amino acid X significantly modulates this process. It should be emphasized that the new glycoforms corresponded exclusively to hyperglycosylated forms Exg1 and were not contaminated with other glycoforms such as the unprocessed precursor, pro-Exg1, that elutes in the same region in our HPLC system. In addition, our chromatographic and electrophoretical analysis also revealed the presence of at least two populations of hyperglycosylated molecules that differ in the amount of carbohydrate. The elucidation of their precise nature and biosynthetic requirements could be relevant to the mode of action of mannan polymerases.

With regard to the nature of the Xaa residue, we have found that the native amino acid, glutamic acid, is the weakest promoter of hyperglycosylation of sequon 2. Substitution of Glu for any of the rest of amino acids improves to a variable degree the hyperglycosylation probabilities of the corresponding oligosaccharide. Positively charged residues (His, Lys, Arg) were the best promoters of hyperglycosylation, whereas the other negatively charged amino acid (Asp) was also supported poorly the extension of the inner core. Curiously, the same situation is found in sequons 2 (Asn

Asp-Thr), 3 (Asn

Asp-Ser) and 6 (Asn

).
Asp-Thr) of *S. cerevisiae* invertase, all of which carry an Asp in the X position and carry short (Man₈–14-GlcNAc₂) oligosaccharides (15, 38). In contrast to this rule, sequon 1 of invertase, which carries glutamic acid in the second position (Asn 4-Glu-Thr), is hyperglycosylated (Table II). However, this sequon is so close to the amino terminus of the protein, and its local environment has such a high hydropathic index (15) that other constraints may influence its elongation degree. On the other hand, the absence of a negatively charged amino acid does not guarantee an elongated oligosaccharide as expected from the short oligosaccharides attached to the first sequon of Exg1 (Asn₁₆₅-Asn-Ser) or sequons 12 (Asn₃₆₅-Ser-Thr) or 13 (Asn₃₇₉-Thr-Thr) of invertase, which never elongate (Table II). Clearly, in addition to the nature of the amino acid X, other constraints of a different nature (probably related to the tertiary structure or to the charge of surrounding regions) avoid hyperglycosylation of some oligosaccharides. In any case, the fact that the most inhibitory amino acid (Glu) was the one present in the second site of Exg1 suggests that it has been selected to avoid hyperglycosylation.

Negative selection does not seem to be the only way to regulate hyperglycosylation. Other yeast, like *Pichia* sp., do not elongate extensively the N-oligosaccharides that contain 8–14 mannoses, and even heterologous expression of the *S. cerevisiae* SUC2 gene in *P. pastoris* or *P. angusta* results in hypoglycosylation of otherwise hyperglycosylated oligosaccharides (39, 40). The most likely conclusion is that *Pichia* sp. lacks the outer chain synthesizing machinery described in *S. cerevisiae* and accordingly their glycoproteins should have evolved under different constraints. Studies with the invertase of *P. anomala* encoded by the INV1 gene have indicated that this protein contains 10 potential glycosylation sites, all of which carry short oligosaccharides (8–14 mannoses). Only six of these sequons (1, 2, 3, 4, 8, and 9) are fully or partially conserved in the invertase of *S. cerevisiae* (sequons 4, 5, 6, 7, 10, and 12 respectively), whereas the rest do not have a clear counterpart in the latter (39) (Table II). Expression of the *P. anomala* INV1 gene in *S. cerevisiae* yields an invertase carrying all the potential glycosylation sites (10 of them) occupied, but this invertase has oligosaccharides that are even shorter than their counterparts from the autologous enzyme, an indication that the oligosaccharides are not a substrate for the mannosylpolymerases that elongate the inner core in autologous proteins (40). Because the corresponding sequons have evolved in the absence of hyperglycosylation pressures, their sequences could be considered as random, and accordingly a hyperglycosylation signal must require a positive selection. A closer look at both invertases indicates that four of the six fully or partially conserved sequons of the *S. cerevisiae* enzyme (sequons 4, 6, 10, and 12) contain short-chain oligosaccharides (although only one (sequon 6) carries an acidic amino acid (Asp) in the X position) and a fifth one (sequon 5, Asn) is non-glycosylated. In addition,
three of the four hyperglycosylated sequons in Suc2p (1, 9, 11) are not conserved in P. anomala, and the partially conserved one (ScAsn145-Thr versus PaAsn145-Ser-Ser) is glycosylated with a frequency lower than 50% (40) (Table II). Accordingly, it seems that at least some specific sequons in S. cerevisiae have been tailored and selected during evolution to carry long oligosaccharides by a combination of the nature of the acceptor tripeptide and the position at specific places of the protein. In this regard, we have found that bulky aromatic amino acids at the X position, such as Trp and Tyr, also favor hyperglycosylation.

An important issue arising from our experimental approach is to identify the reason why the substitution of an acidic amino acid by a neutral one or vice versa may affect hyperglycosylation of the corresponding oligosaccharide. An obvious answer is that the inner core oligosaccharide elongates when the acceptor mannose added by Och1p is more accessible to the mannosyltransferases that elongate the outer chain. Several possibilities may account for this situation. First, changes in the local environment of the corresponding oligosaccharide. An obvious answer is that the final shape of a secreted glycoprotein is achieved. Core glycosylation takes place in the ER at a moment when the orientation of the inner core, which in turn depends on the orientation adopted by Asn325 is determined by a combination of the nature of this amino acid and in particular by the nature of the amino acid in position 92 (Figs. 4 and 5). The crystal structure of wild type Exg1 glycosylated at Asn165 and Asn225 (29) shows that the glycan at Asn225 lays tangentially to the globular protein monomer. Further extension of the sugar chain seems to be restricted by its proximity to the protein surface. Changes in the orientation of Asn225, particularly those proposed in Fig. 5, are expected to displace the glycan, thus favoring hyperglycosylation. When oriented toward the exterior, the glycan core, which is first modified by the addition of one mannose by the action of Och1p, can be further elongated by the mannosyl polymerase I complex. The crystal structure of Exg1 also suggests that a steric impediment may also be the reason why the first sequon (Asn165) is never elongated.

An interesting model to explain the biosynthesis of either extended or core-type oligosaccharides has been recently proposed (7). In this model, the protein carrying N-glycan chains with the first α1,6-mannose added by Och1p interacts with Mnn9p and depending on the nature of this interaction Mnn9 would add either an α1,6- or α1,2-mannose. Our results support this model and suggest that once the protein-protein interaction between Mnn9 and its substrate is fixed, the orientation of the oligosaccharide plays an important role (Fig. 6). This orientation has been selected to either prevent the elongation of some oligosaccharides or to promote hyperglycosylation of others.

Acknowledgments—We thank Dr. Frank Maley for his generous gift of Eno H. We also thank Dr. Luis Rodriguez for sharing with us the results presented in Table II.

REFERENCES
