A search for hyperglycosylation signals in yeast glycoproteins

Raúl 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, CSIC, Paterna, Valencia, Spain

**Corresponding author:

Germán Larriba
Departamento de Microbiología,
F. de Ciencias
Universidad de Extremadura
06071 Badajoz, Spain
Phone & Fax: 34-924-289428
E-mail: glarriba@unex.es

*This work was supported by grants DGICYT PB98-0889 and MCYT BIO2003-09180-C02-02 to GL, IPR00C033 to RC and MCYT BIO2000-1279-C02-01 to JP
‡Supported by fellowships from Junta de Extremadura

1The abbreviations used are: ER endoplasmic reticulum; MPI mannan polymerase I; MPII mannan polymerase II; CPY, carboxypeptidase Y; Exg1, exoglucanase; proExg1, Exg1 precursor unprocessed by the Kex2 protease; endo H, endoglicosidase H. (ver otras)
SUMMARY

N-oligosaccharides of *Saccharomyces cerevisiae* glycoproteins are classified as core- and mannan-types. The former contain 13-14 mannoses whereas mannan-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.
INTRODUCTION

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, 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 due not only to 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-Man9Glc3 on a lipid carrier (dolichol-PP), ii) transfer of the former to the nascent or recently synthesized protein acceptor, and iii) trimming of the three glucoses and one mannose (for recent reviews see 1, 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 S. cerevisiae, some of the protein-attached oligosaccharides leaving the ER (GlcNAc2-Man8) 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 X 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 (N92-N93-T94-S95), the first one (Thr) was almost completely glycosylated, but the second (Ser) was barely glycosylated, if at all (15). Furthermore, a change in the tetrapeptide from NNTS to NNSS 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 using again 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).

While studies on the identification of the structural features that influence the degree of occupation of a sequon are scarce and controversial, no attempts have been reported to specifically characterize hyperglycosylation. Gene fusions between either CPY 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 exoglucanase (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). They are attached to both potential glycosylation sites (N165-N166-S167 and N325-E326-S327) present in the polypeptide (17, 24). Exg325 and Exg165 carry a single oligosaccharide attached to the second (N325) and the first (N165) glycosylation sites respectively (17). Exg1a contains 30-40% carbohydrate, 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 the it should be more accessible than the first one to the α-1,6-mannosyltransferase that elongates 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 this 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 Saccharomyces cerevisiae TD28 (MATa ura3-52 ino1-11 can1) and its Δexg1 derivative, CV55, have been described before (17, 24). S. cerevisiae YS57-5A (MATa, och1::LEU2 leu2 ura3 his1 his3) and S. cerevisiae BFY 109-1C (MATa can1-100 ade2-1 his3-11,-15 leu2-3,-112 trp1-1 ura3-1 kex2Δ::HIS3A) were kindly provided by Drs. Y. Jigami and R. Fuller respectively. Yeast cells were maintained in YEPD medium. For the production of external exoglucanase, cells were grown at 28ºC in liquid minimal medium supplemented with amino acids (26), until 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 (Quick Change kit, Stratagene), First, using pRB1 as a template, a SalI/XbaI mutated fragment which included the first site as NES was obtained. It was then used to replace its counterpart in pRB1 and pRB4 in order to generate EXGEE and EXGE0 respectively. Similarly, a KpnI/NaeI mutated fragment including the second site as NNS was used to generate EXGNN and EXG0N from pRB1 and pRB4 respectively. Finally, EXGEN was obtained 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 \textit{KpnI}/\textit{NaeI} fragment was amplified by PCR using suitable oligonucleotides which carry the appropriate mutation and the resulting product was used to replace its counterpart in \textit{EXG00}. All the mutant constructs were confirmed by DNA sequencing.

\textbf{Purification and characterization of Exg1-} Culture supernatants, 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, 7.5mm x 7.5 cm, Tosohaa). Exoglucanase activity was determined using p-nitrophenol-\(\beta\)-D-glucopyranoside (p-NPG) 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.

\textbf{Exg1 structure analysis -} Crystal structure of the \textit{S. cerevisiae} exo-(1,3)-\(\beta\)-exoglucanase has been determined recently (29). Atomic coordinates were retrieved from the Protein Data Bank (PDB accession code 1H4P). Protein structure was analyzed with Swiss-Model and Swiss-PdbViewer (30, 31).

\textbf{RESULTS}

\textit{Analysis of the Exg1 glycoforms secreted by S. cerevisiae expressing mutated versions of EXG1 with altered sequon composition}

Wild type \textit{EXG1} or each one of the \textit{EXG1} constructs carrying all possible combinations of the first and second sequons were cloned in the centromeric plasmid PRS316 (27). These clones were then used to transform strain CV55 (\(\Delta\text{exg1}\)) and the transformants were grown at 28\(^\circ\)C for 15h. 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 \textit{EXG1} and glycosylation mutants is shown in Table 1. Wild type \textit{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 Exg\(_{165}\), whereas mutant 0E yielded an Exg1a-like glycoform and Exg\(_{325}\). Western blot analysis of Exg1b, Exg\(_{165}\) and Exg\(_{325}\) 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), a Exg\(_{165}\)-like enzyme was secreted indicating that the change introduced in the amino acid X of the first sequon (N→E) 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 N corresponds to the first sequon in native Exg1)(form 0N), only about 70% of the activity eluted as an Exg\(_{325}\)-like form whereas the residual 30% eluted, as two peaks, in the Exg1a region (Table 1; Fig. 1A, panel c). This result suggested that the presence of N instead E in the X residue of the second sequon increases the probabilities for elongation of the attached oligosaccharide. We should emphasize that, since 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 1). 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 (see 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%) (Table1; Fig. 1A, panel b), the latter being further characterized as a form carrying two short oligosaccharides by Western blotting (Fig. 2). Finally, mutant NN (both sequons are identical to the first sequon of wild type exoglucanase) behaved as its EN counterpart (Table 1, Fig. 2). Therefore, in all three constructs carrying N 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 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 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 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 ExgIa 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-eluted in HPLC, their deglycosylated products have quite different retention times.

As shown in Fig. 1A (top), following treatment with endo H, authentic form A is converted into a desglycosylated product eluting in HPLC as native ExgIb (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 ExgIb. The absence of deglycosylated form A indicates that all of the exoglucanase eluting in the ExgIa region is 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 Δochl are unable to elongate the inner core but the protein portion of susceptible substrates (alpha factor, or Exg1) are normally processed; accordingly, this mutant only secretes an Exg1b-like form with no traces of Exg1a. YS57-5A cells (Δochl) transformed with the EN construct
Hyperglycosylation signals in yeast glycoproteins

exclusively secreted Exg1b. Since the Δochl mutation does not prevent secretion of form A in a Δkex2 mutant (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 the Exg1a-like forms yielded by constructs NN, EN and 0N smeared when analyzed in SDS-PAGE, where they exhibited a similar molecular size (Fig. 1C, panel a). Obviously, the absence of the first short oligosaccharide (MW 3000) in the 0N 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 be also noted that hyperglycosylated forms derived from constructs EN, NN, and 0N elute clearly into two peaks, a property that we have extended now to wild type Exg1a where it was not as evident due to the low levels of this glycoform. Interestingly, as shown for the construct 0N (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 MP I, MP II or both.

In order 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 only present those 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 makes it indistinguishable from the second intermediate detected during the deglycosylation of Exg1b (17). The most likely explanation is that it 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 oligosacharide 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 0N transformant yielded similar results, except that there were no traces of the small intermediate (Fig. 4b). 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

In order 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 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. In order 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 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 hyperglycosylated 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 N165 and N325 (28). The protein monomer shows (α/β)₈ barrel conformation characteristic of family 5 glycoside hydrolases (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 N325, while an acid residue has the opposite effect. Figure 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 N325 must be critical to allow the action of the glycosylating enzyme machinery. Fig. 8 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 E326 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 E326 by a basic residue (E326H, E326K and E326R) allows an alternative positioning of N325 (Fig. 5 C-E) that directs the growth of the glycosidic chain towards 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 probabilities (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-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 of Exg1, since the second sequon (NES) was 100% core glycosylated regardless if 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. Since both systems appear to be only distantly
related, this coincidence suggest that NWS/T sites are not efficiently recognized by OST, 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 NWS/T sequons also modulate their degree of occupancy.

On the other hand, our results indicate that 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 three-fold the hyperglycosylation frequency of the corresponding core oligosaccharide as compared with the original sequon (NES). This indicates that (i) the environment of a sequon in the folded protein is crucial for its hyperglycosylation potential; and (ii) once this parameter is fixed, the nature of the amino acid X modulates significantly 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) also supported poorly the extension of the inner core. Curiously, the same situation is found in sequons 2 (Asn45-Asp-Thr), 3 (Asn78-Asp-Ser) and 6 (Asn99-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-Thr) is hyperglycosylated (Table 2). However, this sequon is so close to the amino terminus of the protein and its local environment has such a high hydropatic index (+1.5) (15) that other constrains 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 (Asn₁₆₅-Ser-Thr) or 13 (Asn₃₇₉-Thr-Thr) of invertase, which never elongate (Table 2). Clearly, in addition to the nature of the amino acid X, other constrains of different nature, probably related to the tertiary structure or to the charge of surrounding regions, avoids 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 \textit{Pichia sp}, do not elongate extensively the N-oligosaccharides which contain 8-14 mannoses and even heterologous expression of the \textit{S. cerevisiae} \textit{SUC2} gene in \textit{P. pastoris} or \textit{P. angusta} results in hypoglycosylation of otherwise hyperglycosylated oligosaccharides (39, 40). The most likely conclusion is that \textit{Pichia sp} lacks the outer chain synthesizing machinery described in \textit{S. cerevisiae} and, accordingly their glycoproteins should have evolved under different constraints. Studies with the invertase of \textit{P. anomala} encoded by the \textit{INV1} gene have indicated that this protein contains ten 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 \textit{S. cerevisiae} (sequons 4, 5, 6, 7, 10 and 12 respectively), whereas the rest have not a clear counterpart in the latter (39)(Table 2). Expression of the \textit{P. anomala INV1} gene in \textit{S. cerevisiae} yields an invertase carrying all the potential glycosylation sites (10 of them) occupied, but whose oligosaccharides are even shorter than their counterparts from the autologous enzyme, an indication that they are not a substrate for the mannosylpolymerases that elongate the inner core in autologous proteins (40). Since 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 out of the six fully or partially conserved sequons of the \textit{S. cerevisiae} enzyme (sequons 4, 6, 10 and 12) contain short-chain oligosaccharides (although only one -6- carries an acidic amino acid -Asp- in the X position) and a fifth one (sequon 5, Asn₉₃) is non-glycosylated. In addition, three out of the four hyperglycosylated sequons in Suc2p (1, 9, 11) are not conserved in \textit{P. anomala} and the partially conserved one (\textit{ScAsn₁₄₆-Ser-Thr} versus \textit{PaAsn₁₆₅-Ser-Ser}) is glycosylated with a frequency lower than 50% (40)(Table 2). 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 than bulky aromatic amino acids at the X position, such as Trp and Tyr, also favor hyperglycosylation.

An important question arising from our experimental approach is to identify the reason why the substitution of an acidic amino acid by a neutral one or viceversa 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 mannosylpolymerases that elongate the outer chain. Several possibilities may account for this situation. First, changes in the local environment that place the oligosaccharide in a more favorable position; second, the presence of a negative charge near the inner core may impair accessibility of the elongating enzymes; and third, the rate of intracellular transport of the new protein is slowed down because the structural alteration and accordingly it has more time to interact with the Golgi mannosylpolymerases.

Our results clearly support the first possibility, but do not eliminate any of the other two. How the introduction of a negatively charged amino acid at the X position can modify the orientation of the inner core oligosaccharide? There are important differences between core glycosylation and hyperglycosylation that need to be taken into consideration to understand how the final shape of a secreted glycoprotein is achieved. Core glycosylation takes place in the ER at a moment when the nascent protein is still unfolded. Thus, it is not surprising that a statistical analysis of N-glycosylation sites present in more than 500 proteins of known structure reveals that a large number of glycosylated asparagines residues are located at sites of low accessibility (37). In contrast, hyperglycosylation takes place in the Golgi once the polypeptide chain has been fully synthesized and folded. At this moment, the elongation of the mannan chains is subjected to steric constrains imposed by the orientation of the inner core, which in turns depends on the position and orientation of the asparagine side chain. It is obvious that the orientation adopted by N325 is determined by the atomic interaction that take place in the surroundings of this aminoacid and in particular by the nature of the amino acid in position 326 (see Figs. 4 and 5). Crystal structure of wild type Exg1 glycosylated at N165 and N325 (29) shows that the glycan at N325 lays tangentially to the globular protein monomer. Further extension of the sugar chain
Hyperglycosylation signals in yeast glycoproteins seems to be restricted by its proximity to the protein surface. Changes in the orientation of N325, particularly those proposed in Fig. 5, are expected to displace the glycan, thus favoring hyperglycosylation. When oriented towards the exterior, the glycan core, which is firstly 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 (N165) 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 $\alpha$-1,6 mannose added by Och1p interacts with Mnn9p and, depending on the nature of this interaction Mnn9 would add either an $\alpha$-1,6 or $\alpha$-1,2 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.

ACKNOWLEDGEMENTS

We thank Dr. Frank Maley for his generous gift of Endo H. We also thank Dr. Luis Rodriguez for sharing with us the results presented in Table 2.

REFERENCES

1. Orlean, P. (1997) The molecular biology of the yeast Saccharomyces, vol. 3. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 229-362
2. Burda, P., and Aebi, M. (1999) Biochim Biophys Acta 1426, 239-257
3. Ballou C.E. (1990) Methods Enzymol. 185, 440-70
4. Rayner, J. C. y Munro, S. (1998) J. Biol. Chem. 273, 26836-26843
5. Jungmann, J., and Munro, S. (1998). EMBO J, 17, 423-434
6. Jungmann, J., Rayner, J.C., and Munro, S. (1999) J. Biol. Chem. 271, 6579-6585
7. Stolz, J., and Munro, S. (2002) J Biol Chem. 277, 44801-44808
8. Marshall, R.D. (1974) Biochem Soc Symp. 40, 17-26
9. Bause, E., and Lehle, L. (1979) Eur J Biochem. 101, 531-40.
10. Bause, E. (1983) Biochem. J. 209, 331-336
11. Kaplan, H.A., Welply, J.K., and Lennarz, W.J. (1987) Biochim Biophys Acta 906, 161-173
12. Roitsch T., and Lehle, L. (1989) Eur J Biochem. 181, 525-529
13. Gavel, Y., and von Heijne, G. (1990) Protein Eng. 3, 433-442
14. Kasturi, L., Eshleman, J.R., Wunner, W.H., and Shakin-Eshleman, S.H. 1995. J Biol Chem. 270, 14756-14761
15. Reddy, V.A., Johnson, R.S., Biemann, K., Williams, R.S., Ziegler, F.D., Trimble, R.B. and Maley,F. (1988) J. Biol. Chem. 263, 6978-6985
16. Reddy, A., Gibbs, B.S., Liu, Y.L., Coward, J.K., Changchien, L.M., Maley, F. (1999) Glycobiology 9, 547-55
17. Basco, R. D., Hernández, L. M., Muñoz, M. D., Vázquez de Aldana, C., and Larriba, G. (1993) Yeast 9, 221-234
18. Shakin-Eshleman, S.H., Spitalnik, S.L., and Kasturi, L. (1996) J Biol Chem. 271, 6363-6366
Hyperglycosylation signals in yeast glycoproteins

19. Kasturi, L., Chen, H., and Shakin-Eshleman, S.H. (1997). Biochem. J. 323, 415-419
20. Johnson, L.M., Bankaitis, V.A., and Emr S.D. (1987) Cell 48, 875-85
21. Klionsky, D.J., Banta, L.M., and Emr, S.D. (1988) Mol. Cell. Biol. 8, 2105-2116
22. Henriissat, B., and Davies, G. (1997) Curr Opin Struct Biol. 7, 637-644
23. Hernandez, L.M., Olivero, I., Alvarado, E., and Larriba, G. 1992. Oligosaccharide structures of the major exoglucanase secreted by Saccharomyces cerevisiae. Biochemistry 31, 9823-9831
24. Vázquez de Aldana, C.R., Correa, J., San Segundo, P., Bueno, A., Nebreda, A.R., Méndez, E., and del Rey F. (1991) Gene 97, 173-182
25. Basco, R. D., Hernández, L. M., Muñoz, M. D., Olivero, I., Andaluz, E., del Rey, F. and Larriba, G. (1994). Biochem. J. 304, 917-922
26. Hernandez, L.M., Ramírez, M., Olivero, I., and Larriba, G. (1986) Arch. Microbiol. 146, 221-226
27. Sikorski, R.S. and Hieter, P. (1989) Genetics 122, 19-27
28. Cueva, R., Munoz, M.D., Andaluz, E., Basco, R.D., and Larriba, G. (1996) Biochim Biophys Acta 1289, 336-342
29. Taylor, S.C., Ferguson, A.D., Bergeron, J.J., and Thomas, D.Y. (2004) Nat. Struct. Mol. Biol. 11, 128-134
30. Guex, N., and Peitsch, M. C. (1997) Electrophoresis 18, 2714-2723
31. Guex, N., Diemand, A., and Peitsch M. C. (1999) Trends Biochem. Sci. 24, 364-367
32. Ramírez, M., Hernández, L.M., and Larriba, G. (1989) Arch Microbiol 151, 391-398
33. Basco, R. D., Giménez-Gallego, G., and Larriba, G. (1990) FEBS Lett. 268, 99-102
34. Basco R.D., Cueva, R., Andaluz, E., and Larriba, G. (1996) Biochim Biophys Acta 1310, 110-118.
35. Sakon, J, Adney, W.S., Himmel, M.E., Thomas, S.R., and Karplus, P.A. (1996) Biochemistry 35, 10648-60
36. Cutfield, S. M., Davies, G. J., Murshudov, G., Anderson, B. F., Moody, P. C. E., Sullivan, P. A., and Cutfield, J. F. (1999) J. Mol. Biol. 294: 771-783
37. Petrescu, A.J., Milac, A.L., Petrescu, S.M., Dwek, R.A., and Wormald, M.R. (2004) Glycobiology 14:103-104
38. Ziegler, F.D., Maley, F., and Trimble, R.B. (1988) J Biol Chem. 263, 6986-92
39. Perez, J.A., Rodriguez J, Rodriguez L, Ruiz T. (1996) Curr Genet. 29, 234-40
40. Perez, J.A., Rodriguez, J., Ruiz, T., and Rodriguez, L. (2001) Arch Microbiol. 175:189-9
LEGENDS TO FIGURES

Scheme 1. Biosynthesis of N-linked oligosaccharides in *S. cerevisiae*. Upper branch, mannan-type structures. Lower branch, core-type structures. GlcNAc: N-acetylglucosamine; Man: mannose; P: phosphate.

Fig. 1. Top: elution times of the indicated glycoforms of Exg1p. Exo\textsubscript{TUNI}: Exg secreted in the presence of tunicamycin. (A) HPLC profiles of the exoglucanase activity produced by constructs NE (wild type)(a), NN (b), and 0N (c). (B) Western blot analysis of Exg1b, Exg1b-like, underglycosylated or non-glycosylated forms produced by the indicated constructs: 1, EE (wt); 2, NN; 3, EN; 4, E0; 5, 0N; 6, NE(wt); and 7, 00. Arrows on the left side from the top to the bottom indicate the position of authentic Exg1b, Exb\textsubscript{325}, Exb\textsubscript{165} and non-glycosylated Exg. (C) a) Western blot analysis of the hyperglycosylated forms yielded by wild type (lane 1) and constructs EN (lane 2), NN (lane 3) and 0N (lane 4). Hyperglycosylated fractions (8-13 the HPLC column) secreted by the indicated mutants were pooled and analyzed by Western blots using polyclonal antibodies against the protein moiety of Exg1p b) Hyperglycosylated exoglucanase derived from construct 0N (lane 1) was fractionated in HPLC. Fractions 8-10 (lane 2) and 11-13 (lane 3) were pooled and analyzed by Western blots as described above.

Fig. 2. Time course deglycosylation of the hyperglycosylated variant produced by construct NN. a) Analysis of in HPLC. Panel A: Isolated hyperglycosylated fraction. Panels B-G: Hyperglycosylated fraction following treatment with endo H for 1.5, 3, 6, 12, 22 and 40 h respectively. b) Analysis by Western blots c) Quantification of substrates and reaction products from the HPLC eluate.

Fig. 3 (A) Relative amounts of hyperglycosylated Exg produced by the constructs carrying the indicated amino acid at the 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).

Figure 4: Environment of residue N285 (sequon 2) showing neighbouring residues located within a radius of 6 Å

Figure 5: A to E) Representation of permitted orientations of N325 (sequon 2) resulting from the free rotation of its side chain.

Fig. 6. A speculative model to explain the existence of core- and mannan-type oligosaccharides on the basis of the oligosaccharide orientation (see Stolz and Munro, 2002). Once glycoproteins received the first α-1,6 mannose by Och1p, they interact
with the M-PoI complex formed by Mnn9 and Van1p proteins. Only when the orientation is appropriate to interact with the $\alpha$-1,6 mannosyltransferase active site of Mnn9p, the oligosaccharide is extended with an additional $\alpha$-1,6 mannose, subsequent extension requiring the action of Van1p. When the orientation is not appropriate there are two possibilities. Mnn9p could add the $\alpha$-1,2 mannose that acts as a stop signal, since the purified protein has been shown to have also $\alpha$-1,2 mannosyltransferase activity or the interaction is non-productive and the oligosaccharide is released from the complex. In the last case, other $\alpha$-1,2 mannosyltransferases, such as those from the \textit{MNT1/KTR2} family, may add the terminal $\alpha$-1,2 mannose.
Figure 3

A

Hyperglycosylated Exg1 (%)

B

Relative Exg1 activity (%)
Figure 4
Figure 5 (cont.)
Figure 6

Arrival from ER (GlcNAc$_2$Man$_6$)

First mannose added by Och1p

Mannose by M-Pol 1

Alternative addition of second mannose

Extension or release

A-1,6-addition

Addition of a-1,2 mannose

Release

Non-productive interaction

Addition of a-1,2 mannose

A-1,2-addition

Kre2p? Kttrp?
| Exg1 variant | First sequon | Second sequon | Associated glycoforms                        |
|--------------|--------------|---------------|-----------------------------------------------|
| ExgNE        | NNS          | NES           | Exg1a(10%) + Exg1b                            |
| ExgEE        | NES          | NES           | Exg1a(10%) + Exg1b                            |
| ExgNN        | NNS          | NNS           | Exg1a(30%) + Exg1b                            |
| ExgEN        | NES          | NNS           | Exg1a(30%) + Exg1b                            |
| ExgE0        | NES          | QNS           | Exg125-like                                   |
| Exg0N        | QNS          | NNS           | Exg1a-like(30%) + Exg325                      |
| ExgN0        | NNS          | QES           | Exg165                                        |
| Exg0E        | QNS          | NES           | Exg1a-like(10%) + Exg325                      |
| Exg00        | QNS          | QES           | none                                          |
Table 2. Alignment of sequons from Inv1 (*P. anomala*)\(^1\) and Suc2 (*S. cerevisiae*)\(^2\)

| Sequons in PaInv1p | Sequons in ScSuc2\(^3,4\) | Features of oligosaccharides present in invertase from *S. cerevisiae* |
|-------------------|--------------------------|---------------------------------------------------------------------|
| -                 | (4)NET                   | Long (>50 Man)                                                      |
| -                 | (45)NDT                  | Short (8-14 Man)                                                    |
| 1 (112)NNT       | (92)NNT                  | Short                                                               |
| 2 (113)NTS       | (93)NTS                  | None                                                                |
| 3 (119)NDS       | (99)NDT                  | Short                                                               |
| 4 (165)NSS       | (146)NST                 | Long (frequency < 50%)                                              |
| 5 (211)NFT       |                          | -                                                                   |
| 6 (237)NDT       |                          | -                                                                   |
| -                 | (247)NGT                 | Short                                                               |
| 7 (333)NYT       |                          | -                                                                   |
| -                 | (256)NQS                 | Long                                                                |
| 8 (364)NET       | (337)NIS                 | Short                                                               |
| -                 | (350)NTT                 | Long                                                                |
| 9 (398)NAT       | (365)NST                 | Short                                                               |
| -                 | (379)NTT                 | Short                                                               |
| 10 (420)NMT      |                          | -                                                                   |
| -                 | (493)NMT                 | Short                                                               |

\(^1\)From references 37, 38. \(^2\)From references 16, 40.

\(^3\)Numbers in bold identify the order of the sequon starting from the N-terminus

\(^4\)Numbers in parenthesis refer to the position of the Asn of each sequon in the sequences of Inv1 and Suc2

Color codes: yellow, sequons carrying long oligosaccharides in the *S. cerevisiae* invertase; green, conserved sequons carrying short oligosaccharides in the *S. cerevisiae* invertase; blue, empty sequon in the *S. cerevisiae* invertase.
A search for hyperglycosylation signals in yeast glycoproteins
Raúl Conde, Rosario Cueva, Guadalupe Pablo, Julio Polaina and Germán Larriba

J. Biol. Chem. published online July 26, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M406678200

Alerts:
  • When this article is cited
  • When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts