Regulation of N-linked core glycosylation: use of a site-directed mutagenesis approach to identify Asn-Xaa-Ser/Thr sequons that are poor oligosaccharide acceptors

Lakshmi KASTURI*, Hegang CHEN† and Susan H. SHAKIN-ESHLEMAN*‡

*Institute of Pathology and †Department of Epidemiology and Biostatistics, Case Western Reserve University School of Medicine, 2085 Adelbert Road, Cleveland, OH 44106, U.S.A.

N-linked glycosylation can profoundly affect protein expression and function. N-linked glycosylation usually occurs at the sequon Asn-Xaa-Ser/Thr, where Xaa is any amino acid residue except Pro. However, many Asn-Xaa-Ser/Thr sequons are glycosylated inefficiently or not at all for reasons that are poorly understood. We have used a site-directed mutagenesis approach to examine how the Xaa and hydroxy (Ser/Thr) amino acid residues in sequons influence core-glycosylation efficiency. We recently demonstrated that certain Xaa amino acids inhibit core glycosylation of the sequon, Asn^{37}-Xaa-Ser, in rabies virus glycoprotein (RGP). Here we examine the impact of different Xaa residues on core-glycosylation efficiency when the Ser residue in this sequon is replaced with Thr. The core-glycosylation efficiencies of RGP variants with different Asn^{37}-Xaa-Ser/Thr sequons were compared by using a cell-free translation/glycosylation system. Using this approach we confirm that four Asn-Xaa-Ser sequons are poor oligosaccharide acceptors: Asn-Trp-Ser, Asn-Asp-Ser, Asn-Glu-Ser and Asn-Leu-Ser. In contrast, Asn-Xaa-Thr sequons are efficiently glycosylated, even when Xaa = Trp, Asp, Glu or Leu. A comparison of the glycosylation status of Asn-Xaa-Ser and Asn-Xaa-Thr sequons in other glycoproteins confirms that sequons with Xaa = Trp, Asp, Glu or Leu are rarely glycosylated when Ser is the hydroxy amino acid residue, and that these sequons are unlikely to serve as glycosylation sites when introduced into proteins by site-directed mutagenesis.

INTRODUCTION

N-linked glycosylation is one of the most common forms of protein processing in eukaryotes [1]. The pattern of N-linked glycosylation in glycoproteins is important because the number and position of N-linked oligosaccharides added to a protein can have profound effects on protein expression, structure and function [2–7]. N-linked glycosylation begins with the transfer of the core oligosaccharide, $Glc_3Man_9GlcNAc_2$, to an Asn residue in the protein chain by the enzyme, oligosaccharyltransferase [1,8,9]. Oligosaccharyltransferase is located in the endoplasmic reticulum membrane, with the active site of the enzyme near the endoplasmic reticulum membrane on the luminal side [9–11]. Core glycosylation usually occurs as the glycosylation site on a nascent protein enters the endoplasmic reticulum lumen [10,12–14].

The tripeptide signal or sequon, Asn-Xaa-Ser/Thr, is generally required for N-linked glycosylation [15]. It is clear, however, that other factors also control core glycosylation, because many sequons are not glycosylated [16] or are glycosylated inefficiently [3,17–22]. These factors include the local amino acid sequence (i.e. the residues near the Asn residue) and the accessibility of the sequon for glycosylation [23]. A variety of experimental approaches have been used to characterize the impact of local amino acid residues on core glycosylation. These approaches include comparisons of the amino acid sequences near glycosylated and non-glycosylated sequons (protein surveys), the use of peptides as oligosaccharide acceptors or inhibitors, and the analysis of the co-translational glycosylation of sequons in recombinant proteins [23]. Those studies reveal that Pro residues usually block glycosylation when they occur at the Xaa position or the position immediately after a sequon [16,24,25], and that Asn-Xaa-Thr sequons are generally better oligosaccharide acceptors than Asn-Xaa-Ser sequons [11,16,26–28]. Cys residues near sequons can also inhibit glycosylation if they participate in disulphide bonding [23]. Peptide studies also demonstrated an inhibitory effect of an Asp residue at the Xaa position [29,30], and some protein surveys have found that Cys, Trp [11], Asp [15] and Glu [24] residues are uncommon at the Xaa position in glycosylated sequons.

We have used rabies virus glycoprotein (RGP) as a model system to examine systematically the impact of local amino acid residues on core-glycosylation efficiency. For these studies RGP variants were generated with a single sequon at Asn^{37}. Site-directed mutagenesis was used to substitute different amino acid residues near Asn^{37} and the impact of each amino acid substitution on core glycosylation was directly quantified by expressing the variants in a cell-free translation/glycosylation system [27,31]. Recently we compared the core-glycosylation efficiencies of variants with each of the 20 common amino acids substituted at the Xaa position of the sequon, Asn^{37}-Xaa-Ser. The core-glycosylation efficiency of those variants ranged from no glycosylation to full glycosylation, demonstrating that the Xaa residue is an important determinant of core-glycosylation efficiency when Ser is present at the hydroxy (Ser/Thr) position [31]. Here we compare the impact of the Xaa residue in Asn-Xaa-Ser sequons with that of Asn-Xaa-Thr sequons to define further the protein sequences that control N-linked core glycosylation.

MATERIALS AND METHODS

Construction of plasmids encoding RGP variants with amino acid substitutions near Asn^{37}

RGP(1--)X38T39 variants with a Trp, Asp, Glu, Asn, Gly, Arg, His or Ser residue at the Xaa position were generated using a

Abbreviations used: RGP: rabies virus glycoprotein.

‡ To whom correspondence should be addressed.

cassette mutagenesis approach described in detail previously [31]. [(1--) signifies deletion of sequons 2 and 3, leaving only sequon 1.] The RGP(1--)X38T39 variant with a Leu residue at the Xaa position was constructed previously by replacing Ser^{39} in RGP(1--) with Thr [27]. Successful mutagenesis was confirmed in each variant by DNA sequencing.

Expression of RGP variants in a cell-free system

Methods for expression of RGP variants in a cell-free system have been described previously [3,27,31]. Briefly, RNA encoding RGP variants were generated by *in vitro* transcription and translated with a rabbit reticulocyte lysate system supplemented with [^{35}S]methionine and canine pancreas microsomes. After translation, the reactions were treated with proteinase K and analysed by SDS/PAGE and autoradiography [31]. Densitometric analysis was performed with gel autoradiographs exposed in the linear range.

Statistical analysis

Fisher's Exact Test was used for data analysis.

RESULTS AND DISCUSSION

Construction of plasmids encoding RGP variants with amino acid substitutions near Asn^{37}

Wild-type RGP contains three Asn-Xaa-Ser/Thr sequons at Asn^{37}, Asn^{247} and Asn^{319} [33]. In previous studies we modified RGP by site-directed mutagenesis to generate a variant with a single glycosylation sequon, $Asn^{37}Leu^{38}Ser^{39}$ [Figure 1, RGP(1--)] [3]. The sequon in RGP(1--) is glycosylated inefficiently in the cell-free system described above and in transfected Chinese hamster ovary cells [3,27,31]. We also generated RGP(1--) variants in which the Xaa residue (Leu^{38}) was replaced with each of the common 20 amino acids [Figure 1, RGP(1--)X38S39 variants] [31]. Analysis of RGP(1--)X38S39 variants in the cell-free system revealed that the Xaa residue is an important determinant of core-glycosylation efficiency when Ser is the hydroxy amino acid [31].

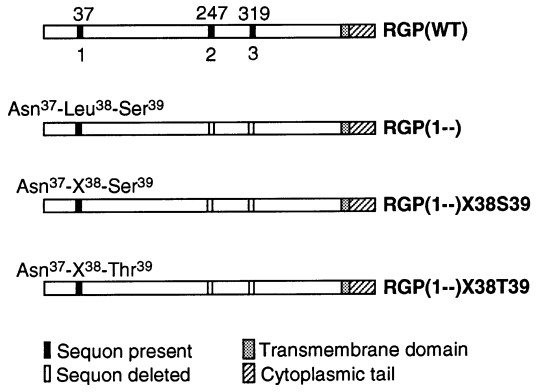

Figure 1 Structure of RGP and RGP variants

The extracellular domain of wild-type RGP contains three sequons at Asn^{37} (sequon 1), Asn^{247} (sequon 2) and Asn^{319} (sequon 3) [RGP(WT)]. RGP(1--) contains a single sequon with the sequence $Asn^{37}Leu^{38}Ser^{39}$ (sequon 1). In RGP(1--), sequons 2 and 3 were deleted by site-directed mutagenesis, replacing the Thr residue in each sequon with Ala. RGP(1--)X38S39 and RGP(1--)X38T39 variants were derived from RGP(1--) by oligonucleotide cassette mutagenesis. Each RGP(1--)X38S39 variant contains a single Asn-Xaa-Ser sequon with a different residue at the Xaa position. Each RGP(1--)X38T39 variant contains a single Asn-Xaa-Thr sequon with a different residue at the Xaa position. Abbreviation: X, Xaa.

In the present study we generated additional variants to compare the effect of the Xaa residue in Asn-Xaa-Ser versus Asn-Xaa-Thr sequons. These variants each contain a single sequon at Asn^{37} with Thr at the hydroxy position [Figure 1, RGP(1--)X38T39 variants]. RGP(1--)X38T39 variants were constructed with a Trp, Glu, Asp, Leu, Asn, Gly, Arg, His or Ser residue at the Xaa position. In a previous study, variants with these Xaa residues were shown to range from inefficient to full glycosylation when Ser is present at the hydroxy position [31].

Expression of RGP variants in a cell-free system

We compared the core-glycosylation efficiency of RGP variants with Asn^{37}-X^{38}-Ser^{39} sequons [RGP(1--)X38S39 variants] and Asn^{37}-X^{38}-Thr^{39} sequons [RGP(1--)X38T39 variants] in a rabbit reticulocyte cell-free translation system supplemented with [^{35}S]methionine and canine pancreas microsomes [34,35]. These microsomes remove the N-terminal signal sequence from RGP variants and add core oligosaccharides to Asn-Xaa-Ser/Thr sequons [3]. Processing of core oligosaccharides is limited in this system, simplifying the analysis of core glycosylation. Before glycosylation analysis, the full set of variants was translated in parallel in the absence of microsomes (without glycosylation) and analysed by SDS/PAGE. This analysis confirmed that none of the amino acid substitutions introduced near Asn^{37} altered the electrophoretic mobility of the protein (results not shown). Translation of each variant was performed in the presence of microsomes to examine core-glycosylation efficiency. In this system the core-glycosylation efficiency of the sequon, Asn-Leu-Ser in RGP, does not change significantly over a 50-fold range of RNA concentration (results not shown).

To compare the core-glycosylation efficiency of RGP variants, the full set of 18 variants (nine variants with Asn-Xaa-Ser sequons and nine variants with Asn-Xaa-Thr sequons) were translated in parallel. To ensure that RNA encoding each variant was translated in the presence of an identical amount of microsomes, the components of the *in vitro* translation system were premixed with microsomes in a single batch and then aliquots were added to tubes containing RNA. After incubation, proteins that were not targeted to microsomes and were therefore not exposed to the glycosylation apparatus were removed from the reactions by proteinase K digestion, as described [31]. Therefore the only RGP products included in the determination of core-glycosylation efficiency were those that entered the microsomal lumen. Proteinase K digestion also removes the non-glycosylated cytoplasmic tail of these proteins, which causes a shift in their electrophoretic mobility; this provides a control for the completeness of proteinase K digestion.

Because each RGP variant contains a single sequon for glycosylation, expression of the variants in this system produces proteins that are either glycosylated with a single core oligosaccharide (G) or not glycosylated (N). The glycosylated and non-glycosylated forms of each variant were separated by SDS/PAGE and analysed by autoradiography, as described [31]. Gel autoradiographs showing the translation products of variants with Asn-Xaa-Ser sequons (Figure 2A) and Asn-Xaa-Thr sequons (Figure 2B) are shown. The residue at the Xaa position of each variant is indicated above each lane, and the migration positions of the glycosylated and non-glycosylated proteins are shown.

To quantify core-glycosylation efficiency, each variant was expressed in the cell-free system in three independent experiments. The amount of protein produced in each translation reaction can vary, reflecting differences in the amount of RNA added.

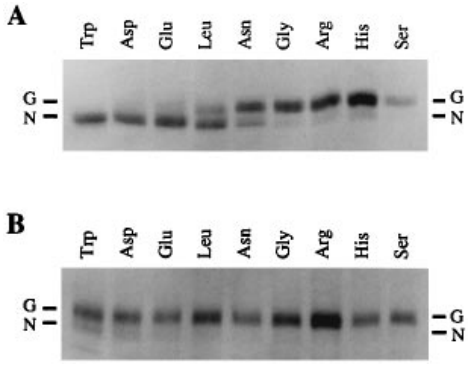

Figure 2 Core glycosylation of RGP variants

RNA encoding each RGP variant was generated by *in vitro* transcription and expressed in a cell-free translation system in the presence of canine pancreas microsomes and [^{35}S]methionine. Translation products were treated with proteinase K and analysed by SDS/PAGE and autoradiography. The residue at the Xaa position in each variant is indicated above each lane. The migration positions of the non-glycosylated protein (N) and the protein glycosylated with a single core-oligosaccharide (G) are indicated. (**A**) Translation of RGP(1--)X38S39 variants with Asn-Xaa-Ser sequons; (**B**) translation of RGP(1--)X38T39 variants with Asn-Xaa-Thr sequons.

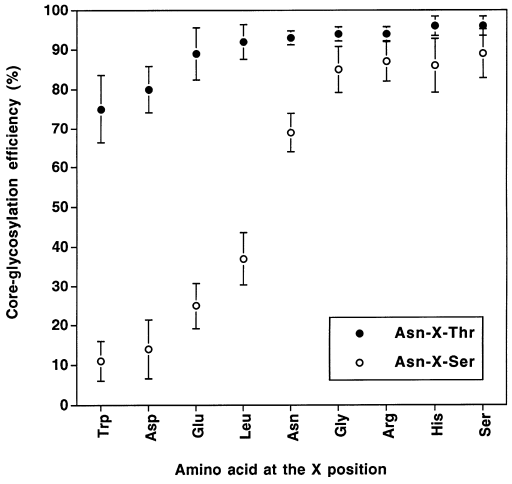

Figure 3 Influence of the Xaa residue on the core-glycosylation efficiency of Asn-Xaa-Thr sequons compared with Asn-Xaa-Ser sequons

RGP variants were analysed in the cell-free system as described in the legend to Figure 2 in three independent experiments. Gel autoradiographs from each experiment were exposed in the linear range and analysed by densitometric scanning. The core-glycosylation efficiency of each variant was calculated as described in the text for each experiment, and the mean core-glycosylation efficiency $\pm$2 S.D. from the three experiments was plotted. Data from variants with Asn-Xaa-Ser sequons (○) and Asn-Xaa-Thr sequons (●) are shown. Abbreviation: X, Xaa.

Therefore the core-glycosylation efficiency of the sequon in each variant was determined by comparing the relative amounts of glycosylated and non-glycosylated protein in each translation reaction. The amounts of glycosylated and non-glycosylated products were quantified by densitometric scanning of gel autoradiographs, and core-glycosylation efficiency was calculated as 100[G]/[G+N], as described [31,32]. For each variant, the mean core-glycosylation efficiency $\pm$2S.D. was determined for the three experiments (Figure 3). The core-glycosylation efficiency observed for each RGP variant was highly reproducible in this system. Results obtained for the RGP(1--)X38S39 variants agreed closely with those reported previously [31].

We find that the presence of a Thr residue at the hydroxy position minimizes the inhibitory effect of unfavourable Xaa residues. When Xaa = Trp, Asp, Glu or Leu, the core-glycosylation efficiency of Asn-Xaa-Ser sequons is markedly inhibited (mean core-glycosylation efficiencies 11%, 14%, 25% and 37%, respectively) (Figures 2A and 3). In contrast, the corresponding Asn-Xaa-Thr sequons are more efficiently glycosylated (mean core-glycosylation efficiencies 75%, 80%, 89% and 92%, respectively) (Figures 2B and 3). The presence of a Thr residue at the hydroxy position also enhanced the core-glycosylation efficiency of sequons with other Xaa residues. These results suggest that Asn-Xaa-Ser sequons are generally more susceptible to inhibition by Xaa residues than are Asn-Xaa-Thr sequons.

Influence of the Xaa residue on core-glycosylation efficiency in Asn-Xaa-Ser compared with Asn-Xaa-Thr sequons in other glycoproteins

To test whether our findings could be generalized to proteins other than RGP, we compared the Xaa residues at Asn-Xaa-Ser with Asn-Xaa-Thr sequons in a set of 38 well-characterized glycoproteins for which the glycosylation status at each sequon was known [16]. The proteins used to determine glycosylation status in this data set were synthesized *in vivo* or in cultured cells. In a previous study, the same sequons were analysed to identify factors that influence glycosylation [16]. That study revealed that sequons with Pro at the Xaa position or the position immediately after a sequon were rarely glycosylated, that glycosylated sequons were more likely to occur close to the N-terminus of proteins, and that Asn-Xaa-Thr sequons were more likely to be glycosylated than Asn-Xaa-Ser sequons. That study did not find an association between the presence of specific amino acid residues at the Xaa position (other than Pro) and glycosylation status when Asn-Xaa-Thr and Asn-Xaa-Ser sequons were considered as a single group [16].

On the basis of results obtained with RGP as a model protein, we postulated that Asn-Xaa-Ser sequons with Xaa = Trp, Asp, Glu or Leu would be poor oligosaccharide acceptors, whereas other Asn-Xaa-Ser/Thr sequons would be likely to be glycosylated, provided that Xaa $\neq$ Pro. To test this hypothesis we analysed the same set of proteins evaluated previously [16]. As previously, sequons in cytoplasmic and transmembrane domains of proteins were excluded from the analysis, as were sequons in highly similar regions of closely related proteins. We further excluded two non-standard sequons (Asn-Xaa-Cys, Asn-Gly-Gly-Ser), and two sequons for which the amino acid sequence was not clearly defined. This yielded a total of 148 sequons for analysis, including 100 glycosylated sequons and 48 non-glycosylated sequons. Twelve of these sequons had a Pro residue at the Xaa position. Of the remaining 136 sequons, 23 had a Trp, Asp, Glu or Leu residue at the Xaa position (Table 1).

The glycosylation status of sequons in this data set was analysed to determine whether the findings obtained for sequons in RGP could be generalized to other proteins. We first compared the glycosylation of Asn-Xaa-Ser sequons with that of Asn-Xaa-Thr sequons, with Xaa = Trp, Asp, Glu or Leu (Table 1). By assuming that these Xaa residues would inhibit glycosylation of Asn-Xaa-Ser sequons but not Asn-Xaa-Thr sequons, we correctly predicted the glycosylation status of 20 out of 23 sequons; 13 out of 15 Asn-Xaa-Thr sequons with these Xaa residues were glycosylated (predictive value 87%), compared with only 1 out of 8 of the Asn-Xaa-Ser sequons (predictive value 88%). These results demonstrate that sequons with Xaa = Trp, Asp, Glu or

Table 1 Glycosylation status of Asn-Xaa-Thr and Asn-Xaa-Ser sequons with a Trp, Asp, Glu or Leu residue in the Xaa position

Seventeen proteins in the data set contained sequons with Trp, Asp, Glu or Leu at the Xaa position. The glycosylation status of 15 Asn-Xaa-Thr and eight Asn-Xaa-Ser sequons with Xaa = Trp, Asp, Glu or Leu was compared (see text). The number of glycosylated and non-glycosylated sequons with each Xaa residue is indicated. Glycosylation was observed for 13 out of 15 of the Asn-Xaa-Thr sequons, but only for 1 out of 8 of the Asn-Xaa-Ser sequons. The proteins analysed include alkaline extracellular protease (*Yarrowia lipolytica*), α_1-antitrypsin (human), α-lactalbumin (rabbit), aspartic proteinase (*Rhizomucor miehei*), ceruloplasmin (human), cholinesterase (human serum), deoxyribonuclease (bovine pancreas), folate-binding protein (cow's milk), immunoglobulin M heavy chain (myeloma, MOPC 104E), immunoglobulin M heavy chain (human Ou, Waldenstroms), lactotransferrin (human), leucine-rich α_2-glycoprotein (human), ovalbumin (hen), pensinogen (hen), peroxidase (horseradish), ribonuclease (hippopotamus pancreas) and von Willebrand factor (human) [16].

	Asn-Xaa-Thr sequons		Asn-Xaa-Ser sequons	
Xaa	Glycosylated	Non-glycosylated	Glycosylated	Non-glycosylated
Trp	2	0	0	0
Asp	1	0	0	4
Glu	3	1	0	1
Leu	7	1	1	2
Total	13	2	1	7

Leu are more likely to be glycosylated when Thr is present at the hydroxy position ($P < 0.05$). The different glycosylation patterns of the Asn-Xaa-Thr and Asn-Xaa-Ser sequons in Table 1 was not explained by an inhibitory effect of Pro at the position after a sequon, because only one of the nine non-glycosylated sequons contained a Pro residue at that position. The proximity of the sequons to the N-terminus also did not seem to influence our findings. The position of the Asn residues in the 13 Asn-Xaa-Thr glycosylated sequons in Table 1 ranged from 17 to 481, with a mean distance from the N-terminus of 231 residues. The position of the Asn residues in the seven Asn-Xaa-Ser non-glycosylated sequons ranged from 21 to 569, with a mean distance from the N-terminus of 225 residues.

The analysis described above was extended to compare the glycosylation status of the eight sequons predicted to be poor oligosaccharide acceptors (Asn-Xaa-Ser sequons with Xaa = Trp, Asp, Glu or Leu) with that of 128 sequons in the data set with favourable sequences; the latter group included Asn-Xaa-Ser sequons with Xaa $\neq$ Trp, Asp, Glu, Leu or Pro and Asn-Xaa-Thr sequons with Xaa $\neq$ Pro. Only 1 out of 8 sequons predicted to be poor oligosaccharide acceptors was glycosylated, compared with 98 out of 128 sequons predicted to be favourable. This difference was highly significant ($P < 0.001$).

As a control for the analysis above, we compared the glycosylation status of Asn-Xaa-Ser sequons with that of Asn-Xaa-Thr sequons with Xaa residues that were predicted to be favourable in either group. First, we compared the glycosylation status of the 20 Asn-Xaa-Ser and 23 Asn-Xaa-Thr sequons in the data set with Xaa = Asn, Gly, Arg, His or Ser; both Asn-Xaa-Ser and Asn-Xaa-Thr sequons with those Xaa residues were efficiently glycosylated in RGP (Figures 2 and 3). As expected, we found no significant difference between these two groups ($P = 0.081$). We also compared the glycosylation status of the 45 Asn-Xaa-Ser and 68 Asn-Xaa-Thr sequons in the data set with any of the 15 Xaa residues predicted to be favourable with either Ser or Thr at the hydroxy position (i.e. Xaa $\neq$ Trp, Asp, Glu, Leu or Pro). As predicted, no significant difference was observed between these two groups ($P = 0.087$).

Novel Asn-Xaa-Ser/Thr sequons have been introduced into recombinant proteins by genetic engineering to study the role of glycosylation in protein folding, transport, stability and proteolytic processing [23]. This approach has also been used to study the topology of multispanning membrane proteins and to modify the biological properties of glycoproteins [23]. Unfortunately, many sequons introduced into proteins are not glycosylated. The findings in the present study suggest that Asn-Xaa-Ser sequons with Xaa = Trp, Asp, Glu or Leu should be avoided when novel sequons are engineered into proteins by site-directed mutagenesis. To address this issue we examined the glycosylation status of 38 novel sequons engineered into nine different recombinant proteins [4,5,7,36–41]. The studies cited provided information about the Xaa and hydroxy amino acid residues in each novel sequon, and determined whether or not each novel sequon was glycosylated. One of the sequons contained a Pro residue at the Xaa position, and one contained Pro at the position after the sequon; neither of those sequons was glycosylated. The 36 remaining sequons included three Asn-Xaa-Ser sequons with Xaa = Trp, Asp, Glu or Leu, and four Asn-Xaa-Thr sequons with these Xaa residues. None of the three sequons with a Ser residue at the hydroxy position was glycosylated, whereas all four sequons with Thr at the hydroxy position were glycosylated. Of the remaining 29 sequons in these proteins that were predicted to be favourable, 24 (83 %) were glycosylated. Although the data for analysis of novel sequons are limited, these findings agree with the experimental results and analysis of naturally occurring sequons in glycoproteins presented above.

Our findings are consistent with those of previous studies that suggest that Asn-Xaa-Thr sequons are generally better oligosaccharide acceptors than Asn-Xaa-Ser sequons [11,16,26–28]. The greater oligosaccharide acceptor activity of Asn-Xaa-Thr sequons has been explained by a model in which the hydroxy amino acid residue plays a direct catalytic role in the transfer of the core oligosaccharide [26,28]. In this model, core glycosylation involves a series of hydrogen bond transfer reactions between oligosaccharyltransferase, the hydroxy amino acid side chain, and the β-amide group of Asn [26,42]. The hydroxy amino acid residue (or Cys, in Asn-Xaa-Cys sequences) is proposed to relay a proton from the Asn residue to the active site of oligosaccharyltransferase, rendering the β-amide group of the Asn residue more susceptible to nucleophilic attack on the dolichol–oligosaccharide complex [26,28,42]. The basicity of the hydroxy amino acid would therefore influence the efficiency of oligosaccharide transfer, with Thr > Ser > Cys [26,28]. Other investigators suggest a related model, in which interactions between the Asn residue and hydroxy amino acid residue involve the peptide backbone as well [43,44]. The findings from the current study demonstrate that the presence of a Thr residue at the hydroxy position of a sequon exerts a dominant effect on core glycosylation, minimizing the inhibitory effect of unfavourable Xaa residues.

Our study identifies four Xaa residues as particularly unfavourable (Trp, Asp, Glu and Leu) and demonstrates that Asn-Xaa-Ser sequons with these Xaa residues are unlikely to be glycosylated. Unfavourable Xaa residues might decrease the efficiency of core glycosylation by lowering the affinity of a sequon for oligosaccharyltransferase or by impairing hydrogen bond transfer reactions required for core oligosaccharide transfer. Because the spatial relationship of the Asn and hydroxy amino acid residues in a sequon might be critical for oligosaccharide transfer [25,26,29,43–47], large, hydrophobic Xaa residues (e.g. Trp, Leu, Phe and Tyr) might inhibit core glycosylation by producing an unfavourable local protein con-

formation [31]. In contrast, small Xaa residues that do not constrain protein conformation, such as Gly, are associated with efficient core glycosylation. Our studies also reveal that negatively charged Xaa residues (Asp and Glu) inhibit glycosylation, whereas positively charged Xaa residues (Lys, Arg and His) are favourable [31]. The presence of a negatively charged amino acid residue at the Xaa position might impair the ability of oligosaccharyltransferase to bind simultaneously to the sequon and the negatively charged dolichol-*PP*-oligosaccharide precursor [29–31]. Interestingly, sequons containing hydroxy amino acid residues (Ser and Thr) or Cys at the Xaa position were found to be excellent oligosaccharide acceptors, whereas those with amide groups (Asn and Gln) at the Xaa position were glycosylated less efficiently [31]. Further characterization of the structure and enzymic mechanism of oligosaccharyltransferase will help to clarify the role that individual amino acid residues play in oligosaccharide addition.

The general relevance of our findings with RGP as a model protein is supported by our analysis of other proteins synthesized *in vivo* and in cultured cells. This information should improve predictions of the glycosylation patterns of proteins from their amino acid sequences and should facilitate the design of novel recombinant glycoproteins. Because efficient core glycosylation also requires that a sequon is accessible to oligosaccharyltransferase during protein synthesis, factors such as protein folding and the position of a sequon in a protein should also be considered in the design of recombinant proteins [23].

This work was supported in part by funds from the National Blood Foundation and the Center for AIDS Research at Case Western Reserve University (NIH # AI36219).

REFERENCES

1 Kornfeld, R. and Kornfeld, S. (1985) Annu. Rev. Biochem. **54**, 631–664
2 Shakin-Eshleman, S. H. and Spitalnik, S. L. (1993) Trends Glycosci. Glycotechnol. **5**, 355–368
3 Shakin-Eshleman, S. H., Remaley, A. T., Eshleman, J. R., Wunner, W. H. and Spitalnik, S. L. (1992) J. Biol. Chem. **267**, 10690–10698
4 Gallagher, P., Henneberry, J., Wilson, I., Sambrook, J. and Gething, M.-J. (1988) J. Cell. Biol. **107**, 2059–2073
5 Machamer, C. E. and Rose, J. K. (1988) J. Biol. Chem. **263**, 5948–5954
6 Roberts, P. C., Garten, W. and Klenk, H.-D. (1993) J. Virol. **67**, 3048–3060
7 Omura, H., Otsu, M., Yoshimori, T., Tashiro, Y. and Kikuchi, M. (1992) Eur. J. Biochem. **210**, 591–599
8 Hubbard, S. C. and Ivatt, R. J. (1981) Annu. Rev. Biochem. **50**, 555–583
9 Kelleher, D. J., Kreibich, G. and Gilmore, R. (1992) Cell **69**, 55–65
10 Nilsson, I. M. and von Heijne, G. (1993) J. Biol. Chem. **268**, 5798–5801
11 Kaplan, H. A., Welply, J. K. and Lennarz, W. J. (1987) Biochim. Biophys. Acta **906**, 161–173
12 Bergman, L. W. and Kuehl, W. M. (1977) Biochemistry **16**, 4490–4497
13 Glabe, C. G., Hanover, J. A. and Lennarz, W. J. (1980) J. Biol. Chem. **255**, 9236–9242
14 Rothman, J. E., Katz, F. N. and Lodish, H. F. (1978) Cell **15**, 1447–1454
15 Marshall, R. D. (1974) Biochem. Soc. Symp. **40**, 17–26
16 Gavel, Y. and von Heijne, G. (1990) Prot. Engin. **3**, 433–442
17 Plummer, T. H. and Hirs, C. H. W. (1964) J. Biol. Chem. **239**, 2530–2538
18 Pohl, G., Kallstrom, M., Bergsdorf, N., Wallen, P. and Jornvall, H. (1984) Biochemistry **23**, 3701–3707
19 Thim, L., Bjoern, S., Christensen, M., Nicolaisen, E. M., Lund-Hansen, T., Pedersen, A. H. and Hedner, U. (1988) Biochemistry **27**, 7785–7793
20 Stevenson, S. C., Wang, S., Deng, L. and Tall, A. R. (1993) Biochemistry **32**, 5122–5126
21 Curling, E. M. A., Hayter, P. M., Baines, A. J., Bull, A. T., Gull, K., Strange, P. G. and Jenkins, N. (1990) Biochem. J. **272**, 333–337
22 Inouye, S., Guo, Y., Ling, N. and Shimasaki, S. (1991) Biochem. Biophys. Res. Commun. **179**, 352–358
23 Shakin-Eshleman, S. H. (1996) Trends Glycosci. Glycotechnol. **8**, 115–130
24 Mononen, I. and Karjalainen, E. (1984) Biochim. Biophys. Acta **788**, 364–367
25 Bause, E. (1983) Biochem. J. **209**, 331–336
26 Bause, E. and Legler, G. (1981) Biochem. J. **195**, 639–644
27 Kasturi, L., Eshleman, J. R., Wunner, W. H. and Shakin-Eshleman, S. H. (1995) J. Biol. Chem. **270**, 14756–14761
28 Picard, V., Ersdal-Badju, E. and Bock, S. C. (1995) Biochemistry **34**, 8433–8440
29 Imperiali, B. and Shannon, K. L. (1991) Biochemistry **30**, 4374–4380
30 Bause, E. and Hettkamp, H. (1979) FEBS Lett. **108**, 341–344
31 Shakin-Eshleman, S. H., Spitalnik, S. L. and Kasturi, L. (1996) J. Biol. Chem. **271**, 6363–6366
32 Shakin-Eshleman, S. H., Wunner, W. H. and Spitalnik, S. L. (1993) Biochemistry **32**, 9465–9472
33 Wunner, W. H., Larson, J. K., Dietzschold, B. and Smith, C. L. (1988) Rev. Infect. Dis. **10** (Suppl 4), 771–784
34 Pelham, H. R. B. and Jackson, R. J. (1976) Eur. J. Biochem. **67**, 247–256
35 Scheele, G. (1983) Methods Enzymol. **96**, 94–111
36 Nakamura, S., Takasaki, H., Kobayashi, K. and Kato, A. (1993) J. Biol. Chem. **268**, 12706–12712
37 Wright, A., Tao, M., Kabat, E. A. and Morrison, S. L. (1991) EMBO J. **10**, 2717–2723
38 Iorio, R. M., Glickman, R. L. and Sheehan, J. P. (1992) J. Gen. Virol. **73**, 1167–1176
39 Guan, J.-L., Machamer, C. E. and Rose, J. K. (1985) Cell **42**, 489–496
40 Ladenheim, R. G., Seidah, N. and Rougeon, F. (1991) Eur. J. Biochem. **198**, 535–540
41 Baldwin, A. N. and Shooter, E. M. (1995) J. Biol. Chem. **270**, 4594–4602
42 Bause, E., Warmuth, R. and Wesemann, M. (1989) in Tenth International Symposium on Glycoconjugates. Symposium Proceedings, Jerusalem, Israel (Sharon, N., Lis, H., Duksin, D. and Kahane, I., eds.), pp. 38–39
43 Imperiali, B., Shannon, K. L., Unno, M. and Rickert, K. W. (1992) J. Am. Chem. Soc. **114**, 7944–7945
44 Imperiali, B., Shannon, K. L. and Rickert, K. W. (1992) J. Am. Chem. Soc. **114**, 7942–7944
45 Beeley, J. G. (1977) Biochem. Biophys. Res. Commun. **76**, 1051–1055
46 Aubert, J.-P., Helbecque, N. and Loucheux-Lefebvre, M.-H. (1981) Arch. Biochem. Biophys. **208**, 20–29
47 Bause, E., Hettkamp, H. and Legler, G. (1982) Biochem. J. **203**, 761–768

Received 7 August 1996/18 November 1996; accepted 26 November 1996