Characterization of the O-Glycosylation Sites in the Chorionic Gonadotropin β Subunit in Vivo Using Site-directed Mutagenesis and Gene Transfer

Tadashi Sugahara, Mary R. Pixley, Fuad Farest, and Irving Boime

From the Department of Molecular Biology and Pharmacology, Washington University School of Medicine, St. Louis, Missouri 63110

Human chorionic gonadotropin (CG) is a member of a family of glycoprotein hormones which are heterodimers containing two nonidentical subunits: a common \( \alpha \) and a hormone-specific \( \beta \) subunit. One of the distinguishing features of the CG\( \beta \) subunit is the presence of four serine acceptors clustered within the last 25 amino acids. We previously demonstrated that this carboxyl-terminal region is important for maintaining its biologic half-life, and when the sequence was genetically fused to either the common \( \alpha \) or follicitropin \( \beta \) subunits, O-glycosylation was observed. Because this carboxyl-terminal sequence is located at the end of the subunit, we considered this region a convenient model for studying O-linked glycosylation in domains containing multiple serine recognition sites. A CG\( \beta \) gene was engineered in which the N-linked sites were inactivated to eliminate background from those carbohydrate groups. Using this construct, we made a series of truncation and amino acid substitutions of acceptor serines, and these mutants were transfected into Chinese hamster ovary cells. O-Glycosylation was determined by \(^{3}H\) glucosamine incorporation and glycanase sensitivity of the products on SDS-polyacrylamide gels. We show that the O-linked sites comprise independent repetitive regions in which each acceptor serine has a recognition signal bounded by the next carboxy acceptor serine within four to five amino acids. It is also apparent that recognition of one site is not dependent on the glycosylation of another acceptor. Amino acid mutations in the acceptor regions demonstrated the importance of proline as a necessary feature for O-linked recognition in the CG\( \beta \) subunit sequence.

The placental hormone, human chorionic gonadotropin (hCG), and the pituitary hormone lutropin are members of the glycoprotein hormone family that also include follitropin and thyrotropin. They are heterodimers formed by a noncovalent association of a common \( \alpha \) subunit and a hormone-specific \( \beta \) subunit which determines receptor specificity (1). A unique feature of the hCG\( \beta \) subunit is the carboxyl-terminal peptide (CTP) which bears four serine-linked oligosaccharides (2–4). One potential function of the CTP is to prolong plasma half-life of hCG (5–8) since its deletion from the hCG\( \beta \) subunit by site-directed mutagenesis reduces the in vivo potency 3-fold compared to wild type hCG (6). Previously, we demonstrated that when the CTP is fused to either the carboxyl terminus of the follicle stimulating hormone \( \beta \) subunit (7) or to the common \( \alpha \) subunit (8), it was O-glycosylated. This suggested that all of the information needed for O-glycosylation is contained within the CTP sequence.

The O-linked oligosaccharides clustered in the CG\( \beta \) carboxyterminal region are composed of sialylated galactoseN-acetylgalactosamine structures (4, 9). Attachment of N-acetylgalactosamine initiates the process and, although there is controversy regarding the precise point in the secretory pathway where this occurs (10; see references therein), recent experiments using immunoelectron microscopy identified UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase in the cis-Golgi compartment (10). The addition of galactose and sialic acid presumably occurs in the trans-Golgi. While the minimum sequence requirement for N-glycosylation of asparagine residues is Asn-Xaa-Ser/Thr (where Xaa is any amino acid except proline), it is not known how the Ser/Thr sites are selected for O-glycosylation. Several investigators have examined this process using in vitro systems comprised of synthetic peptides containing the glycosylation sites and purified UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase to define the functional acceptor unit. These studies showed that both type and position of amino acids near the glycosylated serine or threonine are important for glycosylation (11–19). Although there is no clear consensus sequence analogous to the N-glycosylation sites, there are some preferences in the flanking amino acids such as proline, serine, threonine, and alanine.

Few studies have been performed in vivo (cell-culture) to examine O-linked recognition (8, 20, 21). Such investigations are important since there are multiple steps (and compartments) involved in the maturation of O-linked oligosaccharides. Based on our previous data, we considered the CG\( \beta \) subunit a convenient model for studying O-glycosylation in cells because the domain which contains the O-linked carbohydrate is located at the carboxyl terminus, and because the presence or absence of this region has no dramatic effect on the overall folding of the subunit (6, 7). Thus, mutations within the CTP sequence should create minimal perturbations in the protein. Here we address whether or not glycosylation of the four acceptor serines in the CTP is governed by a single determinant or if multiple repetitive recognition units are involved. The amino acids flanking the acceptor sites were also examined in an attempt to characterize the recognition sequence. For these studies, a construct was engineered in which the N-
linked sites were inactivated to eliminate any background from those carbohydrate groups. This vector was used as a substrate for mutagenesis, and the resulting mutants were transfected into Chinese hamster ovary cells. Here we show that the four O-linked sites are independent repeat units and that O-glycosylation of one site is not dependent on the recognition of another O-glycosylated acceptor.

**EXPERIMENTAL PROCEDURES**

Restriction enzymes used to prepare vectors were purchased from Promega or Boehringer Mannheim. The enzymes used for PCR were obtained from Stratagene (Pfu DNA polymerase). The DNA vector, M13mp19 was obtained from New England Biolabs. Oligonucleotides were prepared by the Washington University Protein and Nucleic Acid Chemistry Laboratory (St. Louis, MO). [$^35$S]Cysteine (1,000 Ci/mmol) was purchased from ICN Biochemicals (Irvine, CA). Cell culture media and reagents were prepared by the Washington University Center for Basic Cancer Research. Dialyzed calf serum, G418, fetal calf serum, and immunoprecipitant were obtained from Life Technologies, Inc. Polyclonal antiserum directed against hCG subunit was prepared in our laboratory.

**RESULTS**

**Truncated CTP Mutants**—The hCGβ subunit has two N- and four O-linked oligosaccharides, and two forms of the subunit are expressed in CHO cells; this is apparently caused by the presence of one or two N-linked oligosaccharides (24, 29). Compared to the lysate form, one of the distinguishing features of the secreted protein is a mobility shift of about 10 kDa which is primarily due to the addition of O-linked chains immediately prior to secretion. This difference in electrophoretic mobility is the basis for the assay of O-linked acceptor efficiency of the mutants. Because the heterogeneity created by the N-linked carbohydrates could complicate the analysis (see below), we constructed a mutant (CGβΔAsn (1 + 2)), which lacks the N-linked consensus sequences (24). We first made a series of truncated mutants containing stop codons at various sites in the CTP to: (a) calibrate the system for relating protein migration and the number of sites O-glycosylated (b) assess if the serine acceptor sites are independent from one another (see Table I). Three variants containing a stop codon at either residue 122 (T22), 128 (T28), and 133 (T33) were constructed (Table I). These mutants were shuttled into the eukaryotic expression vector, pM2HA (23, 24) containing the CGβ gene. To perform the analyses without interference from the N-linked carbohydrates on O-linked carbohydrate analysis, we used a mutant CGβ gene (ΔAsn (1 + 2)) which lacks Asn-X-(Thr/Ser) consensus sequences at the Asn13 and Asn30 acceptor sites (24). The entire exon II cassette of all the mutants was sequenced to ensure that the mutations were correct and that there were no misincorporations during the PCR.

**Transfection and Cell Culture**—These constructs were transfected into CHO cells as described previously (25), and stable transfectants were selected by growing in culture medium containing 0.25 μg/ml neomycin analogue G418. Expression of the mutants was detected by immunoprecipitation of metabolically labeled cells (see below). All stably transfected CHO cell lines were maintained in medium A (Ham's F12 medium supplemented with penicillin (100 units/ml), streptomycin (100 μg/ml), and glutamine (2 mM)) containing 5% (v/v) FCS and 0.125 mg/ml G418 in a humidified 5% CO2 incubator.

Transfection and Cell Culture—These constructs were transfected into CHO cells as described previously (25), and stable transfectants were selected by growing in culture medium containing 0.25 μg/ml neomycin analogue G418. Expression of the mutants was detected by immunoprecipitation of metabolically labeled cells (see below). All stably transfected CHO cell lines were maintained in medium A (Ham's F12 medium supplemented with penicillin (100 units/ml), streptomycin (100 μg/ml), and glutamine (2 mM)) containing 5% (v/v) FCS and 0.125 mg/ml G418 in a humidified 5% CO2 incubator.

**Metabolic Labeling and Protein Analysis**—Cells were plated into 12-well dishes (300,000–350,000 cells/well) in 1 ml of medium A supplemented with 5% FCS 1 day before labeling. For continuous labeling experiments, cells were washed twice with medium B (cysteine-free medium A supplemented with 5% dialyzed calf serum) and labeled for 6 h in 1 ml of this medium B containing 20 μCi/ml [$^35$S]cysteine. Media and cell lysates were prepared and treated as described (26). All immunoprecipitates were resolved on 15% SDS-polyacrylamide gels by the method of Laemmli (27). Gels were soaked for 10 min in 1 M sodium salicylate, dried, and autoradiographed. Cells were also metabolically labeled with [3H]glucosamine, and labeled GalNAc was determined by HPLC on acid hydrolyzed protein using an Aminex Bio-Rad column according to Green et al. (28).

**RESULTS**

**Truncated CTP Mutants**—The hCGβ subunit has two N- and four O-linked oligosaccharides, and two forms of the subunit are expressed in CHO cells; this is apparently caused by the presence of one or two N-linked oligosaccharides (24, 29). Compared to the lysate form, one of the distinguishing features of the secreted protein is a mobility shift of about 10 kDa which is primarily due to the addition of O-linked chains immediately prior to secretion. This difference in electrophoretic mobility is the basis for the assay of O-linked acceptor efficiency of the mutants. Because the heterogeneity created by the N-linked carbohydrates could complicate the analysis (see below), we constructed a mutant (CGβΔAsn (1 + 2)), which lacks the N-linked consensus sequences (24). We first made a series of truncated mutants containing stop codons at various sites in the CTP to: (a) calibrate the system for relating protein migration and the number of sites O-glycosylated (b) assess if the serine acceptor sites are independent from one another (see Table I). Three variants containing a stop codon at either residue 122 (T22), 128 (T28), and 133 (T33) were constructed (Table I). If a variant is not O-glycosylated, the electrophoretic mobility of lysate and media should be the same. The
Asn mutant containing the complete CTP unit undergoes the lysate/medium change as discussed above (Fig. 1, lanes 3 and 4). There is a progressive decrease in mobility as the amount of deleted sequence is increased; however, it is evident that the loss of a significant portion of the CTP sequence as in the T33 mutant does not eliminate all of the acceptor activity in the remaining serine sites (lanes 5 and 6). The secreted form of the T22 mutant exhibits no apparent molecular weight change (lanes 9 and 10) even though it contains the normal O-glycosylation site at Ser121, whereas the T28 mutant which contains an additional six amino acids including the Ser127 acceptor exhibits a lysate/medium modification (lanes 7 and 8); thus, at least one site appears O-glycosylated in T28 (see below). The mutant T33 has three potential serine O-glycosylation sites and manifests a molecular weight change greater than that seen in mutant T28. The increase in the apparent mass corresponds to an additional O-linked carbohydrate unit. These results suggest that each acceptor site in the CTP comprises an independent recognition unit.

To address whether absence of the N-linked carbohydrate affec-
ted recognition, the corresponding CTP mutants containing the two intact Asn-acceptor sites expressed in CHO cells were treated with endoglycosidase F (Fig. 1B). Because this enzymatic activity partially releases the N-linked oligosaccharides, not O-linked carbohydrates, the digested proteins should be equivalent to the (∆Asn (1 + 2)) mutants. In the absence of Endo F, both the lysate and media proteins of the wild type and the CTP mutants display multiple forms. The lysate/medium profiles of the treated samples with Endo F are identical to that seen with ∆Asn (1 + 2) variants (Fig. 1A). As expected, the digestion product of the intra- and extracellular forms of the T22 mutant were identical, indicating no modification occurred (data not shown). These data imply that the lack of the N-linked oligosaccharides do not affect acceptor recognition in these mutants.

Characterization of the CTP Modification—The above data show that the lysate/media modifications are confined between the Ser121 and carboxyl terminus 145. To show that the changes in the mobility are due to O-linked glycosylation, samples of media from the truncated mutants (T22, T28, and T33) were treated with O-glycanase (Genzyme) (Fig. 2). Because sialic acid substitution on Gal–GalNAc blocks glycanase hydrolysis, samples of T33 and T28 were first treated with neuraminidase which reduces the size of the truncated and wild type forms (panel B, lanes 3, 7, and 11). Subsequent addition of O-glycanase further decreases the mass to the corresponding lysate form (lanes 4, 8, and 12). As expected, the enzymes had no effect on the migration of the unmodified T22 mutant (panel B, lanes 13–16).

Further evidence for the presence of O-linked carbohydrate in these truncated mutants is the incorporation of the carbohydrate precursor [3H]glucosamine (Fig. 3). Cells were labeled overnight in media containing 50 μCi/ml of [3H]glucosamine. Because the N-linked sites are mutated, incorporated label will reflect only the presence of O-linked carbohydrate. The secreted forms of ∆Asn, T33, T28, but not T22, incorporate 3H (lanes 3–6), and they co-migrate with the corresponding 35S-labeled proteins (lane 2). Moreover, when the O-linked oligosaccharides were acid-hydrolyzed from the ∆Asn (1 + 2) mutant and analyzed by HPLC (28), two major peaks corresponding to GalNAc and GluNaC in a 15:1 ratio were observed. These data are in agreement with the predominance of GalNAc, not GluNaC in the native CGBP O-linked carbohydrates. Taken together, the data show that the changes in mobility of the secreted mutants are due to the presence of O-linked carbohydrate.

Point Mutations in CTP and Characterizations of a Single Acceptor Site—While data with the T28 mutant appeared to define a single serine recognition site, we could not exclude that the three tandem serine residues at 118–120 and the terminal Ser127 were also alternative acceptors. To address this issue, a
modified T28 mutant was constructed in which serines 118–120 were converted to alanine (S3T28) (Fig. 4A, lanes 7 and 8). Despite this change, the lysate/medium shift characteristic of T28 (lanes 11 and 12) is still observed. Further, when a variant was constructed in which only the acceptor serine 121 was mutated to alanine (ΔSer1 + T28) (Fig. 4B, lanes 5 and 6), the acceptor activity was abolished. These data show that in the T28 mutant only Ser121 is the acceptor and that the terminal Ser127 is not glycosylated. It is also evident that the entire complement of serine acceptor sites are glycosylated in the S3 variant of CGβ ΔAsn(1 + 2), i.e. containing the complete CTP sequence (Fig. 4A; lanes 5 and 6). Thus, serine residues 118–
Newly synthesized glycoproteins undergo numerous modifications in the secretory pathway; for O-glycosylation, several compartments are involved. In many cases, the O-linked sites

120 on the amino side of Ser\(^{121}\) are not required for recognition at Ser\(^{121}\) or the other acceptor sites. Moreover, a ΔSer1 mutant does not affect glycosylation of the other serines (data not shown) which implies O-linked acceptor activity of the downstream sites is not dependent on proximal recognition of Ser\(^{121}\). It appears then that Ser\(^{120}\) defines the amino-terminal border of the CTP.

Having identified at least one recognition site in the T28 mutant, we created a series of stop and amino acid substitution mutants between the T22 and T28 mutants to further define the structural requirement for the acceptor activity of Ser\(^{121}\) (Fig. 5A). A mutant truncated at amino acid 123 (T24) was not glycosylated (Fig. 5B, lanes 11 and 12); the presence of the additional Lys-Ala was not sufficient for recognition. However, if a variant included the adjacent Pro\(^{124}\) (T25), some glycosylation was detected, although it was incomplete as evident by the presence of the unglycosylated form (lanes 9 and 10). Mutants which result in extending the protein from Pro\(^{125}\) to Ser\(^{127}\) become more efficiently glycosylated; the intensity of the upper band increases at the expense of the lower band (lanes 4, 6, 8, and 10). Adding the next serine acceptor site, i.e. Ser\(^{127}\) results in quantitative O-glycosylation of Ser\(^{121}\). Even though the T28 mutant has the Ser\(^{127}\) glycosylation site, the apparent molecular weight of this mutant is comparable to that observed for the truncated mutants T25-T27. Thus, while Ser\(^{127}\) in the T28 mutant is not glycosylated, it is required for maximal acceptor activity of Ser\(^{121}\), which implies that Ser\(^{127}\) defines the carboxyl border of the Ser\(^{121}\) recognition sequence.

Although acceptor activity of Ser\(^{121}\) is dependent on the intervening sequence between 121 and 127, it is not clear if the nature of the amino acid, i.e. proline and/or the length of the sequence is critical. To address this question, the Pro in the partial acceptor mutant T25 was changed to several different amino acids (Fig. 5C). If length is the sole determinant, then acceptor activity should remain unaffected by the amino acid changes, whereas if the proline at 124 is crucial, changing this residue should affect O-glycosylation. No detectable acceptor activity was seen when the Pro was mutated to Arg (lanes 3 and 4) or Asp (not shown). The glycosylation efficiency of variants containing either Leu, Ala, or Gin was 30–50% of that seen in T25, and extending the sequence from 122 to 127 with only Ala replacements did not increase the extent of glycosylation over that seen with T25 Ala mutant (Fig. 5D). These data show that on the carboxyl side of the Ser\(^{121}\) acceptor site, the type of amino acid is an essential feature for maximal glycosylation efficiency and that increasing the peptide length is insufficient to overcome the lack of the Pro residues.

We examined the sequence requirements for the second acceptor site at Ser\(^{127}\). The migration of T33 stop mutant corresponds to a CTP derivative containing two O-linked sites. This indicates that the sequence Ser\(^{127}\) to Ser\(^{132}\) encompasses an additional recognition unit (Fig. 6). If the T33 mutant is truncated at Ser\(^{130}\) (T31), two bands are detected (lane 6); the mobility shift of the lower band is the same as that of the T28 stop mutant, and that of the upper band is comparable to the unmodified T33 mutant. No nonglycosylated form of either of these mutants is observed. Thus, Ser\(^{127}\) in the T31 mutant is a weak acceptor in the absence of the Pro\(^{131}\) and Ser\(^{132}\) residues. Moreover, Ser\(^{132}\) of the T33 mutant is not an acceptor because the mobility of the secreted T33 mutant is the same as that of the upper band of the T31 mutant devoid of Ser\(^{132}\). Although we cannot exclude the possibility that Ser\(^{130}\) in the T31 mutant is an alternative acceptor, based on our data, its carboxyl-terminal location makes this unlikely. These results further support the conclusion that the recognition signals for the multiple serine acceptors in the CTP is defined on the carboxyl side of each acceptor by the adjacent serine glycosylation site.

**DISCUSSION**

FIG. 4. Identification of the glycosylation site in the mutant T28. A, alanine mutations in the CTP including serines 118–120 (S3) or serine 120 (S20) of the full-length ΔAsn construct (lanes 1–6) or in the T28 truncated form (lanes 7–12). Lanes 1, 2 and 11, 12 show the migration of the nonmutated full-length CTP and T28 forms, respectively. In B, the serine at 121 in T28 was converted to alanine (ΔSer1; lanes 5 and 6). The migration of the control ΔAsn (lanes 1 and 2) and T28 (lanes 3 and 4) are shown.
are clustered, and it is not clear if recognition of multiple sites involves overlapping sequences or if the acceptors are each defined by discrete signals. Here we used the CTP sequence in the CGβ subunit as a model to address this question since it contains four closely spaced O-linked sites at the carboxyl end of the protein. Truncated mutants terminating at each Ser acceptor were constructed and expressed in CHO cells. The data show that the CTP region is not a single O-glycosylation signal, but it is a domain composed of multiple recognition units bounded by a Ser-to-Ser motif. A question that arises is whether or not linkage of the initial GalNAc occurs progressively, i.e., in an ordered sequence or randomly. Our data favor the latter because, when Ser121 was changed to Ala, glycosylation of the remaining sites in the cluster was not impaired. In

**Fig. 5. Modification of acceptor activity of Ser121.** A describes a series of truncated mutants (see also Table I) where stop codons were created at the following amino acids: 127 (T27), 126 (T26), 125 (T25), 124 (T24), and 122 (T22). The autoradiograph in B shows the corresponding lysate (L) and medium (M) forms of these mutants. C shows that proline is necessary for maximal O-glycosylation of Ser121. The carboxyl-terminal proline in T25 (lanes 1 and 2) was converted to Arg (lanes 3 and 4), Ala (lanes 5 and 6), Leu (lanes 7 and 8), or Gln (lanes 9 and 10). D describes a series of alanine (Ala) substitutions on the T25 template. The construct containing an Ala substitution for Pro124 in T25 is designated (ala)1, i.e.

-121122 123124
Ser-Lys-Ala-Ala

The mutants (ala)1 to (ala)4 refer to increasing chain length to 127 amino acids. The mutant T28 which contains (ala)4 terminates with Ala instead of the acceptor Ser at 127.
addition, the independent behavior of the individual sites in the truncated mutants appears to preclude a coupled mechanism for sequential O-linked addition.

We attempted to characterize a recognition sequence required for Ser-linked glycosylation in the CTP. There have been numerous investigations using in vitro systems containing oligopeptides and purified GalNAc transferase to study this problem (13–19). Efforts to identify a consensus sequence using such systems has been difficult because of the absence of most secondary and tertiary structure in the peptide substrates and the presence of multiple GalNAc transferases (30, 31). It is of interest that a GalNAc transferase was recently characterized using a CGβ-derived peptide containing the Ser121, Ser127, and Ser132 acceptor sites (30). While our data do not assess if one or more enzymes are responsible for the observations seen here, the homogeneous changes observed favor a single GalNAc transferase. Because of the relative discrete changes observed, we could test certain predictions concerning the recognition sequences based on the previously reported in vitro observations. It was proposed that accessibility to the serine acceptor site is a key factor for recognition by the GalNAc transferase (32). This could be a major consideration for mucins where the density of O-linked acceptor is high, but the CTP, with its proline-induced β turns, is fully exposed to solvent, and thus to the glycosyltransferases. Moreover, accessibility cannot be the exclusive determinant because in the mutant truncated at residue 124, glycosylation at Ser121 required the Pro124. Substitution of Ala, Gin, or Leu at residue 124 resulted in marked reduction of acceptor efficiency, and mutating Pro to Arg or Asp eliminated acceptor activity. This result shows that a proline residue within at least three amino acids carboxy to Ser121 is required for recognition. These results are consistent with studies of the amino acid sequences surrounding the known O-glycosylation sites (derived from the National Biomedical Research Foundation Protein Data Base) (11, 12) which indicated the need for one or more proline residues within five amino acids carboxy to the acceptor Thr/Ser. Moreover, the acceptor efficiency of a T27 mutant, in which the proline residues were all replaced with Ala (Ser-Lys-Ala-Ala-Ala-Ala-Ala) was comparable to the T25 mutant containing the Pro124 → Ala change. These data show that in addition to accessibility, the function of the Pro residue is critical, and that chain length cannot be a sole determinant. Moreover, that replacement of the proline with a charged residue eliminated acceptor activity is consistent with the recent transfection studies of Nehrke et al. (21) using a unique reporter construct comprised of sequences encoding a single O-glycosylation site and epitopes that facilitate isolation of the glycosylated protein.

Data were presented that charged residues in the vicinity of the O-glycosylation site significantly depressed glycosylation efficiency when the reporter is expressed in COS-7 or MCF-7 cells. Thus, depending on the position in the sequence, charged amino acids may quench alternative Ser/Thr sites in the CTP, e.g. Ser119/120/130 and Thr140.

Although the CTP is readily accessible to the glycosyltransferases, not all of the Ser/Thr residues between 118 and 145 are glycosylated, e.g. the adjacent serines 118–120 at the amino-terminal side of the Ser121 acceptor site. Moreover, neither of these residues are acceptors, even when Ser121 in T28 was eliminated. Although Ser/Thr residues adjacent to the normal acceptor sites may be less efficiently glycosylated in vitro (13, 18, 19), the absence of acceptor function by Ser118–120 suggests that this region is the base of a β turn and is not as readily exposed to the glycosyltransferase(s). Given the importance of the 26–110 disulfide bond in the CGβ subunit for assembly with the α subunit (33), Ser118–120 may represent a spacer for the major structure and the acceptor region of Ser121 in the O-linked rich domain. Thus, the folding induced by this disulfide bond could be a factor in defining the amino end of the CTP region. This is consistent with recent mutagenesis studies of erythropoietin which show that the conformation of the protein affects O-linked acceptor efficiency (20). Our data suggest that the GalNAc transferase is very precise for a particular site. This implies that the transfer of carbohydrate to the native acceptor suppresses the recognition at other local sites, or the presence of GalNAc at a Ser abolishes its function as a recognition sequence for an adjacent acceptor Ser.

Transfected cells expressing CGβ provide a feasible system for studying structure-function of site-specific O-glycosylation. One major issue is whether the sequences required for the initial step in O-glycosylation, i.e. GalNAc addition, also encode information for the maturation of the growing oligosaccharide
chain. This model could address the influence of the flanking protein sequences and the effect of cell type on the oligosaccharide structure by inserting the CTP sequence at defined regions in a variety of proteins. Our cellular approach, together with mass spectroscopy of purified protein and oligosaccharides, will further address the specificity of recognition both for Thr and Ser O-glycosylation as well as the enzymatic mechanisms involved.

Acknowledgments—We thank Dr. Jacques U. Baenziger and Dr. David Ben-Menahem for critical review of the manuscript and Susan Carnes for her assistance in its preparation. We are grateful to Drs. Susan Daniels-McQueen and Jacques U. Baenziger for their expertise in the GalNAc analysis.

REFERENCES
1. Pierce, J. G., and Parsons, T. F. (1981) Annu. Rev. Biochem. 50, 465–495
2. Birken, S., and Canfield, R. E. (1977) J. Biol. Chem. 252, 5386–5392
3. Keutmann, H. T., and Williams, R. M. (1977) J. Biol. Chem. 252, 5393–5397
4. Kesler, M. J., Mise, T., Ghai, R. D., and Bahl, O. P. (1979) J. Biol. Chem. 254, 7909–7914
5. Kalyan, N. K., and Bahl, O. P. (1983) J. Biol. Chem. 258, 67–74
6. Matzuk, M. M., Hsueh, A. J. W., LaPolt, P., Tsafirri, A., Keene, J. L., and Boime, I. (1990) Endocrinology 126, 376–383
7. Fares, F. A., Sugahara, T., Hsueh, A. J. W., and Boime, I. (1995) Mol. Endocrinol. 9, 54–63
8. Corless, C. L., Matzuk, M. M., Ramabhadran, T. V., Krichevsky, A., and Boime, I. (1987) J. Biol. Chem. 262, 11713–11718
9. White, T., Bennett, E. P., Takio, K., Sørensen, T., Boime, I., and Clausen, H. (1995) J. Biol. Chem. 270, 24166–24173
10. Hill, H. D., Jr., Schwzyer, M., Steinman, H. M., and Hill, R. L. (1977) J. Biol. Chem. 252, 3799–3804
11. Suganuma, N., Matzuk, M., and Boime, I. (1989) J. Biol. Chem. 264, 19302–19307