PERGAMON

Journal of Steroid Biochemistry & Molecular Biology 75 (2000) 147–158

The Journal of Steroid Biochemistry & Molecular Biology

www.elsevier.com/locate/jsbmb

Glycosylation of the murine estrogen receptor-α

Xiaogang Cheng [a,b], Gerald W. Hart [a,*]

[a] *Department of Biological Chemistry, School of Medicine, Johns Hopkins University, 725 N Wolfe St., Baltimore, MD 21205-2185, USA*
[b] *Graduate Program of Department of Biochemistry and Molecular Genetics, The University of Alabama at Birmingham, Birmingham, AL 35294, USA*

Received 22 May 2000; accepted 15 September 2000

Abstract

O-linked *N*-acetylglucosamine (*O*-GlcNAc) is a highly dynamic and abundant modification found on nuclear and cytoplasmic proteins of nearly all eukaryotes. *O*-GlcNAc addition is required for life at the single cell level and is analogous to protein phosphorylation in most respects. In a previous study (M.S. Jiang, G.W. Hart, A subpopulation of estrogen receptors are modified by *O*-linked *N*-acetylglucosamine. J. Biol. Chem. 270 (1997) 2421–2428), we demonstrated that a subpopulation of the murine estrogen receptor-α (mER-α) is modified by *O*-GlcNAc at Thr^{575}. Here we mutated mER-α to convert Thr^{575} and Ser^{576} to Val and Ala, respectively. Surprisingly, this glycosylation-site mutant is still extensively modified by *O*-GlcNAc. Analyses of glycopeptides identified two additional sites of modification on mER-α, at Ser^{10} and Thr^{50} near the N-terminus. The major glycosylation sites are within or near PEST regions, suggesting that *O*-GlcNAc may regulate mER-α turnover.

Keywords: Estrogen receptor; *O*-glycosylation; *O*-GlcNAc; PEST domain; Post-translational modification; Phosphorylation; Estrogen

1. Introduction

Estrogen receptor (ER) is a member of the ligand inducible transcription factor family [2]. Like other steroid receptors, ER consists of several distinctive functional domains [3]. Upon binding a cognate ligand, e.g. estrogen, ER undergoes conformational changes and subsequently binds its response element sequence upstream or downstream of target gene [4]. During its activation, ER also undergoes hyperphosphorylation mediated by several kinases such as Casein kinase II [5], MAP kinase [6] and pSrc60 [7]. Ligand induced activation triggers the rapid proteasome dependent degradation of ER [8]. Recently, a homologue of the ER-α, named ER-β [9,10] and some coactivators [11–13] responsible for assembly of an ER-associated transcription complex were discovered.

Post-translational modifications, such as phosphorylation, are likely to be involved in ER regulation. A unique post-translational modification on serine or threonine hydroxyl side chains by *N*-acetylglucosamine, *O*-GlcNAc, has been described [14]. Unlike conventional types of protein glycosylation, which occurs mostly in lumenal compartments along the secretory pathway and on the cell surface, *O*-GlcNAc is abundant on cytosolic and nuclear proteins. To date, a large number of proteins ranging from cytoskeletal proteins [15] to transcription factors [16] have been shown to be *O*-GlcNAcylated. An *O*-GlcNAc transferase was recently cloned from rat [17] and human [18]. Knock-out studies in mice show that the addition of *O*-GlcNAc is required for embryonic stem cell viability [19], and in vivo capping studies suggest that *O*-GlcNAc turnover is

Abbreviations: CHCA, α-cyano-4-hydroxylcinnamic acid; CID-MS, collision induced dissociation of tandem mass spectrometry; galactosyltransferase, Galβ (1-4) galactosyltransferase; Gal, galactose; HPLC, high performance liquid chromatography; LC/ESI–MS, liquid chromatography coupled electrospray ionization mass spectrometry; MALDI-TOF, matrix-assisted laser desorption ionization time of flight; mER-α, murine estrogen receptor-α; 575-mER-α, mER-α *O*-GlcNAc site mutant; *O*-GlcNAc, *O*-linked *N*-acetylglucosamine; PAGE, polyacrylamide gel electrophoresis; RP, reverse phase; SDS, sodium dodecyl sulfate; TFA, trifluoroacetic acid; WGA, wheat germ agglutinin; wt, wild type.

* Corresponding author. Tel.: +1-410-6145993; fax: +1-410-6148804.

E-mail address: gwhart@jhmi.edu (G.W. Hart).

0960-0760/01/$ - see front matter
PII: S0960-0760(00)00167-9

also essential [20]. Analogous to phosphorylation, *O*-GlcNAc is highly dynamic in a variety of cell responses [21,22], suggesting that it may play regulatory roles. Although the direct biological functions of *O*-GlcNAc remain to be determined, the experimental evidence supporting its importance is rapidly mounting. Many in vivo *O*-GlcNAc attachment sites have been mapped. Reciprocal occupancy of the same hydroxyl moiety by *O*-GlcNAc and by *O*-phosphate on some proteins suggests a 'yin-yang' relationship between these two post-translational modifications [23,24].

Since the ER signal transdunction pathway is involved in many aspects of cell regulation, we chose ER as a model to study the role of the *O*-GlcNAcylation. Our previous studies [1] have shown that ER-α proteins isolated from various sources are *O*-GlcNAcylated and that one major in vivo site on mER-α is at Thr^{575} near the carboxyl terminus. In this study, we report that two additional in vivo *O*-GlcNAcylation sites on mER-α expressed in Sf9 insect cells. These sites are located in PEST domains [25], known to be involved in rapid protein turnover, suggesting a possible role for *O*-GlcNAc in mER-α degradation.

2. Materials and methods

2.1. mER-α O-GlcNAcylation site mutant

To mutate the *O*-GlcNAc site at Thr^{575}, the mutagenesis system Altered Sites I (Promega, Madison, WI) was used. Briefly, wt–mER-α cDNA graciously provided by Dr M.G. Parker was excised from SP65 vector and subcloned into pAlter I vector via *Eco*RI site with proper orientation. The mutagenesis primer (CAG CTG GCC ACC GTG GCG TCC ACT TCA GCA C) changing Thr^{575} to Val along with Ser^{576} to Ala was synthesized and used in the mutagenesis process. The mutant was designated 575-mER-α (Fig. 1).

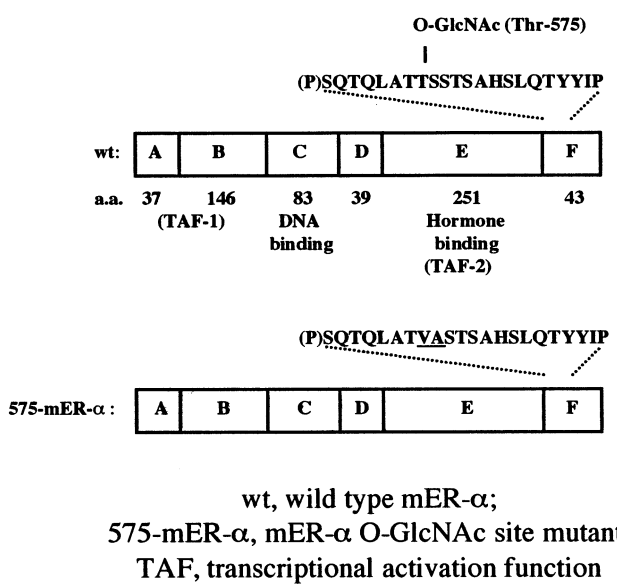

Fig. 1. Schematic of mER-α constructs used in this study. The wt–mER-α cDNA was mutated at 575 (Thr → Val) and 576 (Ser → Ala). The mutated mER-α was designated 575-mER-α. Both cDNAs were engineered with a PCR mediated method to incorporate His_6 tag at the carboxyl terminal end.

To facilitate protein purification, His_6 tag was fused to the carboxyl terminus of either wt–mER-α or 575-mER-α using an established polymerase chain reaction (PCR) method [26]. A 3′ mER-α specific primer annealed to the carboxyl terminus of mER-α encoding His_6 tag and *Hin*dIII site was synthesized and used, in conjunction with a 5′ mER-α primer covering the first amino acid and encoding *Bam*HI site, to amplify mER-α cDNAs. The PCR amplified products were purified, digested with *Bam*HI and *Hin*dIII, and subcloned into the baculovirus transfer vector pBlueBac3 (Invitrogen, San Diego, CA). The final engineered mER-α cDNAs in pBlueBac3 were verified by automated sequencing with serial internal primers covering the entire coding region.

2.2. WGA affinity chromatography on in vitro translated mERs

These analyses were performed as described previously [27]. 575-mER-α cDNA was subcloned into SP65 via an *Eco*RI site, and then in vitro transcribed/translated with coupled rabbit reticulocyte lysate transcription/translation system (Promega, Madison, WI) using SP6 RNA polymerase in the presence of ^{35}S-Met (translation grade, Amersham, Piscataway, NJ). Equal aliquots of translated mixtures were applied to 0.2 ml WGA columns, washed, eluted with 1 M Gal and 1 M GlcNAc. Every fraction (5%) was counted by liquid scintillation counting. For sodium dodecyl sulfate-polyacrylamidegel electrophoresis (SDS-PAGE), proteins were precipitated with 10% (w/v) trichloroacetic acid (TCA), extracted three times with ethanol to remove TCA and solubilized in SDS loading buffer, and examined on 8% SDS-PAGE by autoradiography.

2.3. Expression and purification of mERs from Sf 9 cells

The recombinant baculovirus was generated with the transfection module (Invitrogen, San Diego, CA). The mER-α cDNAs in baculovirus transfer vector were cotransfected into Sf9 cells with wild type AcNPV DNA according to manufacturer's instructions. The recombinant virus was purified by a standard plaque assay. The working stock virus was amplified and titrated to infect approximately 2×10^6 ml^{-1} log phase insect Sf9 cells at 5–10 pfu ml^{-1} in cell suspension.

The infected Sf9 cells were harvested 2 day's post-transfection and washed twice with phosphate-buffered saline (pH 7.4) buffer. The cell pellet was lysed in 6 M guanidine–HCl/20 mM Tris–HCl (pH 7.9)/0.1 M NaCl and sonicated. The mER-α proteins were purified from 20 000 × ***g*** clear lysate with metal affinity resin (Clontech, Palo Alto, CA) in a batchwise fashion. The eluates from the metal affinity resin were purified with preparative SDS-PAGE (Prep-cell, Bio-Rad, Hercules, CA) according to manufacturer's instructions. The proteins were verified by Western blot, pooled, dialyzed, and lyophilized prior to probing for terminal GlcNAc residues by galactosyltransferase labeling.

2.4. Western blot and autoradiography

Western blotting was conducted by the following standard procedures. Proteins were resolved on 8% SDS-PAGE and transferred to PVDF membranes (Bio-Rad). Membranes were probed with a monoclonal antibody H222 directed against human ER (Abbott Laboratories, Abbott Park, IL) at a 1:2000 dilution in 10 mg ml^{-1} bovine serum albumin. The immunoreactive protein was detected with a horseradish peroxidase conjugated secondary antibody followed by ECL (Amersham). For tritiated samples, the gels was stained with Coomassie Blue R250, impregnated with 1 M salicylic acid for 30 min, dried under vacuum, and exposed to X-ray film at −70°C overnight.

2.5. Galactosyltransferase labeling of purified mERs

O-GlcNAc moieties on purified mER proteins were labeled with UDP-[^{3}H]Gal in a reaction catalyzed by galactosyltransferase, as described in the reference [27]. The labeled protein was desalted over a 1.5 × 30 cm Sephadex G50 column equilibrated with 50 mM ammonium formate and 0.1% SDS to remove unincorporated UDP-[^{3}H]Gal, freeze dried, resuspended in a small volume of water, and precipitated with cold acetone at a ratio of 8/1 (v/v). The tryptic peptides were generated by digestion of labeled proteins with 2% (w/w) trypsin (sequencing grade, Boehringer Mannheim, Indianapolis, IN) in 0.1 M bicarbonate (pH 7.4) at 37°C overnight.

2.6. Reverse phase HPLC of tryptic peptides

The first dimension was developed on a C_{18} column (4.6 × 250 mm, Microsorb, Rainin, Woburn, MA) using a linear gradient of acetonitrile (0–60%, v/v) in 0.05% TFA over 90 min at a flow rate 1 ml min^{-1} on a Rainin high performance liquid chromatography (HPLC) system. The eluates were collected (1 ml per fraction) and counted (1%). Peptides were detected at 214 nm. The tritium peaks were pooled, desalted with Sep-Pak C_{18} cartridges (Millipore, Bedford, MA), and lyophilized. Second and third dimension RP–HPLC were conducted using a 0–40% gradient of acetonitrile or other shallow gradients in 0.05% TFA on a C_2/C_{18} column (3.2 × 100 mm) employing a SMART HPLC system (Pharmacia).

For glycosylation site-mapping, smaller peptides were generated using proline specific endopeptidase (0.2 U, Seikagaku Corp., Tokyo, Japan) in the presence of 0.1 M sodium phosphate buffer (pH 7.0) at 37°C for 2 h. Digestions with glutamic acid specific endopeptidase (Glu-C, 0.1 U, Boehringer Mannheim) were carried out in the presence of 2 M urea with 50 mM Tris (pH 7.0) at 37°C overnight.

2.7. Determination of O-GlcNAc sites

The identity of labeled glycopeptides was determined by matrix assisted laser desorption ionization (MALDI) time of flight (TOF) mass spectrometry (Voyager, PerSeptive Biosystems, Framingham, MA). In addition, a liquid chromatography coupled electrospray ionization (LC/ESI–MS) using a capillary C_{18} column linked to an ion-trap mass spectrometer (LCQ, ThermoQuest, Needham Heights, MA) or gas phase Edman sequencing was used to determine the amino acid sequence of the peptides. A standard manual Edman degradation reaction was carried out, as described previously [27], to identify the exact amino acid residue linked to *O*-GlcNAc.

3. Results

3.1. Glycosylation of the mER-α O-GlcNAcylation site mutant

Our previous studies have demonstrated that a subpopulation of ER-α is *O*-GlcNAcylated and that a major site of *O*-GlcNAcylation on mER-α is Thr^{575} [1]. To investigate the roles of *O*-GlcNAcylation on mER-α, we mutated the *O*-GlcNAc site Thr^{575} to Val and the possible *O*-GlcNAc site Ser^{576} to Ala to generate an *O*-GlcNAc site mutant, designated 575-mER-α (Fig. 1). In vitro translated, ^{35}S-methionine-labeled 575-mER-α and wt–mER-α were subjected to WGA affinity chromatography, which specifically recognizes terminal GlcNAc moieties on proteins [27]. 575-mER-α was retained and eluted with 1 M GlcNAc by WGA, and the extent of binding was similar to wt–mER-α, suggesting that 575-mER-α is still extensively glycosylated (Fig. 2).

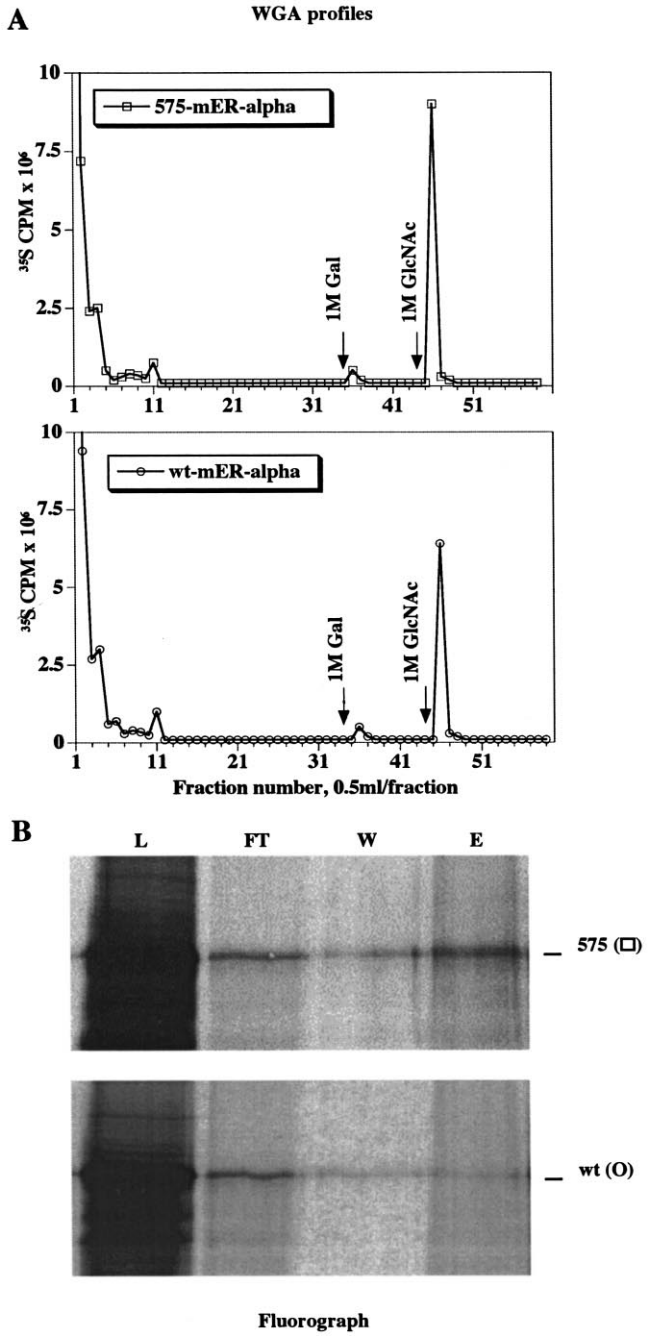

Fig. 2. Characterization of 575-mER-α by WGA chromatography. The mER-α cDNAs were in vitro translated with rabbit reticulocyte lysate in presence of ^{35}S-Met. The translated proteins were individually loaded onto WGA columns and challenged with 1 M Gal. Glycosylated pools of ERs were eluted with 1 M GlcNAc. Fractions (0.5 ml) were collected and monitored by liquid scintillation counting. (A) ^{35}S-WGA chromatography profiles. Upper panel, 575-mER-α; lower panel, wt–mER-α. (B) Proteins in each major fraction were precipitated and resolved by 8% SDS-PAGE gel. The gels were dried and exposed to X-ray film. Lane L, load; lane FT, flow through; lane W, 1 M Gal challenge; lane E, 1 M GlcNAc elute.

3.2. Purification of estrogen receptors expressed in Sf 9 cells

Prior studies have validated the expression of recombinant proteins in insect Sf9 cells as an efficient way to generate large quantities of protein for the study of *O*-GlcNAcylation of low abundance proteins, such as transcription factors [28,29]. While stoichiometry appears variable, *O*-GlcNAc site occupancy in insect cells is the same as that seen in mammalian cells. To simplify protein purification, we made recombinant virus to express His_6-tagged 575-mER-α as well as His_6-tagged wt–mER-α in insect Sf9 cells. After one step purification with metal chelate affinity chromatography under denaturing conditions, the protein from Sf9 cell crude lysate was purified to over 90% homogeneity (Fig. 3, lane E). The eluted protein from metal affinity resin was further purified by preparative SDS-PAGE. Fractions containing mER-α protein from the SDS-PAGE were confirmed by Western blot (Fig. 3, lane E–P). After the second purification step, mER-α migrated as a single band with the correct molecular weight (66 kDa) by analytical SDS-PAGE (Fig. 3).

3.3. Detection of O-GlcNAc on tryptic peptides

We utilized a well-established method, galactosyltransferase labeling using UDP-[^{3}H]Gal as the donor substrate, to detect *O*-GlcNAc moieties on *O*-GlcNAc site mutant 575-mER-α. The purified 575-mER-α protein was well-labeled (Fig. 4A). Combined with the results from WGA lectin chromatography, we concluded that the mutant 575-mER-α is still extensively *O*-GlcNAcylated. Therefore, we proceeded to map the additional *O*-GlcNAcylation sites on mER-α.

3.4. Comparison of glycopeptide tryptic maps between wt–mER-α and 575-mER-α

The [^{3}H]Gal-labeled mER-α proteins were digested with trypsin to generate tryptic fragments. We established C_{18} RP–HPLC tryptic maps for both 575-mER-α and wt–mER-α (Fig. 4B and C). The tryptic map of the mutant 575-mER-α (Fig. 4B) reveals that there are two major tritium peaks designated as glycopeptides A and B, even though the primary site Thr^{575} was deleted. The tryptic map of wt–mER-α also shows at least two major tryptic glycopeptides (Fig. 4C). The similar elution positions of major glycopeptides from wt–mER-α and 575-mER-α mutant suggested that the same peptides are *O*-GlcNAcylated. We proceeded to verify Thr^{575} as one of *O*-GlcNAcylation sites on wt–mER-α.

The major tritium fraction of wt–mER-α (Fig. 4C) was further purified using RP–HPLC (Fig. 5A). The purified tritium fraction was subjected to MALDI-TOF mass spectrometry. Two masses were observed (Fig. 5B). The values 5124.6 and 5490.4 m/z correspond to unmodified peptide (calculated 5101.6 m/z plus one sodium addition) from the carboxyl terminus of wt–mER-α and one Gal-labeled *O*-GlcNAcylated peptide

(calculated 5124.6 + 366 m/z), with a difference of 365.8 representing Gal-GlcNAc (calculated 365.3 m/z). The identity of this peptide was further confirmed by gas phase sequencing. To facilitate manual Edman degradation, this peptide was further digested with proline specific endopeptidase, and resulting smaller peptides were purified with a shallow gradient of acetonitrile on a C_{18} RP column (Fig. 5C). The tritium counts from Gal-labeled glycopeptide were released at the eighth cycle of manual Edman degradation (Fig. 5D), which corresponds to Thr^{575} of wt–mER-α.

3.5. Mapping O-GlcNAc sites on 575-mER-α

To map *O*-GlcNAcylation sites on 575-mER-α, glycopeptide peak A (Fig. 4B) was further purified with a shallow gradient of acetonitrile as the second dimension (Fig. 6A). Although the UV profile shows the additional separation of the peptides, only one major tritium peak was observed, indicating that tritium peak A consists of only one glycopeptide. To determine the identity of glycopeptide A, the tritium rich fraction was analyzed using MALDI-TOF mass spectrometry. As shown in Fig. 6B, three related mass species marked as 2660.5, 2675.9, and 3040.3 m/z, corresponding to unmodified, oxidized, and glycosylated and oxidized peptides, were observed. The observed value of 2660.5 m/z corresponds to the unmodified peptide representing amino acids 9–32 (calculated of 2658.1 m/z) of 575-mER-α. The identity of the oxidized peptide was further confirmed by collision induced dissociation tandem mass spectrometry (CID-MS/MS, Fig. 6C). Although peptide A consists of 24 residues, the serine residue at the second position is the only potential *O*-GlcNAc site. To verify this, manual Edman degradation was employed (Fig. 6D). As expected, most of the tritium counts were released in the second cycle, indicating that glycopeptide A is modified by *O*-GlcNAc at Ser^{10}.

For glycopeptide B in 575-mER-α, a similar approach using MALDI-TOF was attempted. The MALDI-TOF spectrum of glycopeptide B gave only a weak signal and poor quality data (data not shown). The second dimension separation of glycopeptide B, using a shallow gradient of acetonitrile, also gave little separation (data not shown). The high acetonitrile concentration required for the elution implied that glycopeptide B was relatively large. Therefore, the amino terminus of glycopeptide B was directly determined using gas phase sequencing. The result revealed that glycopeptide B is the largest tryptic fragment (11 kDa) from amino acids 38 to 146 of 575-mER-α. Based on the sequence of the fragment, glycopeptide B was digested with endoproteinase Glu-C and the digestion mixture was separated using C_{18} RP–HPLC. After digestion, the elution time of glycopeptide Bon RP–HPLC was shifted from its original 75–46 min. The resulting tritium rich fraction designated as Glu-C B was directly examined using ESI/CID-MS/MS. Three peptides corresponding to the Glu-C digested 11 kDa 575-mER-α tryptic fragments were found in the fraction. They were peptide representing amino acids 42–56 (Fig. 7B), peptide containing amino acids 47–61(data not shown), and the peptide containing amino acid 138–155 (data not shown). Due to the selective loss of the acid labile carbohydrate modification after the repetitive reverse phase purification processes, we were not able to detect the related glycopeptides by mass spectrometry. However, manual Edman degradation indicates that most of the tritium label was released at the ninth cycle of the reaction (Fig. 7C). Since only the peptide amino acid 42–56 has Thr at the ninth position, we conclude that the second *O*-GlcNAc site is at Thr^{50}.

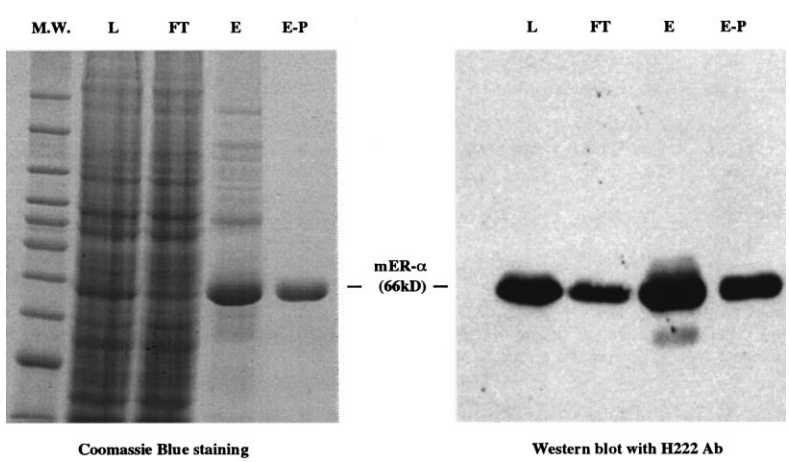

Fig. 3. Purification of His_6 tagged mER-α from Sf9 cells. mER-α expressed in Sf9 cells was purified with metal affinity resin (lane L, FT, and E are load, flow through and elute) and subsequently purified with preparative SDS-PAGE (lane E–P is the prep-cell elute). Aliquots of each fraction were resolved on an analytical 8% SDS-PAGE gel. The gel was either stained with Coomassie Blue or immunoblotted with monoclonal antibody H222 against ER.

4. Discussion

By mutating the major *O*-GlcNAcylation site of mER-α at Thr[575], we were able to identify two additional *O*-GlcNAcylation sites on mER-α at Ser[10] and Thr[50]. These *O*-GlcNAcylation sites on mER-α are summarized in Fig. 8. Interestingly, all these *O*-GlcNAcylation sites on mER-α are located in either N- or C-terminal regions of the protein. This pattern is often seen on other *O*-GlcNAcylated proteins such as neurofilaments [15], synapsin [30], and cytoskeleton proteins [31]. According to the proposed structure of the ER [32], it appears that these sites exist on the surface of the protein and are clustered in three-dimensions allowing for efficient binding to WGA.

The Thr[575] site on mER-α is located on the C-terminal tryptic fragment, which covers most of the F domain of mER-α. Thr[575] is also contained within a long stretch of Ser and Thr residues; only two of which were mutated in the mutant 575-mER-α. Our initial thought, based on the tryptic map comparison shown in Fig. 4, was that this localized region might still be *O*-GlcNAcylated in the mutant and that the modification sites could be on other adjacent Ser or Thr residues not mutated. However, when we made and analyzed a truncated mutant, in which the entire C-terminal tryptic fragment was deleted (designated as the F domain deletion mutant), we found that this deletion mutant is still extensively *O*-GlcNAcylated (data not shown). The tryptic map of [^{3}H] Gal-labeled peptides from the F domain deletion mutant is also similar to one of the mutant 575-mER-α (data not shown). From our site mapping data, it is now evident that the glycopeptide A co-migrates on RP–HPLC with the glycopeptide bearing the Thr[575] site.

An overview of in vivo *O*-GlcNAc sites is reviewed in Table 1. There is, as yet, no obvious primary consensus motif for *O*-GlcNAcylation. It appears that the *O*-GlcNAc transferase preferentially targets some regions with the identity of PEST sequences. PEST sequences enriched with Pro, Glu, Ser, and Thr residues are found in many proteins with short half-lives such as Fos [33]

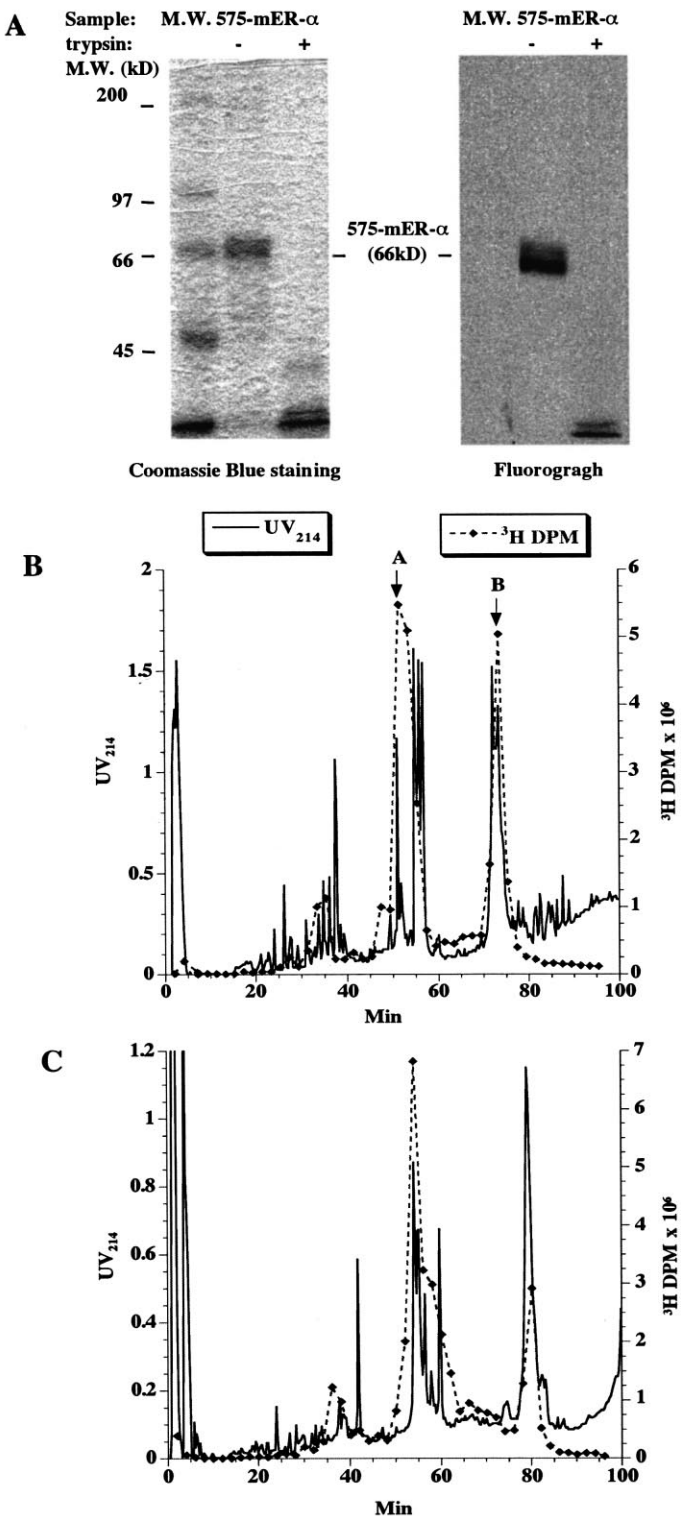

Fig. 4.

Fig. 4. Galactosyltransferase labeling and tryptic maps of wt–mER-α and 575-mER-α. (A) Purified 575-mER-α was labeled with UDP-[^{3}H]Gal using galactosyltransferase and subsequently desalted over a Sephadex G50 column. The protein was precipitated by cold acetone, resuspended in 0.1 M sodium bicarbonate (pH 7.0), and digested with trypsin at a ratio of 1:50 (w/w). Aliquots of each sample were examined by 8% SDS-PAGE gel. The gel was stained with Coomassie Blue R250, impregnated with 1 M salicylic acid, dried down, and exposed to X-ray film at −70°C overnight. (B) [^{3}H]Gal-labeled 575-mER-α tryptic peptides were applied to a C_{18} RP column. Peptides were eluted with a 90-min linear gradient of 0–60% acetonitrile (v/v) in 0.1% TFA from 6 to 96 min at a flow rate of 0.1 ml min^{-1}. Fractions (0.2 ml) were collected and 1% of each was counted. Two major tritium peaks were designated as A and B, respectively. (C) Wild type mER-α was purified and labeled under the same condition as the mutant 575-mER-α. [^{3}H]Gal-labeled wt–mER-α tryptic glycopeptides were applied to a C_{18} RP column. The column was developed under the same condition as panel B.

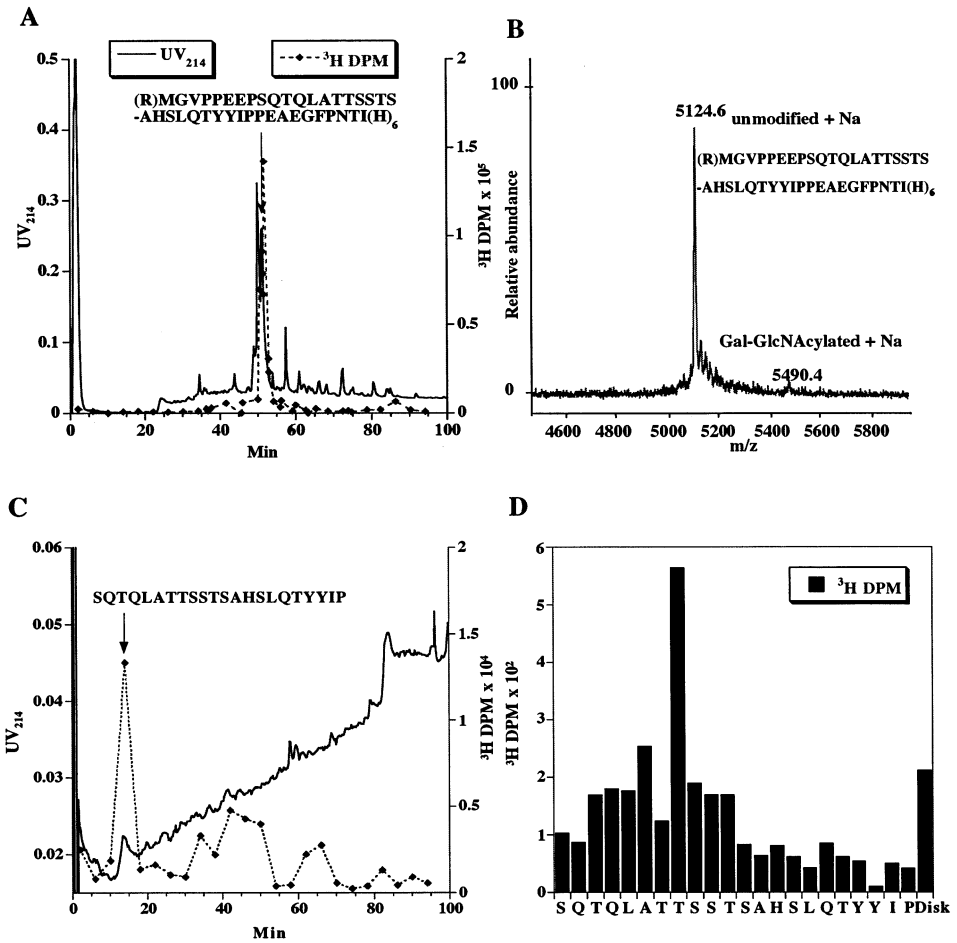

Fig. 5. Thr^{575} is the major *O*-GlcNAcylation site on wt–mER-α. (A) The major [^{3}H] Gal-labeled wt–mER-α tryptic fragment eluting from a C_{18} RP column at 55 min (Fig. 4C). This was further purified by a shallow gradient of 15–45% acetonitrile in 0.1% TFA over 90 min at a flow rate of 0.1 ml min^{-1} on a C_{18} RP column. The fraction carrying the most tritium counts was subjected to gas phase sequencing. The identity of the glycopeptide is indicated. (B) An aliquot of the tritium peak fraction from panel A was mixed with the matrix CHCA (5 mg ml^{-1} in 0.3% TFA/50% acetonitrile) and detected by MALDI TOF. Two mass species were detected and analyzed. (C) The major [^{3}H]Gal-labeled mER-α tryptic C-terminal fragment was digested with proline specific endopeptidase at 37°C overnight and purified by a shallow gradient of 0–30% (v/v) acetonitrile in 0.1% TFA over 90 min at a flow rate of 0.1 ml min^{-1} on a C_{18} RP column. (D) The fraction of [^{3}H]Gal-labeled glycopeptide from panel C was coupled to Sequelon-AA disk (Millipore) and subjected to manual Edman degradation. The tritium counts released from each cycle were monitored using liquid scintillation counting. The major *O*-GlcNAcylation site on wt–mER-α was identified at Thr^{575}.

and ornithine decarboxylase [34]. An algorithm was developed to determine whether a protein contains a PEST region [24]. In the case of the mER-α, the region carrying the Thr^{575} site is a strong PEST candidate with high PEST score (9.62). The other two sites reported in this study are also within the boundaries of weaker PEST regions. ER-α is degraded in a hormone-dependent manner through the ubiqutin–proteasome pathway [8]. A recent study shows that the 26S proteasome is required for ER-α and coactivator turnover and for efficient ER-α transactivation [35]. Since the PEST motif has been implicated in mediating rapid protein degradation through the same pathway, *O*-GlcNAcylation of PEST regions in mER-α may be involved in regulating degradation by preventing rapid protein turnover.

Our preliminary characterization of the mutant 575-mER-α indicated that there is no significant difference between wt–mER-α and 575-mER-α in terms of DNA binding (data not shown), consistent with previous studies on the F domain deletion mutant of ER-α [36]. Interestingly, some PEST domains are conditional signals requiring phosphorylation to fulfill their function [37,38]. Given increasing evidence for a reciprocal relationship between phosphorylation and *O*-GlcNAcylation [23,24], a simple deletion mutant may not reveal the effects mediated by the interplay of phosphorylation and *O*-GlcNAcylation on the same hydroxyl moiety of Ser or Thr residues. One better approach to examine a potential role for phosphorylation is to generate a site specific aspartic acid or glutamic acid mutant that mimics constitutive phosphorylation [39,40].

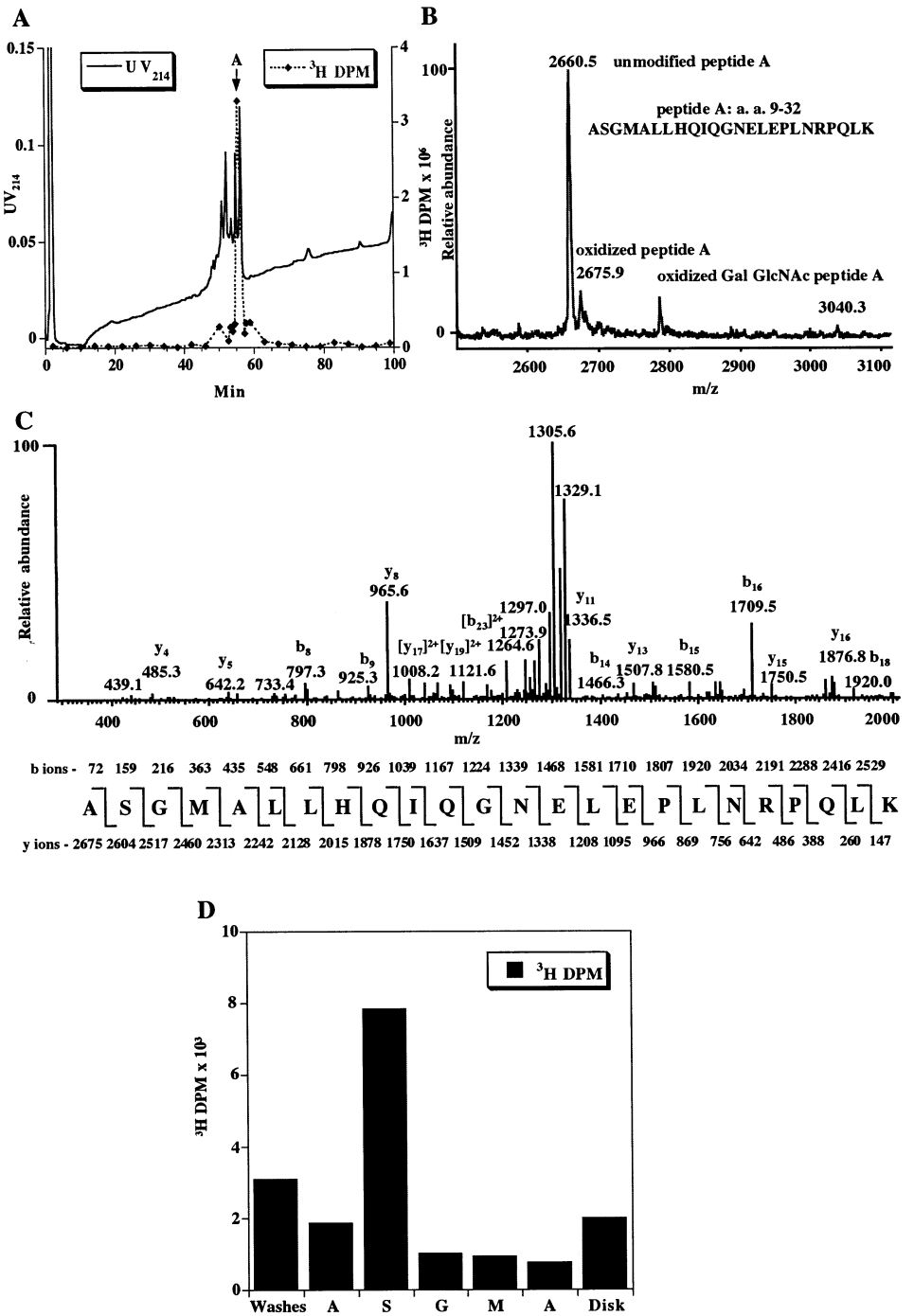

Fig. 6. Ser[10] is the *O*-GlcNAcylation site on glycopeptide A. (A) [^{3}H] Gal-labeled 575-mER-α tryptic glycopeptide A, eluting from a C_{18} RP column at 52 min (Fig. 4B), was further purified by a shallow gradient of 15–45% (v/v) acetonitrile in 0.1% TFA over 90 min at a flow rate of 0.1 ml min^{-1} on a C_{18} RP column. (B) An aliquot of the tritium peak fraction from panel A was mixed with CHCA as reported in Fig. 5 and detected by MALDI-TOF. Three mass species were detected and their deduced identity is indicated. (C) An aliquot of the fraction A from the second dimension (A) was subjected to LC/ESI-MS/MS analysis. For clarification, predicted b-and y-ions of the oxidized peptide are listed below the spectrum. The methionine at the position four was oxidized resulting in a mass addition of 16 Da. The matched ions are indicated in the spectrum. Some unmatched ions are probably due to either internal cleavage or loss of H_2O during the fragmentation process. (D) The fraction of [^{3}H]Gal-labeled glycopeptide A from panel A was coupled to Sequelon-AA disk (Millipore) and subjected to manual Edman degradation. The tritium counts released in each cycle were monitored using liquid scintillation counting. The *O*-GlcNAc site on glycopeptide A was identified at Ser[10].

In addition to regulating protein turn over and acting as an antagonist of phosphorylation, *O*-GlcNAc has been proposed to modulate protein–protein interactions. In p53, the glycosylated protein appears to increase its DNA binding through an *O*-GlcNAc mediated intramolecular interaction [41]. Recent studies

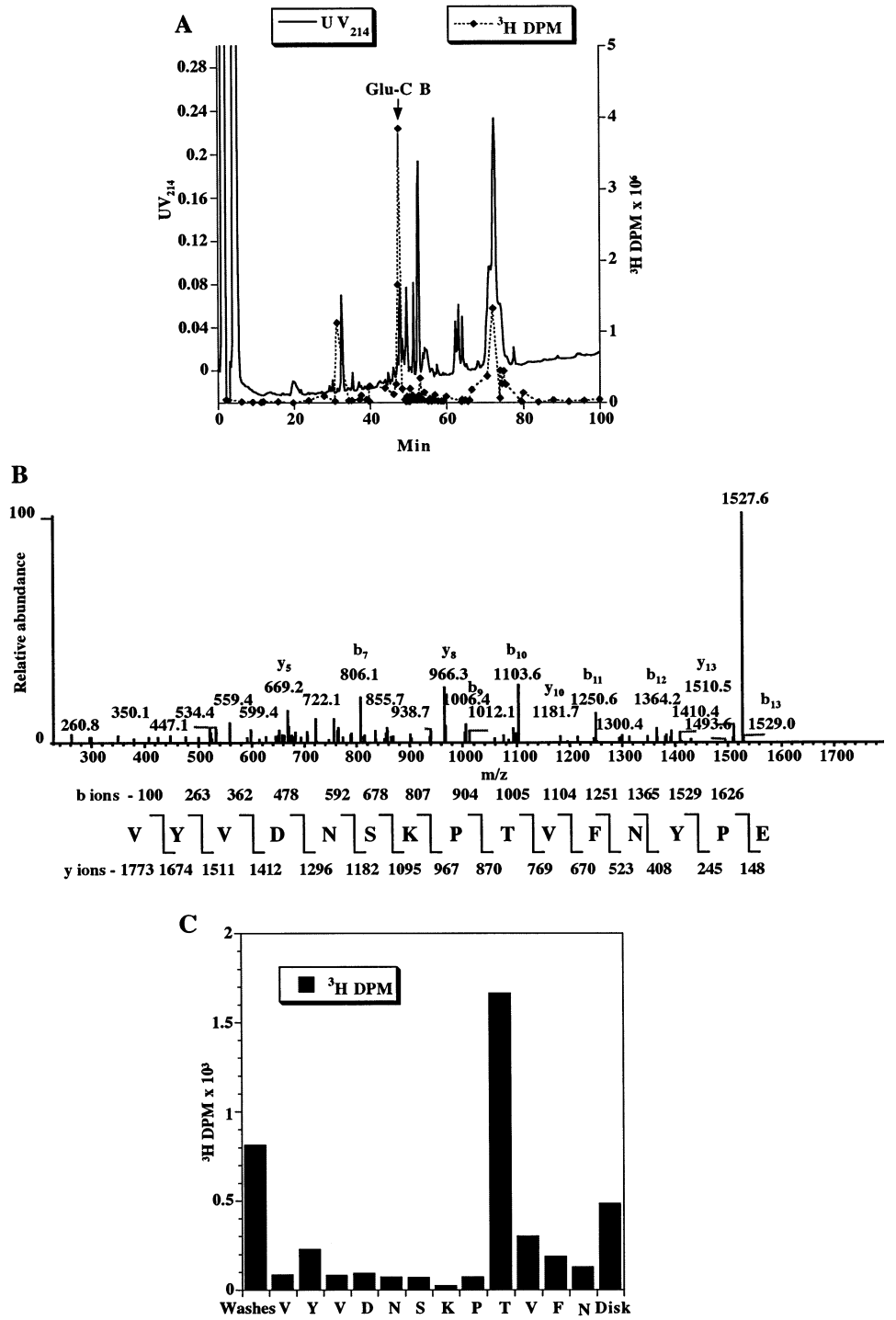

Fig. 7. Thr[50] is the *O*-GlcNAcylation site on glycopeptide B. (A) [^{3}H]Gal-labeled 575-mER-α glycopeptide B eluted at 74 min (Fig. 4B) was further digested with endopeptidase Glu-C in presence of 2 M urea and applied to a C_{18} RP column. The column was developed with a 90-min linear gradient of 0–60% (v/v) acetonitrile in 0.1% TFA at a flow rate of 0.1 ml min^{-1}. The major tritium peak was designated as Glu-C B. (B) An aliquot of the fraction containing Glu-C B was subjected to LC/ESI-MS/MS analysis. The ESI-MS/MS spectrum of the peptide modified by *O*-GlcNAc is shown. For clarification, predicted b- and y-ions of the peptide are listed below the spectrum. (C) The Glu-C B fraction was coupled to Sequelon-AA disk (Millipore) and subjected to manual Edman degradation. The tritium counts released in each cycle were monitored using liquid scintillation counting. The *O*-GlcNAc site on glycopeptide B was identified at Thr[50].

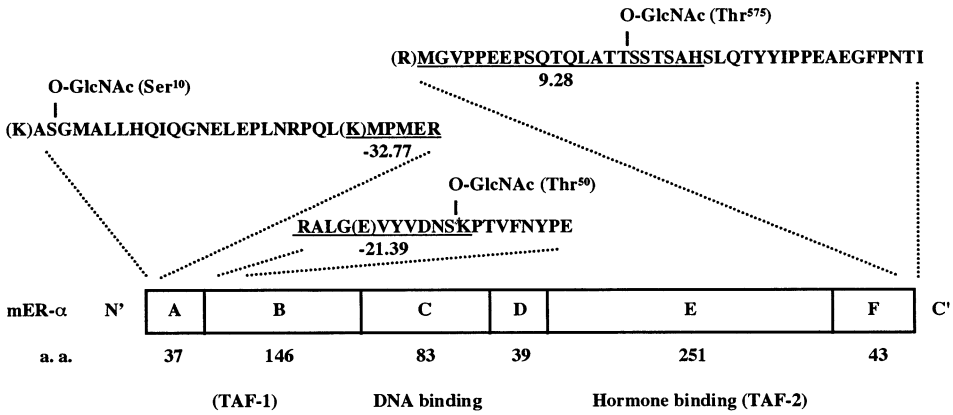

Fig. 8. Summary of *O*-GlcNAcylation sites on mER-α. mER-α's domain structure is indicated. N′ and C′ indicate amino terminal and carboxyl terminal ends. In this study, two *O*-GlcNAc sites on mER-α were identified near amino terminus at Ser^{10} and Thr^{50}. The PEST scores regarding the regions underlined near *O*-GlcNAc sites are also indicated.

Table 1
Summary of *O*-GlcNAcylated proteins and in vivo attachment sites[a]

O-GleNAcylated proteins	Attachment sites	References
Nuclear pore p62	MAGGPADTSDP	[42]
Erythrocyte p65	DSPVSQPSL	[43]
Erythrocyte band 4.1	TTTSETPSS	[44]
αA-crystallin	AIPVSREEK	[45]
αB-crystallin	EEKPAVTAAPK	[22]
Talin	LVDPAXTQ; ANOLTNDY	[31]
Serum response factor	TTSTIQTAP; YLAPVSASV; PSAVSSAD; TQTSSSGTVELPATIM	[46]
RNA polymerase II	S/T P S/T SP; TPTTPN; SPTSPT	[24]
CMV Tegument BPP	PSVPVSGS; PPSSTAK	[29]
Neurofilament-L	YVETPRVHISSV; SGYSTAR; SAPVSSSLSV	[15]
Neurofilament-M	GSPSTVSSSYK; QPSVTISSK; VPTETR; VAGSPSSGFR	
Neurofilament-H	TSVSSVSAS	
Keratin-18	PVSSAASVY; ISVSRSTS	[47]
c-Myc	FELLPTPPLSPSR	[23]
Sp-1	QGVSLGQTSS	[41]
Synapsin I	VSGASPGGQQR; QAGPPQATR; QTTAAAAATF; ASTAAPVASP; AAPQQSASQATPMTQGQGR; LPSPTAAPQQSASQATPMTQGQGR	[39]
Estrogen receptor-α	SQTQLATTSSTSAHSLQTYYIP; ASGMALLHQIQGNELEPLNRPQLK; VYVDNSKPTVFNYPE	[1], This study
Estrogen receptor-β	MNYSVPSSTGNLEGGPVR	[48]

[a] *O*-GlcNAc sites are marked with underlines.

suggest that the presence of *O*-GlcNAc inhibits both Sp1 dimerization and interaction of the Sp1 transactivation domain with TATA binding protein associated factor TAF110 in vitro [42]. Consistent with these findings, our data shows more effective labeling on Thr^{50} site in the mutant 575-mER-α (comparison between Fig. 4B and C). This observation likely results from the change of protein conformation due to the

lack of *O*-GlcNAc moieties on Thr^{575} in 575-mER-α, which makes the mutant protein more accessible to glycosyltransferase probing.

O-GlcNAc is also postulated to play a regulatory role at the transcription level [14]. RNA polymerase II and all of its transcription factors examined to date have been characterized to contain *O*-GlcNAc [16,24]. Not surprisingly, the two *O*-GlcNAc sites mapped here are located at the transactivation domain of the mER-α, raising the possibility that *O*-GlcNAc may also be involved in regulating estrogen responsive gene expression during the transcriptional activation process. Localization of these sites allows functional studies by site-directed mutagenesis.

Acknowledgements

We thank Marion Kirk at the UAB Comprehensive Cancer Center Mass Spectrometry Facility for assistance to analyze mass spectrometry data. We also thank Shane Arnold, Robert Cole, Frank Comer, Glendon Parker, Keith Vosseller, Yuan Gao, and Natasha Zachara in the Hart Lab for critical reading of the manuscript. This work was supported by the National Institutes of health Grant CA42486 (to G.W. Hart).

References

[1] M.S. Jiang, G.W. Hart, A subpopulation of estrogen receptors are modified by *O*-linked *N*-acetylglucosamine, J. Biol. Chem. 272 (1997) 2421–2428.
[2] R. Evans, The steroid and thyroid hormone receptor superfamily, Science 240 (1988) 889–895.
[3] M.G. Parker, Structure and function of estrogen receptors, Vitam. Horm. 51 (1995) 267–287.
[4] V. Kumar, P. Chambon, The estrogen receptor binds tightly to its responsive element as a ligand-induced homodimer, Cell 55 (1988) 145–156.
[5] S.F. Arnold, J.D. Obourn, H. Jaffe, A.C. Notides, Phosphorylation of the human estrogen receptor by mitogen-activated protein kinase and casein kinase II: consequence on DNA binding, J. Steroid Biochem. Mol. Biol. 55 (1995) 163–172.
[6] S. Kato, H. Endoh, Y. Masuhiro, T. Kitamoto, S. Uchiyama, H. Sasaki, S. Masushige, Y. Gotoh, E. Nishida, H. Kawashima, et al., Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase, Science 270 (1995) 1491–1494.
[7] S.F. Arnold, D.P. Vorojeikina, A.C. Notides, Phosphorylation of tyrosine 537 on the human estrogen receptor is required for binding to an estrogen response element, J. Biol. Chem. 270 (1995) 30205–30212.
[8] Z. Nawaz, D.M. Lonard, A.P. Dennis, C.L. Smith, B.W. O'Malley, Proteasome-dependent degradation of the human estrogen receptor, Proc. Natl. Acad. Sci. USA 96 (1999) 1858–1862.
[9] G.G. Kuiper, E. Enmark, M. Pelto-Huikko, S. Nilsson, J.A. Gustafsson, Cloning of a novel receptor expressed in rat prostate and ovary, Proc. Natl. Acad. Sci. USA 93 (1996) 5925–5930.
[10] G.B. Tremblay, A. Tremblay, N.G. Copeland, D.J. Gilbert, N.A. Jenkins, F. Labrie, V. Giguere, Cloning, chromosomal localization, and functional analysis of the murine estrogen receptor-β, Mol. Endocrinol. 11 (1) (1997) 353–365.
[11] V. Cavailles, S. Dauvois, P.S. Danielian, M.G. Parker, Interaction of proteins with transcriptionally active estrogen receptor, Proc. Natl. Acad. Sci. USA 91 (1994) 10009–10013.
[12] D.P. Edwards, Coregulatory proteins in nuclear hormone receptor action, Vitam. Horm. 55 (1999) 165–218.
[13] H. Shibata, T.E. Spencer, S.A. Onate, G. Jenster, S.Y. Tsai, M.J. Tsai, B.W. O'Malley, Role of co-activators and co-repressors in the mechanism of steroid/thyroid receptor action, Recent Prog. Horm. Res. 52 (1997) 141–165.
[14] G.W. Hart, Dynamic *O*-linked glycosylation of nuclear and cytoskeletal proteins, Ann. Rev. Biochem. 66 (1997) 315–335.
[15] D.L. Dong, Z.S. Xu, G.W. Hart, D.W. Cleveland, Cytoplasmic *O*-GlcNAc modification of the head domain and KSP repeat motif the neurofilament protein neurofilament-H, J. Biol. Chem. 271 (1996) 20845–20852.
[16] S.P. Jackson, R. Tijian, *O*-glycosylation of eukaryotic transcription factors: implications for the mechanisms of transcriptional regulation, Cell 63 (1988) 155–165.
[17] L.K. Kreppel, M.A. Blomberg, G.W. Hart, Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique *O*-GlcNAc transferase with multiple tetratricopeptide repeats, J. Biol. Chem. 272 (1997) 9308–9315.
[18] W.A. Lubas, D.W. Frank, M. Krause, J.A. Hanover, *O*-linked GlcNAc transferase is a conserved nucleocytoplasmic protein containing tetratricopeptide repeats, J. Biol. Chem. 272 (1997) 9316–9324.
[19] R. Shafi, S.P. Iyer, L.G. Ellies, N. O'Donnel, K.W. Marek, D. Chui, G.W. Hart, J.D. Marth, The *O*-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny, Proc. Natl. Acad. Sci. USA 97 (2000) 5735–5739.
[20] D.M. Snow, J.H. Shaper, N.L. Shaper, G.W. Hart, Determination of β1,4-galactosyltransferase enzymatic activity by capillary electrophoresis and laser-induced fluorescence detection, Anal. Biochem. 271 (1999) 36–42.
[21] K.P. Kearse, G.W. Hart, Lymphocyte activation induces rapid changes in nuclear and cytoplamic glycoproteins, Proc. Natl. Acad. Sci. USA 88 (1991) 1701–1705.
[22] E.P. Roquemore, M.R. Chevrier, R.J. Cotter, G.W. Hart, Dynamic *O*-GlcNAcylation of the small heat shock protein αB-crystallin, Biochemistry 35 (1996) 3578–3586.
[23] T.Y. Chou, G.W. Hart, C.V. Dang, c-Myc is glycosylated at threonine 58, a known phosphorylation site and a mutational hot spot in lymphomas, J. Biol. Chem. 270 (1995) 18961–18965.
[24] W.G. Kelly, M.E. Dahmus, G.W. Hart, RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by *O*-GlcNAc, J. Biol. Chem. 268 (1993) 10416–10424.
[25] M. Rechsteiner, S.W. Rogers, PEST sequences and regulation by proteolysis, Trends Biochem. Sci. 21 (1996) 267–271.
[26] F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl, Metal–chelate affinity chromatograph. Curr. Protein Mol. Biol. 2 (1993) 10.11.8.
[27] E.P. Roquemore, T.Y. Chou, G.W. Hart, Detection of *O*-linked *N*-acetylglucosamine (*O*-GlcNAc) on cytoplasmic and nuclear proteins, Methods Enzymol. 230 (1994) 443–460.
[28] N.O. Ku, M.B. Omary, Expression, glycosylation, and phosphorylation of human keratins 8 and 18 in insect cells, J. Biol. Chem. 270 (1995) 11820–11827.
[29] K.D. Greis, W. Gibson, G.W. Hart, Site-specific glycosylation of the human cytomegalovirus tegument basic phosphoprotein (UL32) at serine 921 and serine 952, J. Virol. 68 (1994) 8339–8349.
[30] R.N. Cole, G.W. Hart, Glycosylation sites flank phosphorylation sites on synapsin I: *O*-linked *N*-acetylglucosamine residues are localized within domains mediating synapsin I interactions, J. Neurochem. 73 (1999) 418–428.

[31] J. Hagmann, M. Grob, M.M. Burger, The cytoskeletal protein talin is *O*-GlcNAcylated, J. Biol. Chem. 267 (1993) 16679–16687.
[32] M. Beato, P. Herrlich, G. Schutz, Steroid hormone receptors: many actors in search of a plot, Cell 83 (1995) 851–857.
[33] C. Tsurumi, N. Ishida, T. Tamura, A. Kakizuka, E. Nishida, E. Okumura, T. Kishimoto, M. Inagaki, K. Okazaki, N. Sagata, et al., Degradation of c-Fos by the 26S proteasome is accelerated by c-Jun and multiple protein kinases, Mol. Cell. Biol. 15 (1995) 5682–5687.
[34] L. Ghoda, T. van Daalen Wetters, M. Macrae, D. Ascherman, P. Coffino, Prevention of rapid intracellular degradation of ODC by a carboxyl-terminal truncation, Science 243 (1989) 1493–1495.
[35] D.M. Lonard, Z. Nawaz, C.L. Smith, B.W. O'Malley, The 26S proteasome is required for estrogen receptor-α and coactivator turnover and for efficient estrogen receptor-α transactivation, Mol. Cell 5 (2000) 939–948.
[36] F. Pakdel, P. Le Goff, B.S. Katzenellenbogen, An assessment of the role of domain F and PEST sequences in estrogen receptor half-life and bioactivity, J. Steroid Biochem. Mol. Biol. 46 (1993) 663–672.
[37] K. Brown, S. Gerstberger, L. Carlson, G. Franzoso, U. Siebenlist, Control of I κ B-α proteolysis by site-specific, signal-induced phosphorylation, Science 267 (1995) 1485–1488.
[38] M.L. MacKichan, F. Logeat, A. Israel, Phosphorylation of p105 PEST sequence via a redox-insensitive pathway up-regulates processing of p50 NF-κB, J. Biol. Chem. 271 (1996) 6084–6091.
[39] S.P. Collins, M.D. Uhler, Cyclic AMP- and cyclic GMP-dependent protein kinases differ in their regulation of cyclic AMP response element-dependent gene transcription, J. Biol. Chem. 274 (1999) 8391–8404.
[40] F. Bornancin, P.J. Parker, Phosphorylation of protein kinase C-α on serine 657 controls the accumulation of active enzyme and contributes to its phosphatase-resistant state, J. Biol. Chem. 272 (1997) 3544–3549.
[41] P. Shaw, J. Freeman, R. Bovey, R. Iggo, Regulation of specific DNA binding by p53: evidence for a role for *O*-glycosylation and charged residues at the carboxy-terminus, Oncogene 12 (1996) 921–930.
[42] M.D. Roos, K. Su, J.R. Baker, J.E. Kudlow, *O*-glycosylation of a Sp1-derived peptide blocks known Sp1 protein interactions, Mol. Cell. Biol. 17 (1997) 6472–6480.
[43] G.W. Hart, R.S. Haltiwanger, G.D. Holt, W.G. Kelly, Nucleoplasmic and cytoplasmic glycoproteins, Ciba Found. Symp. 145 (1989) 102–118.
[44] M. Inaba, Y. Maede, *O*-N-acetyl-D-glucosamine moiety on discrete peptide of multiple protein 4.1 isoforms regulated by alternative pathways, Biol. Chem. 264 (1989) 18149–181550.
[45] E.P. Roquemore, A. Dell, H.R. Morris, M. Panico, A.J. Reason, L.A. Savoy, G.J. Wistow, J.S. Zigler Jr., B.J. Earles, G.W. Hart, Vertebrate lens alpha-crystallins are modified by *O*-linked N-acetylglucosamine, J. Biol. Chem. 267 (1992) 555–63.
[46] A.J. Reason, H.R. Morris, M. Panico, R. Marais, R.H. Treisman, R.S. Haltiwanger, G.W. Hart, W.G. Kelly, A. Dell, Localisation of *O*-GlcNAc modification on the serum response transcription factor, Biol. Chem. 267 (1992) 16911–169210.
[47] N.O. Ku, M.B. Omary, Identification and mutational analysis of the glycosylation sites of human keratin 18, J. Biol. Chem. 270 (1995) 11820–118227.
[48] X.G. Cheng, J. Zaia, R.N. Cole, G.W. Hart, Alternative O-glycosylation/O-phosphorylation of the murine estrogen receptor b. Biochem. 39 (2000) 11609–11620.