A Subpopulation of Estrogen Receptors Are Modified by O-Linked N-Acetylglucosamine*

(Received for publication, July 23, 1996, and in revised form, September 20, 1996)

Man-Shiow Jiang‡§ and Gerald W. Hart¶

From the ‡Department of Biological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205 and the ¶Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, Alabama 35294-0005

Estrogen receptors (ER) are ligand-inducible transcription factors regulated by Ser(Thr)-O-phosphorylation. Many transcription factors and eukaryotic RNA polymerase II itself are also dynamically modified by Ser(Thr)-O-linked N-acetylglucosamine moieties (O-GlcNAc). Here we report that subpopulations of murine, bovine, and human estrogen receptors are modified by O-GlcNAc.

O-GlcNAc moieties were detected on insect cell-expressed, mouse ER (mER) by probing with bovine milk galactosyltransferase, followed by structural analysis. Wheat germ agglutinin-Sepharose affinity chromatography readily detected terminal GlcNAc residues on subpopulations of ER purified from calf uterus, from human breast cancer cells (MCF-7), or from mER produced by in vitro translation. These data suggest that greater than 10% of these populations of estrogen receptors bear O-GlcNAc. Site mapping of insect cell-expressed mER localized one major site of O-GlcNAc addition to Thr-575, within a PEST region of the carboxyl-terminal F domain. Based upon their relative resistance to both hexosaminidase and to in vitro galactosylation, O-GlcNAc moieties appear to be largely buried on native mER. This dynamic saccharide modification, like phosphorylation, may play a role in modulating the dimerization, stability, or transactivation functions of estrogen receptors.

The estrogen receptor (ER) is a ligand-activated transcription factor that modulates specific gene expression by binding to estrogen response elements (ERE) (1). ER and other members of the nuclear receptor superfamily of ligand-regulated transcription factors have discrete domains responsible for ligand binding, DNA binding, nuclear localization, dimer formation or activation of transcription (1). However, the mechanisms of how gene transcription is modulated by steroid hormone binding to these receptors remains largely unknown.

Almost all members of the steroid hormone receptor superfamily are known to be phosphoproteins (2–8). ER from mouse and human become hyperphosphorylated on serine residues in the presence of estradiol (3–5, 7). Tyrosine phosphorylation of ER has also been reported (9–12). Phosphorylation of ER has been postulated to modulate dimerization, interactions with other proteins, transcriptional activity, DNA-binding activity, nuclear localization, and other functions (2).

Ser(Thr)-O-linked N-acetylglucosamine (O-GlcNAc; 13, 14) is a form of dynamic (15–18) intracellular protein glycosylation, which can be reciprocal with Ser(Thr)-O-phosphorylation (for review, see Ref. 19). For example, the phosphorylation and O-GlcNAcylation of the carboxyl-terminal domain of RNA polymerase II (20) and of Thr-58 on the c-myc oncogene (21, 22) are mutually exclusive events. O-GlcNAc has the attributes of a regulatory modification analogous to phosphorylation (15, 23). O-GlcNAc is ubiquitous in all eukaryotes, thus far examined. Unlike more conventional types of protein glycosylation, O-GlcNAc is nearly exclusively localized in the nucleoplasm and cytoplasm (24–27). Other O-GlcNAc modified proteins identified to date include nuclear pore proteins (28–31), vertebrate lens protein (32), small heat shock proteins (33), tumor suppressors (41), oncoproteins (21, 22, 42), the β-amyloid precursor protein (43), and a number of eukaryotic RNA polymerase II transcription factors (44–46). Importantly, the v-erb A oncoprotein whose gene is derived from a host cell gene for thyroid hormone receptors is also modified by O-GlcNAc (42). Here we show that the estrogen receptors from murine, bovine, and human sources are all modified by O-GlcNAc and that a major O-GlcNAc site (Thr-575) occurs within a PEST sequence of the carboxyl-terminal F domain.

EXPERIMENTAL PROCEDURES

Materials—UDP-[6-3H]galactose (38 Ci/nmol) was obtained from Amersham Corp. Bovine milk galactosyltransferase and aprotinin were purchased from Sigma, and galactosyltransferase (37 units/ml) was pre-galactosylated as described (24). All other chemicals were of the highest quality commercially available. SP65 (MOR1-599) (47) and baculovirus MOR1-599 (48) were supplied by Dr. M. G. Parker (Imperial Cancer Research Fund, London). PGEM-7zf (+) that contained eight tandem copies of the consensus ERE was provided by R. A. Bambara (University of Rochester) (49). Anti-human estrogen receptor antibody (H222) was the gift of Abbott Laboratories (50).

Cell Culture and Overexpression of the Mouse Estrogen Receptor in Spodoptera frugiperda Cells—SF-9 insect cells were grown in TNH-FH medium, which consists of Grace’s insect medium (Flow Laboratories) supplemented with 10% heat-inactivated fetal bovine serum, plated out at 70% confluence, and infected with MOR1-599 recombinant or a wild type virus at a multiplicity of infection of 5–10 plaque-forming units per...
cell and then incubated for 24—120 h. Cells were harvested, collected by centrifugation, and frozen at —70 °C.

Purification of the Estrogen Receptor with Sequence-specific ERE-DNA Affinity Chromatography—Insect cells (SF-9) were infected with MOR1-599 recombinant. At 3 days postinfection, insect cells were harvested, cell pellets were washed with 20% (v/v) glycerol, 0.2 mM PMSF, 0.5 mM phenylmethylsulfonyl fluoride (PMSF), 50 mM GlcNAc, pH 7.5), passing the suspension 10 times through a 27-gauge needle, and subsequently removing debris by centrifugation at 50,000 × g for 20 min. The high speed microsomal supernatant was prepared by centrifugation at 200,000 × g for 45 min, treated for 20 min with ENHANCE (DuPont NEV), dried under vacuum, and exposed to preflushed X-Omat diagnostic film (Eastman Kodak Co.) at —80 °C.

Silver Staining Gel—Gels were electrophoresed for 14—16 h at 40 V and then silver-stained by the method of Blum et al. (57).

Western Blot Analysis—Samples resolved by SDS-polyacrylamide gel electrophoresis were blotted to nitrocellulose membranes (Millipore Corp., Bedford, MA). Blots were blocked with 3% non-fat dried milk (Carnation) in TBS-T (20 mM Tris-HCl, 150 mM NaCl, 0.2% Tween 20, pH 7.5) for 2 h at room temperature, washed with TBS-T, and probed with anti-ER antibody (monoclonal antibody H222, Abbott Laboratories) at 2 μg/ml as a primary antibody and with horseradish peroxidase-conjugated goat anti-rat IgG (Pierce) (1:5000 as the secondary antibody). Each incubation was carried out at room temperature for 2 h using antibodies diluted with 1% non-fat dried milk in TBS-T. Probed blots were then washed four times for 5 min each with TBS-T. Secondary antibody was detected by chemiluminescence using reagents from the Amersham Corp.

Preparation of the DNA Affinity Resin—A plasmid (gift from Dr. Robert A. Bambara) containing eight tandem copies of the consensus ERE was grown in Escherichia coli. Plasmid DNA was isolated and digested with EcoRI and HindIII according to the method of Peale et al. (49). The 337-base pair ERE fragment was purified and conjugated to Sepharose-4B as described (49, 58).

Carbohydrate Characterizations—Gel-purified [3H]galactose-labeled mER was treated with mild base (0.1 N NaOH, 1 N NaBH₄) for 24 h at 37 °C. The reaction was stopped by adding acetic acid, and the samples were chromatographed over an Sephadex G-50 desalting column. PNGase F-treatment, to remove the N-linked saccharides, was performed as previously described (24). [3H]Gal-labeled ovalbumin, which contains only N-linked oligosaccharides, was treated with PNGase F as a positive control. The alkaline cleavage products were identified as [3H]Galα1–3GalNAcAcitol by a high resolution TSK gel permeation chromatography (48) and high pH anion exchange chromatography with plus anion exchange chromatography (HPAEC-PAD) (Dionex) (59).

Trypsin in Gel Digestion of [3H]Galactose-labeled mER and Isolation of "H-Labeled Glycopeptides—Gel-purified [3H]Galactose-labeled mER was purified by 7.5% SDS-polyacrylamide gel electrophoresis and high pH anion exchange chromatography with plus anemopheric detection (HPAEC-PAD) (Dionex) (59). Purified glycopeptides from the third dimension of RP-HPLC were sequenced by automated Edman degradation at the method of Peale et al. (49). The 337-base pair ERE fragment was purified and conjugated to Sepharose-4B as described (49, 58).

Preparation of Calf and Human Estrogen Receptors—Cytosol (55) was prepared from calf uterine (49) or human breast cancer cells (MCF-7). The estrogen receptors were labeled with [3H]labeled ligand by incubation with charcoal-dextran, an aliquot of labeled cytosol was subjected to 10 ml of Sephadex G-25 mini-column (PD-10, Pharmacia) which was previously washed with buffer (10 mM phosphate buffer, 150 mM NaCl, pH 7.5). Labeled cytosol in phosphate-buffered saline was used for WGA chromatography.

Hexosaminidase to Remove Terminal GlcNAc—Twenty microliters of 2% SDS were added to 20 μl of the translation mixture, and then the sample was boiled for 5 min. Subsequently, 40 μl of 2 × of the reaction buffer (2 × solution: 80 mM Tris-HCl, 8% Triton X-100, pH 7.5) was added and mixed well. Two microliters of 0.5 units/μl jack bean β-n-Acetylhexosaminidase (V-Labs, Inc.) was then added and incubated for 4 h at room temperature. Boiled hexosaminidase was added to the control mixture.

Gel Electrophoresis and Autofluorography—Proteins were resolved by electrophoresis on 7.5% SDS-PAGE according to the method of Laemmli (56). After SDS-polyacrylamide gel electrophoresis, the gels were fixed in 10% acetic acid, 40% methanol, stained with Coomassie Brilliant Blue G-250, dried at —80 °C, and exposed to preflushed X-Omat diagnostic film (Eastman Kodak Co.) at —80 °C.

RESULTS

Overexpression and Purification of Full-length Recombinant Mouse Estrogen Receptor—Overexpression of the full-length mouse estrogen receptor (mER) was examined in insect SF9 cells from 24 to 120 h following infection with the recombinant
Glycosylation of Estrogen Receptors

**O-GlcNAcylation of Estrogen Receptors**

The Recombinant, Baculovirus-expressed Mouse ER Bears O-Linked GlcNAc Monosaccharide Moieties—O-Linkage of GlcNAc to proteins through serine or threonine hydroxyls is sensitive to alkali-induced β-elimination (13) but is resistant to cleavage by the enzyme peptide-N-glycosidase F (PNGase F; 66, 67). More than 95% of the radioactivity on the gel-purified [3H]galactose-labeled ER was released by alkali-induced β-elimination (Fig. 5A). In contrast, identical preparations of ER were resistant to PNGase F treatment and were eluted in the V₀ of the desalting column (Fig. 5B). In contrast, [3H]galactose-labeled ovalbumin, which contains only N-linked oligosaccharides, is quantitatively sensitive to PNGase F treatment (data not shown). The results indicate that GlcNAc is covalently attached to the ER through an O-glycosidic linkage.

To confirm the saccharide structure, we first chromatographed the ER β-elimination products (Vₜ fractions in Fig. 5A) on a TSK-gel filtration column. The released sugars were then analyzed by gas-phase sequencing for 10 cycles. Both fractions yielded the authentic [3H]Gal-GlcNAcitol disaccharides, is quantitatively sensitive to PNGase F treatment (data not shown).

**Purification and Sequence Analysis of [3H]Galactose-Labeled Glycopeptides from Mouse ER**—The [3H]galactose-labeled ER was purified by SDS-polyacrylamide gel electrophoresis, and in-gel digestion was used to obtain tryptic glycopeptides. RP-HPLC C-18 chromatography resolved at least five radioactive peaks (Fig. 6A). The shoulder between peaks 2 and 3 was variable in different preparations. Peaks 1–4 were too small to analyze further. Peak 5 was further purified by a second round of RP-HPLC (Fig. 6B). Fraction 69 was further purified by a third RP-HPLC run (Fig. 6C). The major radiolabeled materials in fractions 92 and 93 (Fig. 6C) were each subjected to gas-phase sequencing for 10 cycles. Both fractions yielded the beginning sequence MVGVP... (Fig. 6C), corresponding to the first 10 amino residues of tryptic peptide MVGVP...
O-GlcNAcylation of Estrogen Receptors

Peptidase Digestion of ER Exposes 2–3-Fold More GlcNAc Residues—Resistance to hexosaminidase digestion of in vitro labeled mER suggests that like certain other O-GlcNAcylated proteins (33, 69), the majority of the O-GlcNAc residues are inaccessible on native mER. We investigated this further by comparing the galactosylation of native mER to the galactosylation of an equal amount of mER tryptic fragments. Tryptic fragments of mER displayed from 2 to 3-fold more galactosylatable GlcNAc residues per mol of protein than native mER. However, the same tryptic glycopeptides were detected regardless of whether the trypsin treatment was performed before or after probing with galactosyltransferase (data not shown).

A Subpopulation of Both Calf Uterine and Human Breast Cancer Cell ER Bind to WGA- Sepharose—[3H]Tamoxifen aziridine was used to covalently label calf uterine ER at its active site (Cys-530; Refs. 55, 70, 71). The tritiated labeled calf uterine cytosol was subjected to WGA chromatography (Fig. 7A). Approximately 5–10% [3H]tamoxifen aziridine-labeled ER was found in GlcNAc-eluting fractions (compare lanes 1 and 3, Fig. 7B). However, analysis by immunoblotting indicated that the percentage of ER that bound to WGA-Sepharose might be substantially higher (compare lanes 1 and 3, Fig. 7C). Virtually none of the radiolabeled material was nonspecifically bound to the Sepharose column (data not shown). Specificity for WGA binding was confirmed since the ER band was not found in 1 M galactose fractions (Fig. 7B, lane 2). Low molecular weight labeled fragments are proteolytic bands found very commonly in cell extracts even though the protease inhibitor (PMSF) was included during homogenization (55).

We used the same method to investigate whether O-GlcNAc residues modify human ER. The postmicrosomal supernatant from MCF-7 human breast cells (55) was labeled with tamoxifen aziridine and then subjected to WGA-Sepharose chromatography. About 10% [3H]tamoxifen-labeled hER also binds WGA-Sepharose and is eluted specifically with GlcNAc (data not shown). These data indicate that a significant subpopulation of both calf ER and hER are also modified by O-GlcNAc.

DISCUSSION

These studies demonstrate that a significant subpopulation (10% or more) of estrogen receptors from mouse, bovine, or human sources are modified by Ser(Thr)-O-GlcNAc, a highly dynamic form of intra cellular glycosylation that is often reciprocal with Ser(Thr)-O-phosphorylation. Given the importance of Ser(Thr)-phosphorylation in the functions of steroid receptors (2), the modification of ER by O-GlcNAc is also likely to have functional significance.

In order to purify a sufficient amount of ER protein for site mapping of O-GlcNAc glycosylation, we overexpressed the recombinant mER in the baculovirus system. Sites used by O-GlcNAc glycosylation in the baculovirus system appear to be the same as in mammalian cells (38). Recently, Greis and Hart...
(38) showed that the human cytomegalovirus tegument basic phosphoprotein is glycosylated in insect cells and that the glycopeptides produced by chemical or enzymatic digestion are the same as those from native basic phosphoprotein isolated from human cytomegalovirus virions. In addition, it has been shown that cytokeratins 8 and 18 expressed as recombinant baculovirus in insect cells are modified by O-GlcNAc and the major sites of glycosylation are the same as those found in human HT29 cells (17).

In initial studies, we could not detect terminal GlcNAc residues on calf ER or recombinant mER that had been purified from an ERE-DNA affinity column. We may have selectively purified the nonglycosylated forms that bind to the ERE ele-
ment. The relatively low stoichiometry of O-GlcNAcylation on ER purified by other means may reflect the following: (i) hexosaminidase removal of O-GlcNAc during homogenization and purification (72); (ii) multiple forms of ER, with only a subset of ligand binding forms of ER actually containing O-GlcNAc; and (iii) differential accessibility of GlcNAc sites assayed by WGA binding and hexosaminidase digestion on ligand binding forms and other forms of ER. It has been postulated that O-GlcNAc is

**FIG. 6.** Reversed-phase HPLC and sequence analysis of the major gel-purified [3H]galactose-labeled mouse ER tryptic peptides. A, first dimension RP-HPLC with an 120-min linear gradient of 0–60% acetonitrile in 0.1% phosphoric acid, 100 mM NaClO₄, pH 2.1. B, second dimension RP-HPLC. Peak 5 (A) was pooled and subjected to a second dimension C-18 column RP-HPLC with an 120-min gradient of 0–40% acetonitrile in 0.1% trifluoroacetic acid. C, third dimension RP-HPLC and manual Edman degradation analysis. The fraction 69 (B) was further purified on a third dimension of RP-HPLC with an 120-min gradient of 0–30% acetonitrile in 0.1% trifluoroacetic acid. Bottom panel, counts released from manual Edman degradation analysis of the fraction 92 in upper panel of C. D, RP-HPLC analysis of prolidase-treated peptides. The fraction 92 (C) was digested with 2 units of proline-specific endopeptidase (closed circle) or mock-treated (open circle). RP-HPLC of prolidase-treated samples was developed with an 80-min gradient of 0–25% acetonitrile in 0.1% trifluoroacetic acid. Bottom panel, counts released from manual Edman degradation analysis of the fraction 18 (upper panel of D). Top panel of A, B, and C, UV profile of eluted tryptic peptides; bottom panel of A and B, middle panel of C, and upper panel of D, tritium profile of eluted peptides. Y axis cpm/x, where x = 100, 50, or 20 indicates the proportion of the sample analyzed for radioactivity.
1M galactose and then the bound material was subsequently eluted was subjected to WGA chromatography. The column was washed with the evolutionary conserved regions C and E contain the DNA and hormone-binding domains, respectively. Two transcription activation regions are located in the region A/B (TAF-1) and in the hormone binding domain (TAF-2) of the ER. The PEST sequence within F domain in mER is which indicates the PEST region could be multiply modified. O-GlcNacylation within the PEST region of F domain may play an important role in regulating the breakdown of ER proteins. Katzenellenbogen and colleagues (73) have deleted the last 42 amino acids (F domain) of the hER and found that the deletion of F domain did not affect transactivation ability, ligand binding affinity, or the phosphorylation pattern of the receptor and suggested that F domain was not involved in the relatively rapid breakdown of ER protein. However, the glycosylation state of domain F in human ER has not been characterized.

Although the functional role of O-GlcNac has not yet been elucidated, evidence supports the hypothesis that O-GlcNac is a post-translationally regulatory modification and could regulate the protein phosphorylation by blocking the identical (20, 22) or nearby Ser/Thr sites (14, 19). ER becomes hyperphosphorylated on Ser residues in a hormone-dependent manner (3, 4, 6, 7). If O-GlcNac and phosphate are reciprocal or their regulation is related, then the ligand-free binding forms of ER may be glycosylated and the ligand binding forms may only contain basal O-GlcNac moieties. Our initial data suggest that the nonglycosylated form of estrogen receptor is selectively bound to an ERE-DNA affinity column, suggesting that the nonglycosylated ER may be the active ERE-binding form. While as yet inconclusive, these results are consistent with Notides and colleagues (7) showing that treatment of the ER with potato acid phosphatase resulted in the dephosphorylation of the p53 tumor suppressor. Thus, a postulate that we will explore is that dynamic O-glycosylation of the ER might play a crucial role in transcriptional regulation by modulating the activity of the ER protein. Given the substantial technical difficulties in removing O-GlcNac via hexosaminidase digestion of native ER, site mapping and mutagenesis studies are underway to elucidate the functions of O-GlcNac on ER.

Acknowledgments—We thank M. G. Parker for kindly sending us MOR1-599 and MOR121-339 (both in SP65 and baculovirus) and SP16. We also thank Dr. R. A. Bambara for generosity in sending us pGEM-Zf(TAF-2) of the ER. The PEST region, is an O-GlcNacylation site on mER. Ser-576 is also likely O-GlcNacylated (Fig. 8 and data not shown), indicating that the PEST region could be multiply modified. O-GlcNacylation within the PEST region of F domain may play an important role in regulating the breakdown of ER proteins. Katzenellenbogen and colleagues (73) have deleted the last 42 amino acids (F domain) of the hER and found that the deletion of F domain did not affect transactivation ability, ligand binding affinity, or the phosphorylation pattern of the receptor and suggested that F domain was not involved in the relatively rapid breakdown of ER protein. However, the glycosylation state of domain F in human ER has not been characterized.

![Diagram](attachment:fig8.png)

**FIG. 8.** Schematic illustration of the O-GlcNac Sites in a PEST region of the carboxyl-terminal F domain in mouse ER. Estrogen receptors have been defined to six distinct functional domains A–F. The evolutionary conserved regions C and E contain the DNA and hormone-binding domains, respectively. Two transcription activation regions are located in the region A/B (TAF-1) and in the hormone binding domain (TAF-2) of the ER. The PEST sequence within F domain in mER is 559RGVVPPEPSQTQLATTSSATSANS. A major O-GlcNac site is located at Thr-575. O-Glycosylation is also likely located at Ser-576 (indicated as *). NH2 terminus and carboxyl terminus are represented as N′ and C′.
