Identification by MALDI-TOF Mass Spectrometry of 17α-Bromoacetamidopropylestradiol Covalent Attachment Sites on Estrogen Receptor α†

Hélène Mattras,‡ Sigrid Aliau,‡ Eric Richard,‡ Jean-Claude Bonnafous,‡ Patrick Jouin,§ and Jean-Louis Borgna*,‡

INSERM U 439, 70 rue de Navacelles, and UPR 9023 CNRS, 141 rue de la Cardonille, 34090 Montpellier, France

Received August 2, 2002; Revised Manuscript Received October 7, 2002

ABSTRACT: Mass spectrometry was used to identify the sites of covalent attachment of [^{14}C]-17α-bromoacetamidopropylestradiol ([^{14}C]17BAPE$_2$, an estradiol agonist) to the ligand-binding domain (LBD) of mouse estrogen receptor α (ERα). A glutathione *S*-transferase (GST)−LBD chimera protein was overexpressed in *Escherichia coli*, using a vector encoding GST fused with a C-terminal portion of mouse ERα (Ser^{313}−Ile^{599}), via a sequence enclosing a thrombin cleavage site (located 14 amino acids ahead of Ser313). [^{14}C]17BAPE$_2$ covalent labeling experiments were carried out on the GST−LBD chimera immobilized on glutathione−Sepharose. After thrombin cleavage of the chimeric LBD, two major [^{14}C]-17BAPE$_2$-labeled species of 34 (∼75%) and 30 kDa (∼25%) were detected by SDS−PAGE and autoradiography. Their identity was assessed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS): two main signals were consistent with the mass of the full-length (Ser^{313}−Ile^{599}) and truncated LBD (Ser^{313}−Ala^{573}), both comprising the extra 14 N-terminal amino acids and covalently bound [^{14}C]17BAPE$_2$ (via HBr elimination). A purified ^{14}C-labeled LBD preparation was trypsinized to identify the covalent attachment sites of 17BAPE$_2$. HPLC of tryptic fragments only revealed two discrete and practically equivalent radioactive fractions. MALDI-TOF MS analysis of these two fractions showed only two signals which exactly matched the molecular masses of the [^{14}C]17BAPE$_2$-alkylated Cys^{534}Lys^{535} and Cys^{421}−Arg^{438} peptides, respectively. Hydrolysis of the second ^{14}C-labeled fraction by *Staphylococcus aureus* V8 Glu-C endoproteinase generated signals typical of alkylated the Cys^{421}−Glu^{423} tripeptide. We concluded that Cys421 and Cys534 were equivalent alternative covalent attachment sites of 17BAPE$_2$ on the LBD. These biochemical data were interpreted using the crystallographic structures of estradiol−LBD and raloxifene− or 4-hydroxytamoxifen−LBD complexes. The covalent attachment to Cys421, Cys534, or both could be interpreted according to the starting structure. Various hypotheses based on the biochemical results and molecular modeling simulations are discussed, with the likely involvement of dynamic interconversion between multiple conformational states of the LBD−17BAPE$_2$ complex.

In mammals, two estrogen receptors (ERα[1] and ERβ) (*1*−*4*) mediate the effects of estrogens and antiestrogens in estrogen target cells. These ERs belong to the nuclear receptor superfamily (*5*, *6*), the members of which display a modular structure, usually involving six domains (A−F). Domains C and E, which constitute the DNA-binding domain and the ligand-binding domain (LBD), respectively, are the most conserved domains in nuclear receptors; e.g., human ERα and ERβ are 96 and 59% identical in amino acid sequence in their DNA-binding domains and LBDs, respectively. Since the mid-1990s, a number of X-ray analyses of crystallized nuclear receptor LBDs, either unliganded or bound to agonist or antagonist ligands, have considerably improved our knowledge in this field (*7*). Several of these studies were performed with ERα (*8*−*12*) or ERβ (*13*, *14*) LBDs. Crystal structures of ERα LBD−ligand complexes show the characteristic, predominantly α-helical fold observed with other crystallized nuclear receptor LBDs (*15*). According to the authors, the ERα LBD which includes 12 α-helices [the ERα helix numbering of Brzozowski et al. (*8*) will be used] and two β-strands is folded into a three-layered antiparallel α-helical sandwich. Agonists [estradiol (E_2) and diethylstilbestrol (DES)] and antagonists [4-hy-

† This work was supported by the "Institut National de la Santé et de la Recherche Médicale" and the "Association pour la Recherche sur le Cancer".

* To whom correspondence should be addressed. Telephone: 33 4 67 04 37 14. Fax: 33 4 67 04 37 15. E-mail: Borgna@u439.montp.inserm.fr.

‡ INSERM U 439.

§ UPR 9023 CNRS.

[1] Abbreviations: ER, estrogen receptor; LBD, ligand-binding domain; LBP, ligand-binding pocket; GST, glutathione *S*-transferase; E_2, estradiol; DES, diethylstilbestrol; OHT, 4-hydroxytamoxifen; Ral, raloxifene; 17BAPE$_2$, 17α-bromoacetamidopropylestradiol; SDS, sodium dodecyl sulfate; TFA, trifluoroacetic acid; CHCA, α-cyano-4-hydroxycinnamic acid; SA, 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid); PBS, phosphate-buffered saline; T_{20}, 20 mM Tris-HCl buffer; V8 proteinase (or V8), *Staphylococcus aureus* V8 Glu-C endoproteinase; PAGE, polyacrylamide gel electrophoresis; HPLC, high-pressure liquid chromatography; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; MS, mass spectrometry. Θ symbolizes an undefined atom or group of atoms displaying a monoisotopic mass very close to 16.

10.1021/bi0205092 CCC: $22.00
Published on Web 12/04/2002

droxytamoxifen (OHT) and raloxifene (Ral)] bind at the same site within the LBD core; the ligand-binding pocket (LBP) is predominantly formed by hydrophobic residues located in six portions of the LBD sequence. Many structural features are common to estrogen and antiestrogen complexes, but a major established difference between structures of crystallized ERα LBD−estrogen and LBD−antiestrogen complexes was found to be related to the orientation of helix 12, which harbors the AF-2 activation domain (*16*, *17*) that is involved in coactivator recruitment (*18*). The orientation of helix 12 in an ER−ligand complex was therefore thought to be responsible for the agonist or antagonist activity of the ligand (*8*). However, this dual positioning of helix 12, according to the type of ligand bound to the LBD, did not seem to be a common phenomenon since helix 12 was found to have an "antiestrogen type" orientation in agonist-bound ERβ LBD (*13*) or mutated ERα LBD (*11*) crystals. These results underscored the high degree of positional variability of this helix in crystallized LBD forms.

The crystallization approach cannot account for the possible occurrence of equilibria between multiple conformational states for these complexes in aqueous medium. Moreover, the stacking of LBD monomers in the crystalline network may change the conformation or even the type of peripheral amino acid residues occurring in solute complexes. It is important to determine the potential conformational states of these complexes according to hormone or antihormone binding. Some of them, even if they are a minority, could especially be involved in coactivator or corepressor recruitment and could thus account for ligand-induced modulation of ER transcriptional activity. Therefore, noncrystallization approaches, i.e., using fluorescent, cross-linking, or specific amino acid reagents or affinity labeling agents to probe the structure of solute ER−ligand complexes, could enhance the overall understanding of the molecular action of estrogens and antiestrogens, and help in the rational design of new selective ER modulators (*19*).

In previous studies, we characterized lamb and human ERα affinity labeling by three closely related estrogenic ligands, derived from E_2 by introducing an 17α-bromoacetamidoalkyl substituent (the alkyl group being methyl, ethyl, or propyl). First using an SH-specific reagent, we showed that cysteinyl residues of lamb ERα were involved in covalent attachment of these electrophilic ligands (*20*). Then the inability of electrophiles to alkylate the Cys417Ala/Cys530Ala double-mutant human ERα[2] strongly suggested that Cys417 and Cys530 constituted the covalent attachment sites of these electrophiles in human ERα (*21*). However, this indirect evidence did not allow us to determine the relative involvement of the two cysteines in electrophile covalent attachment processes.

In the study presented here, using mass spectrometry (MS) analysis, we attempted to directly identify the covalent attachment sites within the mouse ERα LBD of 17α-bromoacetamidopropylE_2 (17BAPE$_2$, Figure 1), one of the previously studied estrogenic ER affinity labels, which

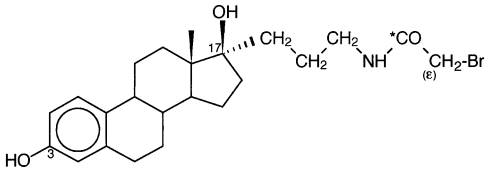

FIGURE 1: Structure of [^{14}C]17BAPE$_2$. The ER affinity labeling agent is derived from E_2 by 17α-substitution. It possesses an electrophilic carbon (terminal bromine-bearing carbon, ε-carbon) and, to facilitate the characterization of the affinity-labeled mouse ERα LBD and fragments, a 88% ^{14}C-enriched carbon (denoted with an asterisk).

displays high affinity for mammal ERαs (K_a estimated to be in the range of 10^8-10^9 M^{-1}, unpublished results). We first report the production in *Escherichia coli* and purification of a chimera protein made of glutathione *S*-transferase (GST) and a C-terminal portion of mouse ERα (Ser^{313}−Ile^{599} sequence, corresponding to the LBD and the adjacent C-terminal region F, and designated LBD). We then describe the affinity labeling of LBD with a ^{14}C-labeled form of the affinity probe, and the characterization of [^{14}C]17BAPE$_2$-labeled LBD. We next report, by means of matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) MS analysis, the identification of covalently modified LBD tryptic peptides, and labeled residues within these peptides. Finally, using available crystalline models of the ERα LBD bound to estrogen or antiestrogen, we try to interpret by molecular modeling the biochemical results obtained.

EXPERIMENTAL PROCEDURES

Materials. 1-Isopropylthio-β-D-galactopyranoside, protease inhibitor cocktail for bacterial cell extracts [containing 4-(2-aminoethyl)benzenesulfonyl fluoride, pepstatin A, *trans*-epoxy succinyl-L-leucylamido(4-guanidino)butane, bestatin, and EDTA], and *Staphylococcus aureus* V8 Glu-C endoproteinase (V8) were from Sigma-Aldrich. Glutathione−Sepharose 4B and thrombin were purchased from Amersham Biosciences. Trypsin (sequencing grade) was from Roche, and the synthetic CysValGluGlyMetValGlu heptapeptide was purchased from Neosystem. Solvents and reagents used for HPLC and MALDI-TOF MS analyses were HPLC grade (for HPLC, Milli-Q grade water was passed through a C18 preparative column and then sterilized). Matrices for MS, α-cyano-4-hydroxycinnamic acid (CHCA) and 3,5-dimethoxy-4-hydroxycinnamic acid [sinapinic acid (SA)], were purchased from Sigma-Aldrich. [6,7-^{3}H]Estradiol ([^{3}H]-E_2, specific activity of 1.6 PBq/mol, radiochemical purity of >98%) was purchased from New England Nuclear, and bromo[1-^{14}C]acetic acid (specific activity of 2.0 TBq/mol, 88% ^{14}C enrichment, radiochemical purity of >98%) from Amersham Biosciences. [^{3}H]E_2 and synthesized [NHCO-^{14}C]-17α-bromoacetamidopropylE_2 ([^{14}C]17BAPE$_2$) were solubilized in absolute ethanol. Solutions were stored at −20 °C in the dark. The purity of solubilized compounds was checked before use by thin-layer chromatography.

GST−LBD Expression Vector. A pGEX-2TK vector encoding glutathione *S*-transferase (GST) fused with the mouse ERα LBD (Ser^{313}−Ile^{599}), via a sequence enclosing a thrombin cleavage site located 14 amino acids ahead of Ser313 (*22*), was provided by V. Cavaillès (INSERM U 540).

Protein Assay and Radioactivity Determinations. Protein concentrations were measured colorimetrically by the method

[2] As mouse ERα possesses four additional residues as compared to human ERα, the residue numbers are different for the two ERs. The mouse numbering is used in the Results and related Figures 2 and 8; the human numbering is used in Experimental Procedures and the Discussion which deal with crystal models of the human ERα LBD.

of Bradford (*23*) using a dye reagent concentrate (Bio-Rad Laboratories) with bovine serum albumin as a standard. The radioactivity of the various samples (10−200 μL) was counted in 4 mL of Emulsifier Safe (Packard) using an LS 1801 Beckman scintillation counter.

SDS−PAGE Analyses and Autoradiography. Electrophoreses were performed using a Mini Protean II Bio-Rad unit with 1 mm gel slabs containing 12% (w/v) acrylamide in the separation gel and 5% acrylamide in the stacking gel, in 0.1% (w/v) SDS, 25 mM Tris-HCl buffer (pH 8.5). Gels were stained with GelCode Blue Stain Reagent (Pierce). Prestained protein molecular weight markers, a broad range from 6.5 to 175 kDa (BioLabs), were used as a standard. ^{14}C-labeled species were analyzed on dried gels using a Fujix BAS 1000 apparatus.

Synthesis and Characterization of [^{14}C]17BAPE$_2$. The procedure described for the preparation of unlabeled 17BAPE$_2$ (*20*) was used to obtain [NHCO-^{14}C]17BAPE$_2$. Briefly, 17α-aminopropylE$_2$, with the phenolic function protected in the form of *tert*-butyldimethylsilyl ether, was condensed with bromo[1-^{14}C]acetic acid in the presence of 1-ethyl-3-[3-(dimethylamino)propyl] carbodiimide hydrochloride. The silylated protection group was then removed by exposure of the compound to tetrabutylammonium fluoride. The [^{14}C]-17BAPE$_2$ that was obtained was purified on silica gel. The compound purity, assessed by thin-layer chromatography, was >95%.

Expression and Extraction of the GST−LBD Chimera. A 10 mL culture of *E. coli* BL21 cells transformed with the pGEX-2TK plasmid encoding the GST−LBD chimera was grown overnight at 37 °C in LB^A (LB medium with 0.1 mg/mL ampicillin). The cell suspension (OD_{550} = 1.2−1.5) was incubated in 450 mL of LB^A at 37 °C until the OD_{600} had reached 0.6. GST−LBD expression was then induced with 1 mM 1-isopropylthio-β-D-galactopyranoside for 2 h at 37 °C. Cells were then pelleted by centrifugation for 30 min at 3500*g* (4 °C). Bacterial pellets were resuspended in a minimal volume of fresh LB, and centrifuged once more. The final pellet was suspended in extraction buffer [20 mM Tris-HCl (T_{20}) (pH 8) containing 100 mM NaCl, 1 mM EDTA, 10 mM dithiothreitol, and protease inhibitor cocktail] at a ratio of 2 mL of buffer to 1 mL of pellet, and then aliquoted (1.5 mL fractions). Aliquots were centrifuged for 15 min at 3500*g* (4 °C); supernatants were discarded and pellets (0.5 mL) snap-frozen and then stored at −70 °C.

For fusion protein extraction, thawed aliquots were slowly resuspended in ice-cooled extraction buffer [1/1 (v/v)]. Nonidet P-40 (final concentration of 0.1%) was added, and then the cell suspension was incubated for 30 min at 0 °C. The cells were ruptured by sonication at 0 °C, using a Bioblock-Scientific Vibra-cell 72405 instrument (six pulses for 6 s, amplitude of 60), and sonication was repeated 15 min later. Cell breakage was controlled by microscopy (Turk reagent), and the sonicate was centrifuged for 10 min at 20000*g* (4 °C).

Binding Assay. Extracts from induced *E. coli* cells were diluted 200-fold with T_{20} (pH 7.5). The protein concentration of diluted extracts was adjusted to 2 mg/mL with bovine serum albumin. Aliquots were incubated for 15 h at 0 °C with 10 nM [^{3}H]E$_2$ in the absence (to measure the level of total binding) and presence (to measure the level of nonspecific binding) of 1 μM unlabeled E$_2$. Aliquots were treated with an equal volume of a charcoal suspension [1% charcoal, 0.1% dextran T_{70} in T_{20} (pH 7.5)] for 15 min at 0 °C; charcoal was pelleted by centrifugation, and concentrations of bound [^{3}H]E$_2$ in supernatants were determined by radioactivity counting.

Purification and Affinity Labeling of the LBD. Bacterial lysate (1.5 mL) was added to 100 μL of glutathione−Sepharose 4B [equilibrated with T_{20} (pH 8.0)] in an ice-cooled Bio-Spin chromatography column (Bio-Rad Laboratories). The mixture was rotated for 15 h at 4 °C, and then the column was drained and carefully washed with phosphate-buffered saline (PBS), equilibrated to pH 8.5, to remove nonspecifically bound proteins. An ∼4-fold excess of [^{14}C]-17BAPE$_2$ (final concentration of 10 μM), relative to the glutathione−Sepharose-bound GST−LBD chimera, was added to the matrix suspended in 1 mL of ice-cooled PBS (pH 8.5) containing 10% dimethylformamide. The mixture was rotated gently for 2 h at 4 °C, washed five times with PBS, and then resuspended in 1 mL of ice-cooled PBS (pH 7.3). Thrombin (1 NIH for a 1 mL mixture) was added to the gel suspension; the mixture was rotated for 15 h at 4 °C and then incubated for 30 min at 4 °C with glutathione (30 mM, pH readjusted to 7.3). The gel was drained and washed twice with 50 μL of PBS (pH 7.3), and the three eluates were pooled. Aliquots of the various samples were analyzed for ^{14}C content, [^{3}H]E$_2$ binding capacity, and protein content (SDS−PAGE).

CysValGluGlyMetValGlu Alkylation with [^{14}C]17BAPE$_2$. Sixty-five microliters of an ethanol solution of [^{14}C]17BAPE$_2$ (1 mM) was added to 195 μL of a solution of the synthetic CysValGluGlyMetValGlu heptapeptide (0.67 mM) in 0.1 M $NaHCO_3$. The mixture was incubated at 20 °C, and equal amounts of heptapeptide were added at 1 and 2 h. After incubation for 3 h, the mixture was submitted to HPLC.

HPLCs. HPLC purifications of [^{14}C]17BAPE$_2$-labeled or unlabeled forms of the GST−LBD chimera and the LBD were performed on a Vydac C4 reversed-phase column (3.2 mm × 250 mm, 5 μm particles, 300 Å pore size) at a flow rate of 500 μL/min, using a Waters gradient system (two 510 pumps, a gradient controller, and a 486 UV detector) equipped with a Shimadzu C-R6A integrator and an Isco Retriever II collector. For compound elution, linear gradients were generated from 10 to 95% CH_3CN over the course of 60−120 min; binary components for these gradients included solvent A [90/10 H_2O/CH_3CN mixture containing 0.05% trifluoroacetic acid (TFA)] and solvent B (5/95 H_2O/CH_3CN mixture containing 0.05% TFA). UV absorption (220 nm) was recorded; 4 min fractions were collected. LBD tryptic fragments were analyzed and purified under the same conditions with a Vydac C4 narrow-bore column (2.1 mm × 250 mm) at a flow rate of 200 μL/min. HPLC purifications of crude and [^{14}C]17BAPE$_2$-labeled heptapeptide and its V8 proteolytic fragments were performed with an Alltima C18 column (4.6 mm × 250 mm, 5 μm particles) at a flow rate of 1 mL/min. Linear gradients from 10 to 95% CH_3CN were generated with solvent A and B over the course of 30−60 min.

Trypsin and Endoproteinase V8 Proteolyses. The radioactive fractions obtained by HPLC purification of the affinity-labeled LBD (or tryptic fragments) or labeled heptapeptide were partially dried to evaporate CH_3CN, and then their pH was adjusted to the appropriate value and the proteolytic

enzyme added. For trypsin proteolysis of the LBD [T_{20} (pH 7.5), enzyme-to-substrate ratio of 1/40], the incubation was carried out for 2 h at 37 °C; for V8 endoproteinase proteolysis of LBD tryptic fragments or labeled heptapeptide [100 mM phosphate buffer (pH 7.8), enzyme-to-substrate ratio of 1/33], the incubation was performed for 15 h at 25 °C.

ZipTip Purification of Affinity-Labeled V8 Proteolytic LBD Fragments. A $ZipTip_{\mu-C18}$ microcolumn (Millipore) was equilibrated with a 50/50 H_2O/CH_3CN mixture. After being washed (three times) with 0.1% TFA, the column was loaded with the V8-proteolyzed sample. The ZipTip was washed (three times) with 0.1% TFA, and then bound peptides were eluted with 4 μL of a 20/80 H_2O/CH_3CN mixture containing 0.02% TFA. Samples were air-dried and then solubilized in 3 μL of 0.1% TFA.

Mass Spectrometry. MALDI mass spectra were obtained on a Bruker BiflexIII time-of-flight mass spectrometer (Bruker-Franzen, Bremen, Germany) equipped with a SCOUT source and Xmass software for data collection and analysis. The matrices were prepared either from a saturated solution of CHCA (analyses of peptides and LBD digests) or from a 4.5 μM solution of SA (LBD analyses) in a 1/1 (v/v) 0.1% TFA/CH_3CN mixture. A 1 μL aliquot of matrix solution was mixed with 1 μL of sample, and 0.4 μL of the mixture was spotted on the probe and allowed to dry at ambient temperature. Desorption and ionization were achieved with an N_2 laser (337 nm, 3 ns pulse width, 3 Hz repetition rate), followed by subsequent gridless delayed extraction. The instrument was used in the positive reflectron mode for synthetic peptides or digests, or linear mode for proteins, with an acceleration voltage of 19 kV. Mass spectra were acquired as a sum of ion signals generated by irradiation of the target with 50−200 laser pulses. Quadratic external calibration was achieved using a mixture of standard peptides containing angiotensin I [$[M + H]^+$, *m/z* (monoisotopic) 1296.69], bombesin ($[M + H]^+$, *m/z* 1619.82), ACTH 1−17 ($[M + H]^+$, *m/z* 2093.09), ACTH 18−36 ($[M + H]^+$, *m/z* 2465.20), and somatostatin 28 ($[M + H]^+$, *m/z* 3147.47). The Swiss-Prot protein data bank was consulted to characterize signals found in the spectra of the mouse ERα LBD and its proteolytic fragments.

Molecular Modeling Studies. Energy minimizations of ERα LBD models were performed using InsightII/Discover software from MSI, and CFF91 as the force field. The starting X-ray crystallographic structures of the human ERα LBD bound to E_2 (PDB entry 2ERE, monomer A), OHT (PDB entry 3ERT, monomer A), or Ral (PDB entry 1ERR, monomer A) (*8*, *10*) were lacking several poorly resolved amino acid residues, and some loop sequences were not specified for LBD−E_2 (residues 330−337 and 461−465) and LBD−Ral (residues 460−469 and 529−534) complexes. Lacking amino acid side chains were introduced in the complexes. All water molecules from crystallographic structures were deleted, with the exception of the one involved in hydrogen bonding of the phenolic OH (E_2 and OHT) or benzothiophene OH (Ral). $17BAPE_2$ was docked into LBD structures derived from crystallized estrogen− and antiestrogen−LBD complexes to account for the covalent attachment of $17BAPE_2$ to Cys421 and Cys534 (the mouse homologues of human Cys417 and Cys530, respectively), while firmly bound to the LBD.

(1) Docking of $17BAPE_2$ into the LBD Structure Derived from Human LBD−E_2 Crystals. $17BAPE_2$ was superimposed over LBD-bound E_2 by aligning the two A rings. This resulted in almost merged steroid nuclei, with very close locations of the two 17β-OH functions. The flexibility of the $17BAPE_2$ 17α-substituent was then used to bring the electrophilic carbon (ε-carbon) into the proximity of the S atom of Cys417. Then steric clashes between $17BAPE_2$ and the LBD were manually suppressed by appropriate rotations of some amino acid side chains and single chemical bonds of the 17α-substituent.

(2) Docking of $17BAPE_2$ into an LBD Structure Derived from Human LBD−OHT Crystals. This was performed by superimposing the $17BAPE_2$ A ring over the phenol ring of LBD-bound OHT. The His524 side chain was rotated so that the distance between δ-N of its imidazole ring and steroidal 17β-O would be suitable for N−OH hydrogen bonding (distance of ∼3 Å). The flexibility of the helix 11−helix 12 loop was then used to position Cys530 close to the $17BAPE_2$ ε-carbon; i.e., the helix 11−helix 12 loop was ruptured at the junction of amino acids 530 and 531; by appropriate rotations of single N−C and C−C bonds, the two open sequences were twisted to bring Cys530 closer to helix-3, and then the two sequences were reconnected. The 17α-substituent and (newly conformed) helix 11−helix 12 loop flexibilities were used to bring the ε-carbon and the S atom of Cys530 closer (distance of ∼3.5 Å); steric clashes between $17BAPE_2$ and LBD were suppressed, as described with the LBD−E_2 complex. OHT was then deleted, and a first minimization was carried out with fixed LBD α-carbons, except those of the helix 9−helix 10 and helix 11−helix 12 loops, and constrained interactions of (i) $17BAPE_2$ 3-O with Glu353 and Arg394 side chains and with the bound water molecule, (ii) $17BAPE_2$ 17β-O with the His524 imidazole ring, and (iii) the $17BAPE_2$ ε-carbon with the S atom of Cys530. A second round of minimization was performed as indicated above, except that the LBD−ligand interaction constraints were suppressed.

(3) Docking of $17BAPE_2$ into an LBD Structure Derived from Human LBD−Ral Crystals. Lacking sequences in the LBD−Ral complex, i.e., amino acids 460−469 (helix 9−helix 10 loop) and 529−534 (helix 11−helix 12 loop), were imported from the crystallized LBD−OHT complex (with a modified helix 11−helix 12 loop conformation, as described in the above paragraph). The modified complex was then minimized with fixed LBD α-carbons, except those of added loops, and constrained interactions of (i) the benzothiophene-linked O of Ral with Glu353 and Arg394 side chains and with the bound water molecule, (ii) the second phenolic O of Ral with the His524 imidazole ring, and (iii) the pyrimidyl N of Ral with the Asp351 side chain. $17BAPE_2$ was then superimposed over Ral to obtain optimal overlap of the two ligands, i.e., (i) matching of $17BAPE_2$ 3-O with the benzothiophene-linked O, with very close locations of the two phenol rings, and (ii) the tight proximity of $17BAPE_2$ 17β-O and the second phenolic O of Ral. The His524 side chain was rotated to allow hydrogen bonding between δ-N of its imidazole ring and $17BAPE_2$ 17β-OH. Flexibilities of the 17α-substituent and the helix 11−helix 12 loop were used to closely position the ε-carbon with the S atom of Cys530 or of Cys417 (distance of ∼3.5 Å). Steric clashes in the two alternative models of the LBD−$17BAPE_2$

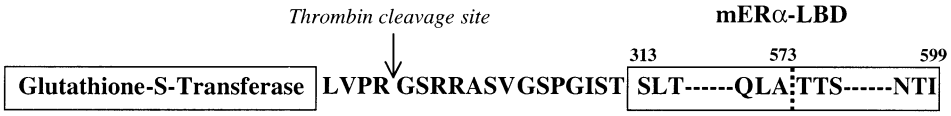

FIGURE 2: Structure of the GST−LBD chimera protein. The chimera encoded by the pGEX-2TK expression vector is schematically represented. Its LBD moiety constituted by the C-terminal half of mouse ERα (Ser^{313}−Ile^{599}) is linked to the GST moiety by an amino acid sequence enclosing a thrombin cleavage site. The major GST−LBD forms expressed in *E. coli* extracts were the full-length form and a C-terminally truncated form ending at Ala573 (cf. the text).

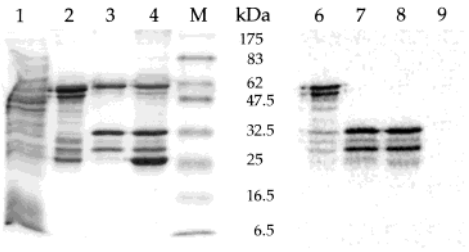

FIGURE 3: SDS−PAGE analysis of the mERα LBD. The cell extract (1.5 mL) from *E. coli*, induced to express the GST−LBD chimera, was rotated for 15 h at 4 °C with glutathione−Sepharose (100 μL). The whole gel was transferred into a column and washed with T_{20} (pH 8.0, 10 mL). Adsorbed species were incubated with or without 10 μM [^{14}C]17BAPE$_2$ (in the absence and presence of 50 μM E_2) for 2 h at 4 °C. Following gel washes, bound species were treated with glutathione, thrombin, or thrombin and glutathione, successively. Aliquots of bacterial extract (lane 1) and proteins released by the action of glutathione (lanes 2 and 6), thrombin (lanes 3 and 7), or thrombin and glutathione successive actions (lanes 4, 8, and 9) were subjected to SDS−PAGE. Proteins from unlabeled samples were revealed with blue stain (lanes 1−4), whereas proteins from samples incubated with [^{14}C]17BAPE$_2$ were visualized by autoradiography (lanes 6−8, samples incubated in the absence of E_2; lane 9, sample incubated in the presence of E_2). Standard molecular mass markers revealed by blue stain are in lane M.

complex were suppressed, as previously described, and then Ral was deleted. Two rounds of minimization, one with and one without constraints, were carried out as described for the minimization with the LBD structure derived from LBD−OHT crystals.

RESULTS

Expression, Purification, and Affinity Labeling of the Mouse ERα LBD. Our goal of identifying the covalent attachment sites of 17BAPE$_2$ (Figure 1) in the ERα LBD by MS analysis required access to large amounts (nanomole range, ~10^{-4} g) of the purified LBD. Expression in *E. coli* of a GST−mouse LBD chimera protein (enclosing a thrombin cleavage site located 14 amino acids ahead of LBD, cf. Figure 2), using a previously described recombinant vector (*21*), allowed us to fulfill these conditions, since (i) binding experiments performed with [^{3}H]E_2 revealed the presence of micromolar concentrations of specific E_2-binding species in soluble bacterial extracts and (ii) incubation with a glutathione−Sepharose absorbent selectively removed these E_2-binding species from extracts (not shown). SDS−PAGE analysis (Figure 3) of crude extracts (lane 1) and of proteins released by glutathione from glutathione−Sepharose (lane 2) showed that four main species (apparent molecular masses of 57, 53, 51, and 28 kDa) were selectively enriched in the glutathione eluate. To differentiate GST−LBD species from other proteins, glutathione−Sepharose-bound proteins were submitted to thrombin action alone or followed by that of glutathione, and then the released proteins were analyzed (lanes 3 and 4, respectively). In both cases, the 57 kDa species was still present, with two new major species (apparent molecular masses of 34 and 30 kDa), whereas the 53 and 51 kDa species were no longer detected. The main difference between the patterns in lanes 3 and 4 was the presence of a prominent band at 28 kDa in the thrombin + glutathione eluate which was absent in the thrombin eluate.

To affinity label the mouse LBD, the glutathione−Sepharose gel loaded with cell extracts was incubated with 10 μM [^{14}C]17BAPE$_2$, a concentration ~4-fold higher than that of E_2 specific-binding sites in adsorbed extracts. Then bound species were released as described above. Following SDS−PAGE, labeled species were revealed by autoradiography. Only two main signals were obtained from the glutathione eluate (lane 6), which precisely corresponded to the 53 and 51 kDa bands in lane 2. The thrombin and thrombin + glutathione eluates exhibited identical autoradiography patterns (lanes 7 and 8, respectively) with two major signals, which accurately corresponded to the 34 and 30 kDa bands in lanes 3 and 4. All of these ^{14}C signals were due to E_2 specific-binding species, since no signal was observed (lane 9) when the labeling step was carried out in the presence of a 5-fold excess of E_2 (relative to the [^{14}C]-17BAPE$_2$ concentration). It is noteworthy that (i) in lanes 6−8 the intensity ratio of signals for the heavy and light species was ~3 and (ii) no ^{14}C signal was observed at levels corresponding to the 57 and 28 kDa bands. The [^{14}C]-17BAPE$_2$ concentration that was used appeared to be quasi-optimal for species labeling, since in another experiment only slight increases (<50%) in signal intensities were obtained when the affinity probe concentration was increased from 6 to 24 μM (not shown).

All of the above results suggested that (i) the 53 and 51 kDa species were two different forms (present at a 3/1 ratio in the glutathione eluate) of the GST−LBD chimera, (ii) the 34 and 30 kDa species were the two corresponding LBD forms, and the 28 kDa species was the GST moiety, all released from the 53 and 51 kDa species by thrombin, (iii) only the LBD moieties of GST−LBD chimeras were specifically alkylated by [^{14}C]17BAPE$_2$, and (iv) the unidentified 57 kDa species may be related to the bacterial chaperonin dnaK involved in protein degradation (*24*) and which seemed to interact with the LBD.

Note that, besides the two major described forms for both GST−LBD and LBD species, two minor forms accounting for less than 10% of labeled species were present in the eluates (Figure 3, lanes 6−8); very weak LBD-like forms were present in the glutathione eluate (lane 6), which could have resulted from partial hydrolysis of GST−LBD forms by a bacterial protease. In some experiments, we observed that the thrombin-induced release of the 30 and 34 kDa species from glutathione−Sepharose was markedly improved when glutathione was subsequently added to the

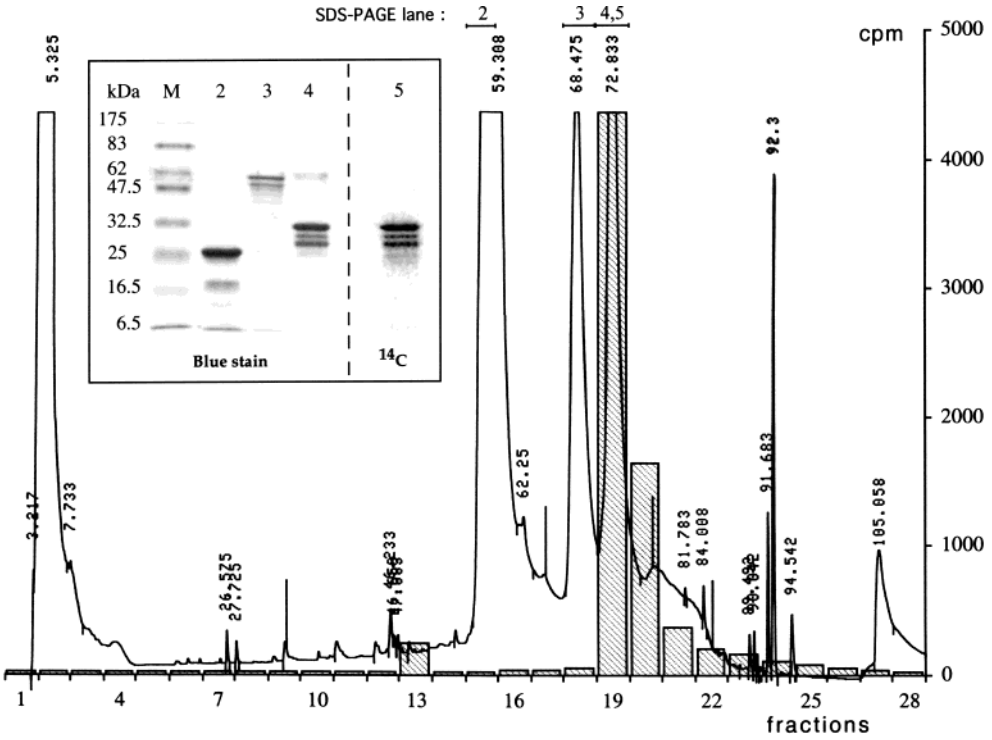

FIGURE 4: HPLC of the [^{14}C]17BAPE$_2$-labeled LBD. *E. coli* cell extracts were incubated with glutathione−Sepharose, and then following gel washes, bound species were exposed to [^{14}C]17BAPE$_2$ in situ. The species released by the action of thrombin and then glutathione were submitted to reverse-phase HPLC. Column elution was performed with a linear CH_3CN gradient. The curve represents the recorded 220 nm absorption according to the elution time, whereas hatched bars indicate the ^{14}C content of the various fractions (counts per minute per milliliter). The contents of major HPLC fractions (15, 18, and 19), analyzed by SDS−PAGE, are shown in the inset: blue staining (lanes 1−4) and autoradiography (lane 5) reveal the unlabeled and labeled species in fractions 15 (lane 2), 18 (lane 3), and 19 (lanes 4 and 5). Standard molecular mass markers are in lane M.

medium. This positive effect could be interpreted in two ways: (i) GST−LBD molecules released by glutathione displayed better accessibility to thrombin, or (ii) glutathione reduced S−S bonds formed between LBD SH groups and those of glutathione−Sepharose or those of glutathione−Sepharose-bound proteins. Regardless of the explanation, in practice we routinely used the successive actions of thrombin and glutathione for the full release of bound species from the absorbent.

Species released from glutathione−Sepharose, following in situ exposure of bound material to [^{14}C]17BAPE$_2$, were submitted to HPLC (Figure 4). Three main components were obtained: the first two (fractions 15, 16, and 18) only included unlabeled species, whereas the third contained practically all of the ^{14}C-labeled species (fractions 19 and 20). SDS−PAGE analysis of these fractions (Figure 4) showed that the first and second components (lanes 2 and 3) mainly contained putative GST (28 kDa) and dnaK (57 kDa) species, respectively, and the third one (lanes 4 and 5) included the ^{14}C-labeled 34 and 30 kDa species together with the two minor companion forms. These results indicated that HPLC was efficient for purifying the LBD forms. To more accurately characterize the two main ^{14}C-labeled LBD forms, whole radioactive fractions were submitted to MALDI-TOF MS analysis (Figure 5). In close accordance with the results of SDS−PAGE blue staining and autoradiography analyses, MS analysis showed only two signals centered at *m*/*z* 34 089 and 31 297 with relative intensities of 70 and 30%, respectively. The *m*/*z* 34 000 component was quite likely due to the $[M + H]^+$ form of the full-length LBD with its 14 extra N-terminal amino acids (Figure 2) and, via dehydrobromination, covalently bound [^{14}C]17BAPE$_2$ (average mass increment of 371.50 Da). The *m*/*z* 31 000 component appeared to correspond to the same species truncated at Ala573. These assumptions were based on acute agreements between (i) the calculated average mass for such species ([M + H], 34198.79 and 31403.78 Da, respectively) and the two experimental *m*/*z* values (deviation of 0.3%) and (ii) the

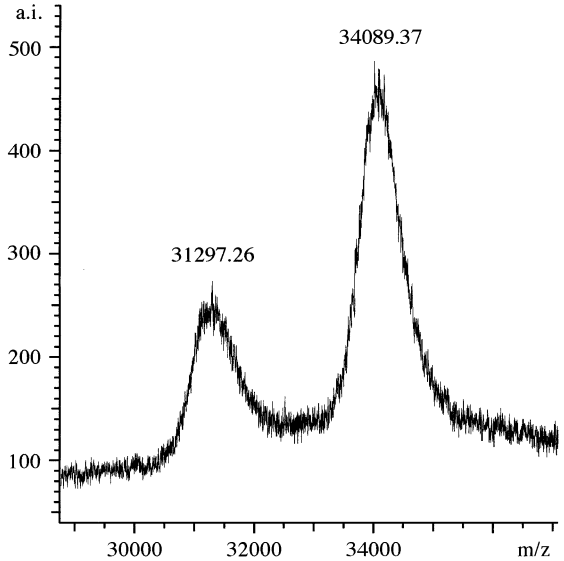

FIGURE 5: MALDI-TOF MS analysis of the [^{14}C]17BAPE$_2$-labeled LBD. The labeled LBD released from glutathione−Sepharose was purified by HPLC, as shown in Figure 4. The main radioactive fractions (19 and 20) were pooled, concentrated, and then analyzed by MS (SA matrix and positive linear mode).

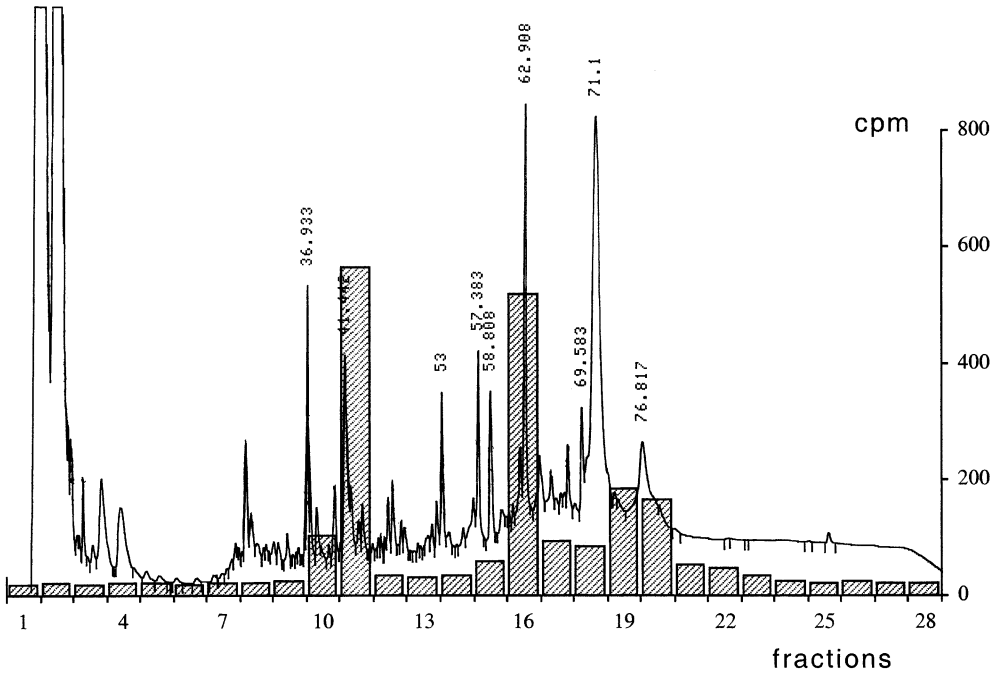

FIGURE 6: HPLC of [^{14}C]17BAPE$_2$-labeled LBD tryptic fragments. The labeled LBD was purified, as shown in Figure 4. The main radioactive fractions (19 and 20) were concentrated, treated with trypsin, and then submitted to reverse-phase HPLC. Column elution was performed with a linear CH_3CN gradient. The curve represents the recorded 220 nm absorption according to the elution time, whereas cross-hatched bars indicate the ^{14}C content of the various fractions (counts per minute per milliliter).

average mass increment ($\Delta M = 2795.01$ Da), corresponding to the postulated lacking Thr^{574}−Ile^{599} (LBD C-terminal) sequence (Figure 2) in the second species and the difference between the two experimental values ($\Delta m/z$ 2792.1, deviation of 0.1%). These interpretations were subsequently confirmed when tryptic fragments of the labeled LBD were analyzed by MALDI-TOF MS: two signals at *m*/*z* 4279.10 and 1483.55 were observed, which accurately corresponded to $[M + H]^+$ ions of full-length Met^{560}−Ile^{599} and truncated Met^{560}−Ala^{573} tryptic peptides. With regard to the truncated form, a similar form of the human ERα LBD, ending at Ala569, the human homologue of Ala573 of mouse ERα, was previously identified in *E. coli* expressing the human ERα LBD (*25*); this truncated form accounted for 23% of the different LBD forms.

Identification of Covalent Attachment Sites of [^{14}C]-17BAPE$_2$ on the Mouse ERα LBD. To identify 17BAPE$_2$-alkylated sites on the LBD, the HPLC-purified labeled LBD preparation was treated with trypsin, and then submitted to HPLC (Figure 6). Numerous species revealed by 220 nm absorption were present in the HPLC fractions, whereas only two nearly equivalent radioactive species were found (fractions 11 and 16, respectively), together with the nonproteolyzed LBD (fractions 19 and 20). MALDI-TOF MS analysis of the various fractions (example shown in Figure 7) allowed the identification of 23 signals corresponding to pseudomolecular $[M + H]^+$ ions of different LBD tryptic peptides (Figure 8), encompassing 97% of the LBD sequence.

Analyses of the two radioactive fractions gave few signals with, in each case, only one signal compatible with the combined masses of an LBD tryptic fragment and a dehydrobrominated form of [^{14}C]17BAPE$_2$ (monoisotopic mass increment of 371.23 Da). The fine structure of each of the two signals (Figure 9) was constituted by two series of well-resolved peaks, due to the occurrence of ^{12}C, ^{13}C, and ^{14}C isotopes in the alkylated species. In the major series, the peaks at *m*/*z* 621.47 and 2371.90 matched the $[M + H]^+$ form of alkylated LBD tryptic peptides (Figure 8) Cys^{534}Lys^{535} (calcd value of 621.36 Da, deviation of 0.02%) and Cys^{421}−Arg^{438} (calcd value of 2371.18 Da, deviation of 0.01%), comprising only ^{12}C isotopes and the ^{14}C isotope of 17BAPE$_2$; the peaks at $\Delta m/z$ +1, +2, +3 ... corresponded to alkylated forms involving one, two, three ... ^{13}C atoms, respectively. The minor series of peaks located at $\Delta m/z$ −2 from the major series simply reflected the presence of 12% [NH-^{12}CO] in [NH-^{14}CO]17BAPE$_2$ (cf. the legend of Figure 9). No signal was found that corresponded to the mass of any LBD tryptic peptide increased by one of the following values: 60.01 Da (addition of dehydrobrominated $BrCH_2$-$^{14}CO_2H$, due to hydrolysis of the amide bond in 17BAPE$_2$, following LBD alkylation), 357.22 Da (elimination of CH_3-Br following LBD alkylation by [^{14}C]17BAPE$_2$ at the S atom of methionine), or 45.99 Da (LBD alkylation at the S atom of methionine, followed by hydrolysis of the amide bond in 17BAPE$_2$). Then the results of HPLC and MALDI-TOF MS analyses strongly indicated that two alternative and quantitatively equivalent sites for covalent attachment of 17BAPE$_2$ were located on the Cys^{534}Lys^{535} and Cys^{421}−Arg^{438} LBD segments, respectively.

In the alkylated Cys^{534}Lys^{535} dipeptide, the covalent attachment site of 17BAPE$_2$ could be unambiguously assigned to the S atom of the Cys residue since, under the affinity labeling conditions used (pH 8.5), half of the Cys SH groups should be under their strongly nucleophilic, S^- form, whereas the Lys ϵ-NH_2 groups should be essentially under their unreactive NH_3^+ form.

In the Cys^{421}−Arg^{438} peptide, apart from the Cys residue, other amino acid residues possess a nucleophilic group, especially OH groups of the two Ser residues and the Thr residue. S atoms of the two Met residues are likely not

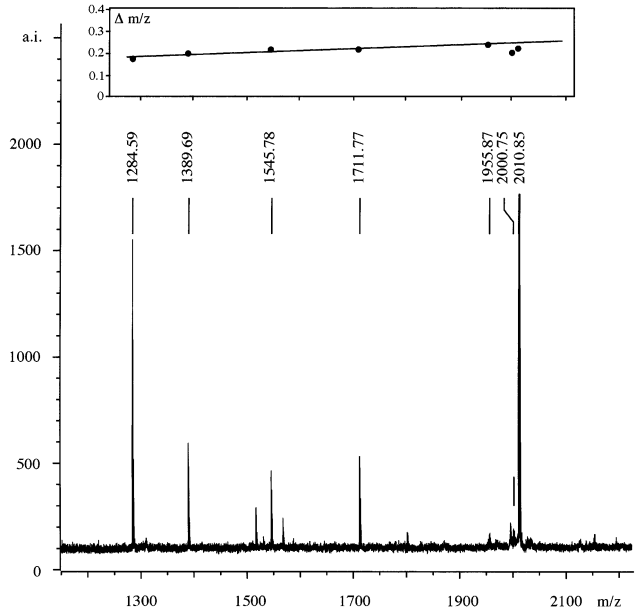

FIGURE 7: MALDI-TOF MS analysis of some unlabeled LBD tryptic fragments. Fractions 14−17 obtained in the HPLC separation of [^{14}C]17BAPE$_2$-labeled LBD tryptic fragments (as shown in Figure 6) were pooled and concentrated for MS analysis (CHCA matrix and positive reflectron mode). The various signals obtained in the range of *m/z* 1200−2200 are shown. *m/z* values which accurately match the monoisotopic mass of the $[M + H]^+$ form of LBD tryptic peptides (as obtained from the SwissProt data bank) are specified: *m/z* 1284.59 (Leu^{406}−Arg^{416}, calcd value of 1284.767 Da), 1389.69 (Leu^{508}−Arg^{519}, calcd value of 1389.894 Da), 1545.78 (Arg^{507}−Arg^{519}, calcd value of 1545.995 Da), 1711.77 (Leu^{406}−Lys^{420}, calcd value of 1711.985 Da), 1955.87 (Ser^{453}−Lys^{471}, calcd value of 1956.105 Da), 2000.75 (Cys^{421}−Arg^{438}, calcd value of 2000.949 Da), and 2010.85 (Asn^{536}−Arg^{552}, calcd value of 2011.068 Da). In the inset, the deviation between the measured *m/z* value and the calculated *m/z* value for the postulated $[M + H]^+$ form of the LBD tryptic peptide is plotted vs the measured value. The good linear correlation that was obtained fully validates the signal attribution.

involved in the covalent attachment of 17BAPE$_2$, as such an event would shift the peptide signal by a Δ*m/z* increment of 357.22 (or 45.99) and not of 371.23 (see above). To determine which nucleophilic atom in the Cys^{421}−Arg^{438} peptide was the covalent attachment site of 17BAPE$_2$ and since three potential V8 proteinase cleavage sites were present on the peptide (Figure 8), V8-induced fragmentation of the ^{14}C-labeled peptide, followed by ZipTip purification of fragments, and then MALDI-TOF MS analysis were performed. Four signals corresponding to the $[M + H]^+$ form of peptide fragments [resulting from peptide cleavage at the second (Glu427) and third (Asp430) V8 proteinase sites] devoid of the affinity label were obtained (not shown). Only a single very faint signal was observed at *m/z* 721.30, which could correspond to the $[M + H]^+$ form of the Cys^{421}−Glu^{423} tripeptide with covalently bound dehydrobrominated [^{14}C]-17BAPE$_2$ (calcd value of *m/z* 721.37); several important signals, which at this point were not yet attributed to definite species, were present (Figure 10).

To obtain conclusive evidence, the synthetic CysValGlu-GlyMetValGlu heptapeptide, constituting the 7-mer N-terminal portion of the Cys^{421}−Arg^{438} peptide, was submitted to alkylation by [^{14}C]17BAPE$_2$. Under conditions more drastic (concerning temperature and compound concentrations) than those used for LBD alkylation (cf. Experimental Procedures), the heptapeptide reacted with [^{14}C]17BAPE$_2$ to give only one labeled species revealed by HPLC (not shown). Treatment of the latter with V8 proteinase gave rise to only one labeled fragment, which was purified by HPLC (not shown). The heptapeptide, the alkylated heptapeptide, and its labeled V8 fragment were analyzed by MS. The compared analysis of the patterns (Figure 11 and Table 1) obtained for the three species revealed their homology; it allowed the identification of the $[M + H]^+$ signals, despite the faint intensity of some of them, as well as additional forms which we could almost exhaustively identify (cf. Appendix). In the heptapeptide (Figure 11A) and labeled heptapeptide (Figure 11B) spectra, two major signals accurately matching the masses of $[M + Na]^+$ and $[M + K]^+$ or $[M + Na + O]^+$ ions were higher (heptapeptide) or even much higher (labeled heptapeptide) than those corresponding to $[M + H]^+$ ions, whereas two minor signals matched the masses of $[M + 2Na - H]^+$ and $[M + Na + K - H]^+$ or $[M + 2Na + O - H]^+$ ions. The existence of such Na- and/or K-containing ions, not observed in the MS analysis of LBD tryptic fragments, probably reflected the high affinity of the heptapeptide and labeled heptapeptide with the ubiquitous Na or K contaminants. For the labeled heptapeptide V8 fragment, the *m/z* values of prominent signals (Figure 11C) homologous to the above-mentioned signals indicated that the peptide moiety of this fragment was constituted by the CysValGlu tripeptide. Other homologous signals were present in the labeled heptapeptide and labeled tripeptide spectra. In each case, one signal (particularly strong for the labeled tripeptide) exactly matched the mass of $[M + H - OH_2]^+$ ions. However, two signals matched the masses of ions differing from $[M + H]^+$ and $[M + Na]^+$ species by a Δ*m/z* increment of very close to −16, and these ions could be derived from the above species by elimination of an undefined atom or group of atoms (designated Θ). All signals obtained with the [^{14}C]-17BAPE$_2$-labeled heptapeptide and labeled tripeptide, but not those obtained with the unlabeled heptapeptide, were preceded at Δ*m/z* −2.00 by weaker satellite signals, which were clearly the mark of 17BAPE$_2$-labeled species.

The results obtained in the analysis of the labeled synthetic heptapeptide and its labeled V8 fragment clearly indicated that the nonattributed signals in the MS analysis of V8 fragments of the labeled Cys^{421}−Arg^{438} peptide (Figure 10) were in fact due to $[M + Na]^+$, $[M + Na - \Theta]^+$, $[M + H - \Theta]^+$, and $[M + H - H_2O]^+$ ions of the labeled Cys^{421}−Glu^{423} tripeptide and labeled Cys^{421}-Glu^{427} heptapeptide (Table 2). Since S^- in Cys residues is a much stronger nucleophile than CO_2^- in Glu residues, as reflected by the absence of a reaction of the C-terminal Glu residue when the heptapeptide was submitted to 17BAPE$_2$ action, we concluded that the S atom of the Cys^{421} residue was the covalent attachment site of 17BAPE$_2$ on the Cys^{421}−Glu^{423} tripeptide, and hence on the Cys^{421}−Arg^{438} LBD tryptic peptide. S atoms of Cys421 and Cys534 constituted the two alternative and equivalent attachment sites of 17BAPE$_2$ on the mouse ERα LBD.

DISCUSSION

This study was devoted to the identification of covalent attachment sites of 17BAPE$_2$, an electrophilic, high-affinity ligand which exhibits estrogenic properties on mouse ERα. After we started with *E. coli* cell extracts containing a GST−

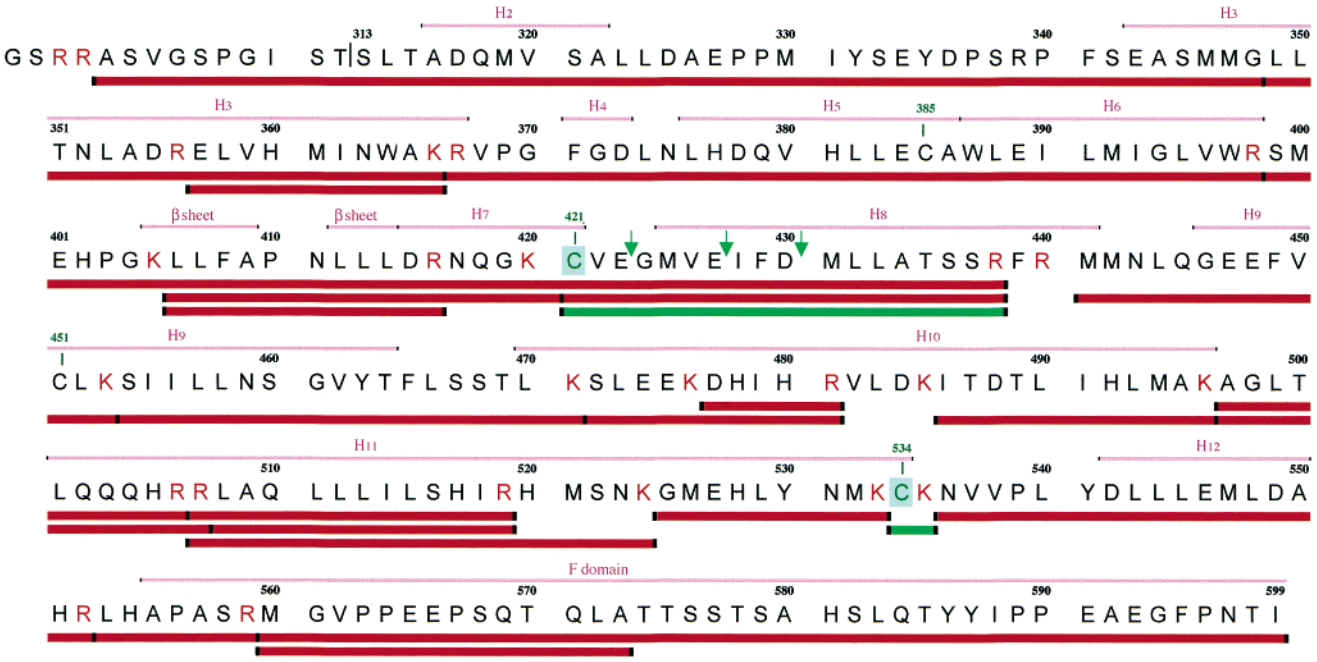

FIGURE 8: Map of LBD tryptic peptides identified by MS analysis. The sequence of the C-terminal moiety of the GST−LBD chimera (starting at the thrombin cleavage site) includes the full mouse ERα LBD encompassing the Ser^{313}−Ile^{599} sequence with the extra N-terminal 14 amino acids following the thrombin cleavage site. As inferred from the structure of the crystallized human ERα LBD−E_2 complex, lacking C-terminal domain F (the Ser^{309}−Pro^{552} human sequence, corresponding to the Ser^{313}−Pro^{556} mouse sequence), the 11 potential α-helices (H_2−H_{12}) and β-sheet which could occur in the mouse ERα LBD are represented (mauve segments above the sequence). Potential trypsin cleavage sites are immediately right of the red letters R and K. Following trypsin hydrolysis of the purified [^{14}C]17BAPE$_2$-labeled LBD, the various peptides identified from signals corresponding to their $[M + H]^+$ forms in MS analyses are materialized under the sequence: unmodified tryptic peptides (red segments) and [^{14}C]17BAPE$_2$-alkylated (via HBr elimination) tryptic peptides (green segments). The three potential V8 proteinase cleavage sites on the Cys^{421}−Arg^{438} tryptic peptide (green) are shown with arrows. The two cysteine residues, which are the covalent attachment sites of [^{14}C]17BAPE$_2$ on the LBD (cf. the text), are shown in green.

Table 1: MS Characteristics of the CysValGluGlyMetValGlu Heptapeptide, the [^{14}C]17BAPE$_2$-Labeled Heptapeptide, and Its Labeled V8 Proteolytic Fragment[a]

ion		CVEGMVE	(CVEGMVE)-APE$_2$	(CVE)-APE$_2$
$[M + H]^+$	*m/z*			
	theoretical	766.31	1137.55	721.37
	experimental	766.30 (51)	1137.94 (10)	721.76 (16)
		experimental Δ*m/z*		
alternative $[M + \text{“X”}]^+$	theoretical Δ*m/z*	CVEGMVE	(CVEGMVE)-APE$_2$	(CVE)-APE$_2$
$[M + H - H_2O]^+$	−18.01	nd[c]	−17.97 (13)	−17.97 (100)
$[M + H - \Theta]^+$	−16.00[b]	−16.00 (0.5)	−15.99 (16)	−15.79[d] (50)
$[M + Na - \Theta]^+$	5.98[b]	nd[c]	nd[c]	6.14[d] (0.5)
$[M + Na]^+$	21.98	21.94 (100)	22.05 (59)	22.06 (58)
$[M + K]^+/[M + Na + O]^+$	37.96/37.98	37.91 (30)	38.02 (100)	38.01 (38)
$[M + 2Na - H]^+$	43.96	43.91 (10)	44.04 (1)	44.03 (10)
$[M + K + O]^+/[M + Na + 2O]^+$	53.95/53.97	nd[c]	54.00 (14)	nd[c]
$[M + Na + K - H]^+/[M + 2Na + O - H]^+$	59.94/59.96	59.89 (10)	60.00 (14)	60.00 (15)

[a] Theoretical and experimental *m/z* values, corresponding to the $[M + H]^+$ ion, of the synthetic CVEGMVE heptapeptide, its [^{14}C]17BAPE$_2$-labeled form, (CVEGMVE)-APE$_2$, and the labeled tripeptide (CVE)-APE$_2$, postulated to be the labeled V8 proteolytic fragment of the latter, are given in the upper part of the table. Alternative ions ($[M + \text{“X”}]^+$), as inferred for the main signals found in the MS spectra of the three above-mentioned species (cf Figure 12), are listed in the left column; in the second and third ion species, Θ symbolizes an undefined atom or group of atoms (e.g., O, NH_2, or CH_4) displaying a monoisotopic mass very close to 16. For each of these ion types, the theoretical Δ*m/z* increment (difference in *m/z* values for $[M + \text{“X”}]^+$ and $[M + H]^+$ ions, equal to the difference in monoisotopic masses of “X” and H) is given. The corresponding experimental Δ*m/z* values in the MS analyses are given for the three peptide species. Values in brackets indicate, for each particular ion, the relative intensity of the corresponding signal in the spectrum of the analyzed species (expressed as a percentage of the most intense signal in the spectrum). [b] Calculated value assuming a monoisotopic mass of 16.00 for Θ. [c] Not detected (due to the absence of the corresponding signal in the spectrum). [d] Slightly inaccurate value (due to the presence of an interfering signal in the spectrum).

LBD chimera, an affinity chromatography step with in situ affinity labeling of the GST−LBD chimera followed by HPLC separation of eluted species led to the highly purified [^{14}C]17BAPE$_2$-alkylated LBD. The use of a ^{14}C-labeled form of 17BAPE$_2$ considerably facilitated the characterization and then purification of the expressed GST−LBD forms and derived LBD forms. Proteolyses of the alkylated LBD followed by HPLC separation and MALDI-TOF MS analysis of LBD fragments clearly identified Cys421 and Cys534 (the mouse homologues of human ERα Cys417 and Cys530, respectively) as alternative and equivalent covalent attachment sites of the electrophile.

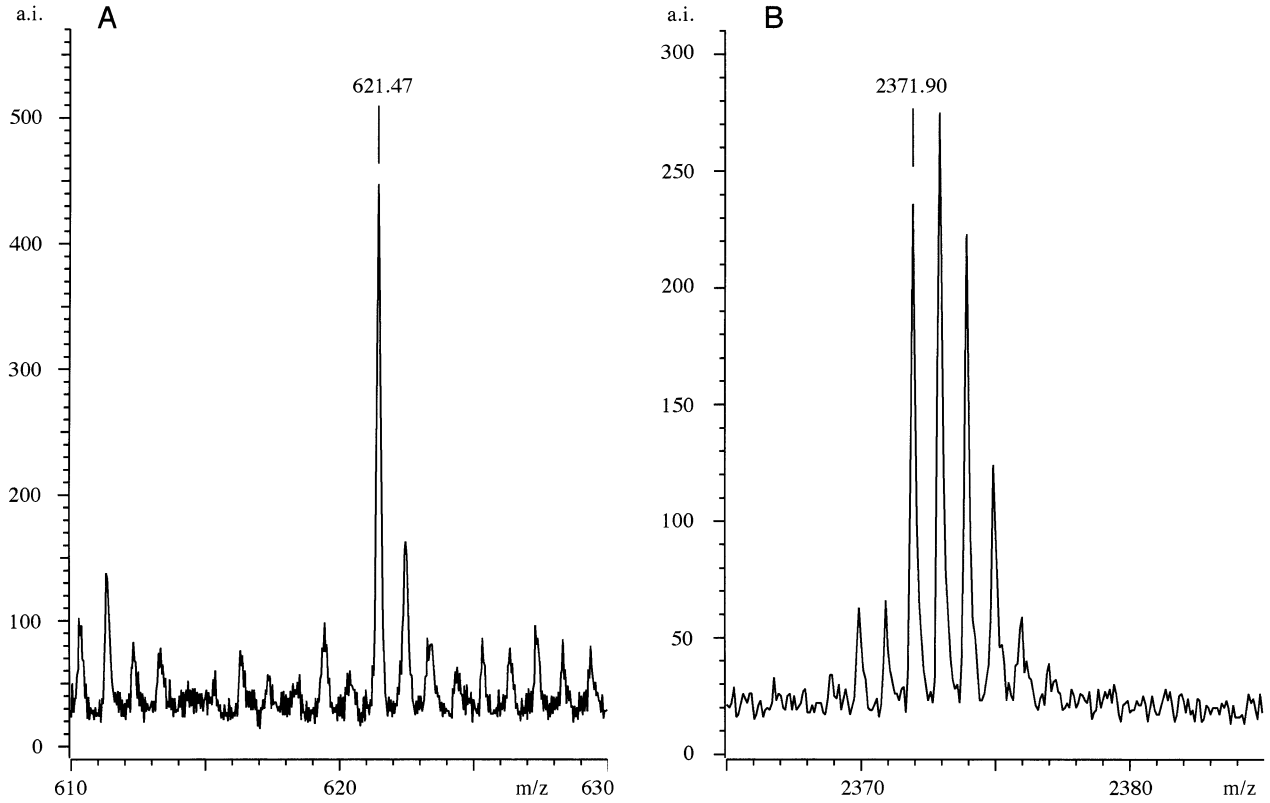

FIGURE 9: MALDI-TOF MS analysis of ^{14}C-labeled species obtained from trypsin hydrolysis of the [^{14}C]17BAPE$_2$-alkylated LBD. The two radioactive fractions (11 and 16 in Figure 6) obtained in the HPLC separation of labeled LBD tryptic fragments were analyzed by MS (CHCA matrix and positive reflectron mode). For each fraction, the obtained signal compatible with a [^{14}C]17BAPE$_2$-alkylated form of a LBD tryptic peptide is shown: (A) analysis of fraction 11 and (B) analysis of fraction 16. Signals at *m/z* 621.47 and 2371.90 accurately match the monoisotopic masses of pseudomolecular $[M + H]^+$ ions of Cys^{534}Lys^{535} and Cys^{421}−Arg^{438} tryptic peptides, respectively, alkylated by the ^{14}C form of 17BAPE$_2$ ([M + H] calcd values of 621.36 and 2372.18 Da for dehydrobrominated adducts, respectively). The companion signals observed just ahead of these two signals (at $\Delta m/z$ −2.00) correspond to tryptic peptides alkylated by nonradioactive 17BAPE$_2$ molecules; their intensities, relative to those of the main signals, closely reflect the proportion of [NHCO-^{12}C]17BAPE$_2$ and [NHCO-^{14}C]-17BAPE$_2$ (1/7.33) in the affinity probe that was used. Note that fine structures of the two major signals, characterized by the relative intensities of $[M + H]^+$, $[M + H + 1]^+$, $[M + H + 2]^+$ ... peaks (due to the occurrence of 1.1% ^{13}C in ^{12}C), reflect the total carbon numbers in alkylated Cys^{534}Lys^{535} (31 unlabeled carbons and one ^{14}C) and alkylated Cys^{421}−Arg^{438} (107 unlabeled carbons and one ^{14}C) LBD tryptic peptides.

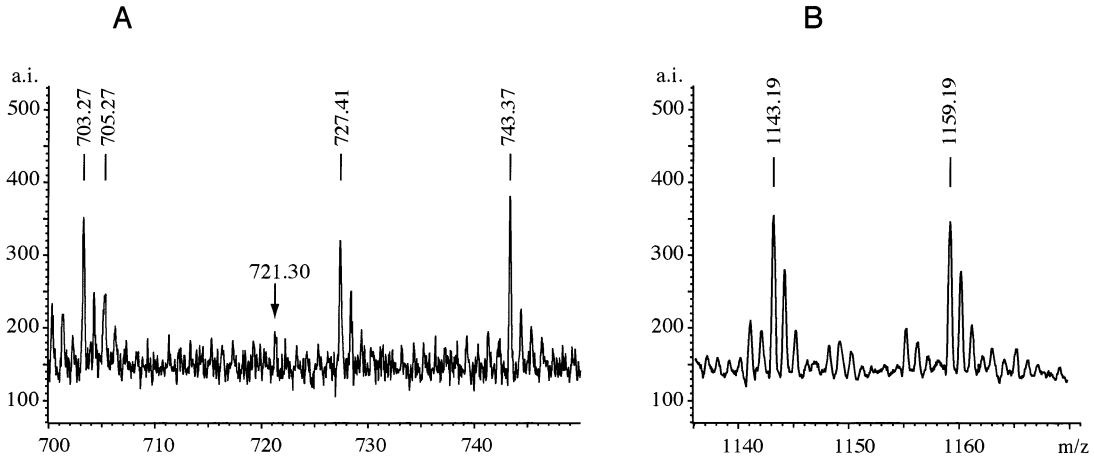

FIGURE 10: MALDI-TOF MS analysis of V8 proteolytic fragments of the [^{14}C]17BAPE$_2$-labeled Cys^{421}−Arg^{438} peptide. The alkylated Cys^{421}−Arg^{438} LBD tryptic peptide (HPLC fraction 16 in Figure 6) was submitted to V8 proteinase hydrolysis, desalted with ZipTip, and then analyzed by MS (CHCA matrix and positive reflectron mode). Signals shown in panels A and B (obtained in two different experiments) characterize $[M + H - H_2O]^+$, $[M + H - \Theta]^+$, $[M + Na - \Theta]^+$, or $[M + Na]^+$ ions of [^{14}C]17BAPE$_2$-labeled Cys^{421}−Glu^{423} (A) and Cys^{421}−Glu^{427} (B) fragments of the alkylated Cys^{421}−Glu^{438} peptide (cf. the text, and the legend of Table 2); only a very faint signal (*m/z* 721.30, shown with an arrow) was obtained for the $[M + H]^+$ ion of the alkylated Cys^{421}−Glu^{423} peptide.

These results validate and supplement those obtained in a previous study (*21*) in which covalent attachment sites of several electrophilic E_2 17α-derivatives, including 17BAPE$_2$, were ascribed to Cys417 and Cys530 of human ERα according to indirect evidence. They were based on the inability of these electrophiles to irreversibly inhibit E_2 binding to the Cys417Ala/Cys530Ala double-mutant ER, whereas the two Cys417Ala and Cys530Ala single mutants,

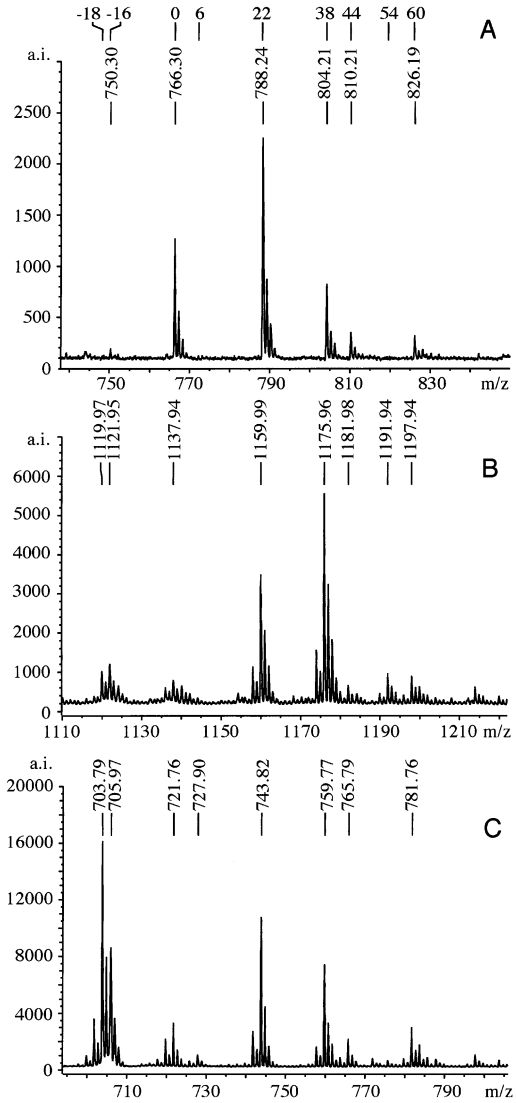

FIGURE 11: MS analysis of the unlabeled CysValGluGlyMetValGlu heptapeptide, the [^{14}C]17BAPE$_2$-labeled heptapeptide, and its labeled V8 proteolytic fragment. The synthetic heptapeptide was labeled with [^{14}C]17BAPE$_2$ in 0.1 M $NaHCO_3$ at 20 °C. Submitting the reaction mixture to HPLC led to only one ^{14}C-containing, 220 nm-absorbing species. An aliquot of the corresponding HPLC fraction was submitted to V8 proteinase hydrolysis. HPLC fractionation of the hydrolysate gave a single fraction showing both UV absorption and a high radioactivity level. The synthetic heptapeptide (A) and the HPLC fractions containing the [^{14}C]-17BAPE$_2$-labeled heptapeptide (B) or its labeled V8 fragment (C) were analyzed by MALDI-TOF MS (CHCA matrix and positive reflectron mode). The obtained spectra are represented so that the *m*/*z* experimental values of signals corresponding to the monoisotopic $[M + H]^+$ ion of the synthetic heptapeptide (measured *m*/*z* 766.30, calcd *m*/*z* 766.31), the [^{14}C]17BAPE$_2$-alkylated heptapeptide (measured *m*/*z* 1137.94, calcd *m*/*z* 1137.55), and the [^{14}C]17BAPE$_2$-alkylated CysValGlu tripeptide, postulated to be the labeled V8 fragment of the latter (measured *m*/*z* 721.76, calcd *m*/*z* 721.37) are aligned. The values displayed at the top of the figure indicate (for A−C spectra) the various Δ*m*/*z* increments (starting from the *m*/*z* values corresponding to the $[M + H]^+$ forms) characterizing ions as opposed to the $[M + H]^+$ ions, i.e., −18 for $[M + H - H_2O]^+$, −16 for $[M + H - \Theta]^+$, 6 for $[M + Na - \Theta]^+$, 22 for $[M + Na]^+$, 38 for $[M + K]^+$ or $[M + Na + O]^+$, 44 for $[M + 2Na - H]^+$, and 60 for $[M + Na + K - H]^+$ or $[M + 2Na + O - H]^+$. Note that, except for the heptapeptide signals, all of the other above-mentioned signals are accompanied by weaker satellite signals at Δ*m*/*z* −2.00, due to heptapeptide and tripeptide species alkylated by nonradiolabeled 17BAPE$_2$ molecules present in [^{14}C]17BAPE$_2$ (cf. the legend of Figure 9).

Table 2: MS Characteristics of [^{14}C]17BAPE$_2$-Labeled Cys^{421}−Arg^{438} V8 Proteolytic Fragments[a]

	m/*z* (Cys^{421}−Glu^{423})-APE$_2$		*m*/*z* (Cys^{421}−Glu^{427})-APE$_2$	
ion	exptl	calcd	exptl	calcd
$[M + H - H_2O]^+$	703.27	703.36	nd[c]	1119.54
$[M - H - \Theta]^+$	705.27	705.37[b]	nd[c]	1121.55
$[M + Na - \Theta]^+$	727.41	727.35[b]	1143.19	1143.53[b]
$[M + Na]^+$	743.37	743.35	1159.19	1159.53

[a] Experimental *m*/*z* values of signals obtained in the MS analysis of alkylated Cys^{421}−Arg^{438} V8 proteolytic fragments (as shown in Figure 10), which accurately matched the calculated values for various ions corresponding to the alkylated Cys^{421}−Glu^{423} tripeptide, (Cys^{421}−Glu^{423})-APE$_2$, and the alkylated Cys^{421}−Glu^{427} heptapeptide, (Cys^{421}−Glu^{427})-APE$_2$, are given. In the second and third ion species, Θ symbolizes an undefined atom or group of atoms displaying a monoisotopic mass very close to 16. [b] Calculated value assuming a monoisotopic mass of 16.00 for Θ. [c] Not detected (due to the absence of the corresponding signal in the spectrum).

like wild-type ER, were susceptible to irreversible E_2 binding inhibition. However, this did not allow us to clarify the following situations. (i) Both Cys417 and Cys530 might be alternative covalent attachment sites of these electrophiles (note that in this case the approach that is used cannot determine the relative involvements of Cys417 and Cys530 in the alkylation processes). (ii) Only one of the two cysteines might be the covalent attachment site while the other becomes an attachment site when the target cysteine is substituted for alanine [a similar case was observed with tamoxifen aziridine, a triphenylethylene antiestrogen, which alkylated the wild-type ER only at Cys530, but quite likely alkylated the Cys530Ala mutant at Cys381 (*26*)]. (iii) Neither of the two cysteines is a covalent attachment site: the inability of electrophiles to alkylate the Cys417Ala/Cys530Ala double mutant would thus result from ER conformational changes, induced by the two mutations, which move off (or hinder) the ER covalent attachment site from the electrophilic center of the bound steroid. Although unlikely, this possibility cannot be totally excluded since a triple cysteine to serine substitution, especially involving both Cys417 and Cys530, induced pronounced structural changes in a crystallized form of the ERα LBD−E_2 complex (*12*).

To date, only Cys417 and Cys530 have been identified as covalent attachment sites of electrophilic ligands on human ERα. Cys530 is the covalent attachment site common to (i) both steroidal [E_2 17α-derivatives (ref *21* and this study)] and nonsteroidal [ketononestrol aziridine (*27*), a diphenylethane derivative] estrogens and (ii) both steroidal [E_2 11*β*-derivatives (*28*)] and nonsteroidal [the aforementioned tamoxifen aziridine (*27*)] antiestrogens. Conversely, as confirmed in this study, Cys417 is the covalent attachment site of only the estrogenic E_2 17α-derivatives. Note that non-cysteinyl residues (but not yet identified) appeared to be involved in the covalent attachment of other antiestrogenic E_2 11*β*-derivatives to lamb ERα (*29*). The unique dual covalent attachment sites of E_2 17α-derivatives on ERαs may result from the particular location of their electrophilic groups relative to the steroid nucleus (17α-orientation) or from their type of activity (they are estrogen agonists, whereas the above-mentioned E_2 11*β*-derivatives are potent antagonists). To investigate this point, it would be interesting to determine the covalent attachment sites of new E_2-derived electrophiles,

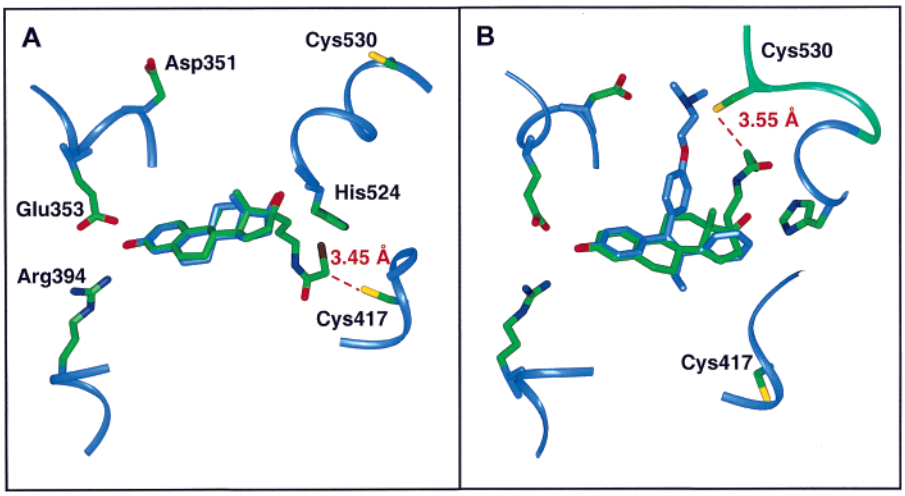

FIGURE 12: Docking of 17BAPE$_2$ into human ERα LBD structures derived from crystallized LBD−E$_2$ and LBD−OHT complexes. (A) 17BAPE$_2$ was manually docked into the fixed LBP structure, deduced from X-ray analysis of the crystallized human LBD−E$_2$ complex, as described in Experimental Procedures. The relative positions of E$_2$ (blue) and that of superimposed 17BAPE$_2$ (green) are shown in the LBP materialized by portions (in blue) of helix 3 (Asp351 and Glu353), helix 6 (Arg394), helix 7 and the helix 7−helix 8 loop (Cys417), and helix 11 (His524 and Cys530). The main characteristics of docked 17BAPE$_2$ are as follows: ϵ-carbon−Cys417 S distance (*d*) of 3.45 Å, 3-OH−Glu353 *d* of 2.31 Å, 3-OH−Arg394 *d* of 3.00 Å, and 17β-OH−His524 *d* of 2.62 Å. (B) 17BAPE$_2$ was docked through a two-step minimization process, into the LBD structure determined for the crystallized human LBD−OHT complex, as described in Experimental Procedures. The relative positions of OHT (blue) and docked 17BAPE$_2$ (green) are shown in the LBP materialized by portions of helix 3, helix 6, helix 7, the helix 7−helix 8 loop, helix 11 (His524), and the helix 11−helix 12 loop (Cys530). LBD original elements are in blue, whereas the helix 11−helix 12 loop, the conformation of which was modified, is in blue-green. The main characteristics of docked 17BAPE$_2$ are as follows: ϵ-carbon−Cys530 S *d* of 3.55 Å, 3-OH−Glu353 *d* of 2.74 Å, 3-OH−Arg394 *d* of 3.01 Å, and 17β-OH−His524 *d* of 3.25 Å.

e.g., estrogens possessing a relatively short 11β-substituent with a terminal bromoacetamido group and 11β-aryl antiestrogens bearing a 17α-bromoacetamidoalkyl group.

Mouse and human ERα LBDs differ by only eight amino acid residues (of more than 250 residues). Hence, with the aim of accounting for the alternative covalent attachments of 17BAPE$_2$ to human ERα Cys417 and Cys530 (i.e., to obtain close locations between S atoms of the two cysteines and the ϵ-carbon of bound compound), 17BAPE$_2$ was docked into available crystallographic structures of the human ERα LBD. In the LBP region, many structural features are common to the four crystallized LBD−estrogen and LBD−antiestrogen complexes (*8*, *10*); especially Glu353, in helix 3, and Arg394, in helix 6, make direct hydrogen bonds with the E$_2$ phenolic OH (or corresponding phenolic OH in DES, OHT, or Ral), and His524, in helix 11, is involved in hydrogen bonding of E$_2$ 17β-OH (or the second phenolic OH of DES or Ral). However, discrete variations were reported between crystallized estrogen and antiestrogen complexes, with some of them specifically affecting the structure of Cys417- and Cys530-bearing segments and then the location of these two amino acids relative to invariant structural elements. LBD structures derived from the crystallized domain bound to E$_2$, OHT, and Ral were therefore used for a manual exploration (followed by minimization) of potential proximities between the 17BAPE$_2$ ϵ-carbon and S atoms of Cys417 and Cys530. This was carried out assuming that (i) LBD helices and β−sheet would not vary and (ii) 17BAPE$_2$ 3-OH and 17β-OH, as observed with their counterparts in crystallized estrogen− and antiestrogen−LBD complexes, would interact with the above-mentioned key amino acids.

Since 17BAPE$_2$ includes the whole E$_2$ structure and displays estrogenic properties (*20*), the LBD structure derived from LBD−E$_2$ crystals (*8*) was first used for the ligand docking. From the steroid moiety of 17BAPE$_2$, positioned like E$_2$ in the LBD−E$_2$ crystals, the flexibility of the 17α-arm (which confers four degrees of freedom to the ϵ-carbon from steroidal C17) allowed us to bring the ϵ-carbon into the proximity of the S atom of Cys417 (distance of ∼3.45 Å), with opposite locations of S and Br atoms, and without any steric clashes between 17BAPE$_2$ and LBP amino acids (Figure 12A). However, we could not obtain a distance of less than 6 Å between the ϵ-carbon and the S atom of Cys530. Since the sum of S and C van der Waals radii is ∼3.36 Å, the distance of 3.45 Å between the ϵ-carbon and the S atom of Cys417 appeared to be very favorable for nucleophilic attack of the latter, whereas the distance of ∼6 Å seems to be unsuitable for attack of the S atom of Cys530.

Due to structural changes (involving helix 7, the helix 7−helix 8 loop, and the helix 11−helix 12 loop) between crystallized LBD−E$_2$ and LBD−OHT complexes (*10*), different results were obtained when 17BAPE$_2$ was docked into the LBD structure inferred from LBD−OHT crystals. (i) The differential orientation of helix 7 moves off Cys417 (still located at the C-terminal end of that helix) from the LBP core and therefore from the ϵ-carbon which, despite the mobility of the 17α-substituent, cannot approach the S atom of Cys417 within <5 Å. (ii) Conversely, the shortening of helix 11 conferred mobility to Cys530 (shift in the helix 11−helix 12 loop); then combining the Cys530 and 17α-substituent mobilities allowed us to bring the S atom of Cys530 and the ϵ-carbon to the very convenient distance of ∼3.55 Å in the minimized complex, while maintaining the interactions of 3-OH and 17β-OH with Glu353, Arg394, and His524, respectively (Figure 12B).

Finally, the potential proximities of the electrophilic center with respect to Cys417 and Cys530 in the LBD−17BAPE$_2$ complex were examined using the structure determined for the crystallized LBD−Ral complex (*8*). In the helix 7−helix

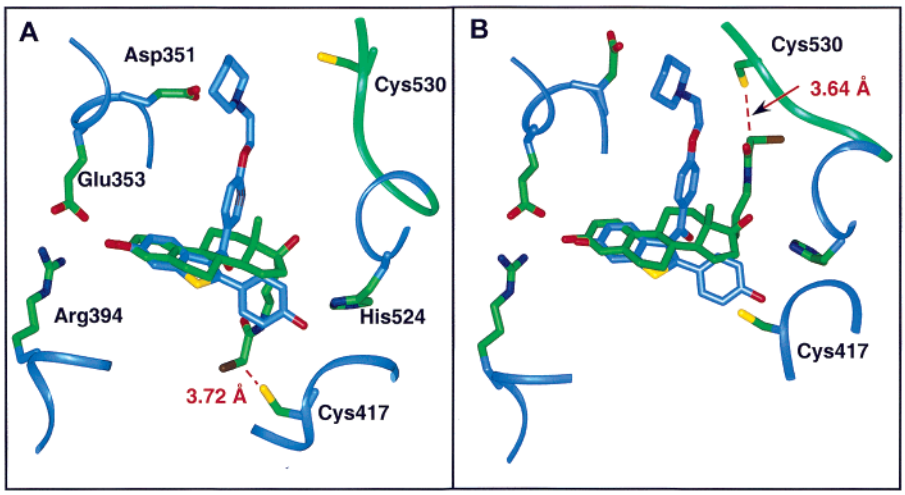

FIGURE 13: Docking of 17BAPE$_2$ into the human ERα LBD structure derived from the crystallized LBD−Ral complex. 17BAPE$_2$ was docked through a two-step minimization process, into the LBD structure determined for the crystallized human LBD−Ral complex (with the helix 11−helix 12 loop imported from the modified LBD−OHT crystallized complex), as described in Experimental Procedures. The docking was performed with the 17α-substituent oriented toward the C-terminal end of either helix 7 (A) or the helix 11−helix 12 loop (B). The relative positioning of Ral (blue) and 17BAPE$_2$ (green) is shown in the LBP materialized by portions of helix 3 (Asp351 and Glu353), helix 6 (Arg394), helix 7 and the helix 7−helix 8 loop (Cys417), helix 11 (His524), and the helix 11−helix 12 loop (Cys530). LBD original elements are in blue, whereas the helix 11−helix 12 loop, the conformation of which was modified, is in blue-green. The main characteristics of docked 17BAPE$_2$ are as follows: (A) ϵ-carbon−Cys417 S distance (*d*) of 3.72 Å, 3-OH−Glu353 *d* of 2.56 Å, 3-OH−Arg394 *d* of 2.88 Å, and 17β-OH−His524 *d* of 4.48 Å and (B) ϵ-carbon−Cys530 S *d* of 3.64 Å, 3-OH−Glu353 *d* of 2.38 Å, 3-OH−Arg394 *d* of 3.07 Å, and 17β-OH−His524 *d* of 3.19 Å.

8 region, the structure of this complex was closer to that of the LBD−E$_2$ complex than that of the LBD−OHT complex. Therefore, using the 17α-substituent flexibility, it was possible to bring the ϵ-carbon close to the S atom of Cys417 (distance of ∼3.72 Å, Figure 13A). In the minimized complex, interactions of 3-OH with Glu353 and Arg394 were maintained, whereas the 17β-OH−His524 interaction was lost. In the reported LBD−Ral structure, the helix 11−helix 12 loop (including Cys530) was lacking. Since Ral and OHT belong to the same class of estrogen antagonists, the lacking loop was imported from the LBD−OHT complex. Then, by means of both 17α-substituent and helix 11−helix 12 loop flexibilities, it was possible to bring the ϵ-carbon close to the S atom of Cys530 (distance of ∼3.64 Å, Figure 13B). In this case, all 3-OH and 17β-OH interactions (with Glu353, Arg394, and His524) were maintained.

Without any extensive molecular dynamics simulation analyses, the results of 17BAPE$_2$ docking into manually modified and minimized structures derived from the various crystallized complexes accounted well for the LBD specific labeling at Cys417 and Cys530: the proximities of the S atoms of the two cysteines and the ϵ-carbon of 17BAPE$_2$ could be obtained without disruption of canonical ligand−ER interactions. The same two cysteines were alkylated when the human ERα LBD, interacting with E$_2$ covalently linked to Sepharose via a 17α-propylthiopropyl arm (*30*), was exposed to the cysteine reagent iodoacetic acid (*31*−*33*); however, the required iodoacetic acid concentration was 500−2000-fold higher than the 17BAPE$_2$ concentration we used, and the reaction time was also ∼10-fold higher. These considerable discrepancies reflect the two types of LBD alkylation processes, i.e., nonspecific with free iodoacetic acid and specific with the E$_2$-linked bromoacetamido group of 17BAPE$_2$.

The fact that only a model deduced from the crystallized LBD bound to Ral [a primarily antiestrogenic compound (*34*)] can account for the simultaneous labeling of Cys417 and Cys530 seems to conflict with the agonist behavior of 17BAPE$_2$. Even though if it is difficult to predict how conformational changes are upset by 17BAPE$_2$ as compared to E$_2$, we favor the hypothesis of alternative labeling of the two cysteines due to equilibria between multiple conformational states of ER−ligand complexes. These interpretations are in agreement with previous crystallographic studies involving a cysteine/serine mutant ERα (*11*) and the partial agonism of some ER ligands (*35*). The fact that the crystal data we used were obtained from the human ER LBD carboxymethylated at its cysteine residues, including Cys417 and Cys530, is an additional handicap for predicting the molecular basis of these conformational changes.

In conclusion, the results obtained suggest that methods other than the crystallization approach could help in gaining further insight into the dynamics of nuclear receptor−ligand interactions: the experimental results and modeling studies point to a view of ER−ligand complexes much more conformationally dynamic than that indicated by the crystal structure of these complexes. A study using specific amino acid reagents is under way to determine, by means of MS analysis, the pattern of modified amino acids in the human ERα LBD according to the type of bound ligand. This ER "footprinting" approach should allow us to establish solvent-exposed surfaces in the various LBD−ligand complexes, information that could be potentially useful for interpreting the agonist and antagonist activities of ER ligands. This work also illustrates that MS is a powerful tool for identifying covalent modifications on nuclear receptors, resulting from either physiological processes (e.g., phosphorylation, glycosylation, acetylation, etc.) or artifactual processes (e.g., affinity labeling to determine nuclear receptor−ligand proximities and specific amino acid derivation to probe conformations of nuclear receptor−ligand complexes).

ACKNOWLEDGMENT

We are grateful to Dr. Vincent Cavaillès for his donation of the plasmid encoding the GST−LBD chimera, and we thank Emmanuelle Demay for MS data acquisition.

APPENDIX

Interpretation of the Signal Diversity in MS Analyses of the Heptapeptide and Derivatives. The occurrence of non-conventional ions in the MS analyses of the synthetic CysValGluGlyMetValGlu heptapeptide, [^{14}C]17BAPE$_2$-labeled heptapeptide, and labeled CysValGlu tripeptide (or labeled Cys^{421}−Arg^{438} V8 proteolytic fragments) deserve commentary.

For these species, most signals attributed to Na- or/and K-containing pseudomolecular ions were stronger than those of classical $[M + H]^+$ ions. This suggests that, due to their particular amino acid sequences (involving hydrophobic residues, one C-terminal Glu residue, and, for heptapeptide species, an additional Glu residue), the heptapeptide and tripeptide species display high affinities for Na^+ and K^+ and therefore form stable complexes with these ubiquitous cation contaminants. Detection of such alkali−metal adducts has previously been reported in MS analyses of various peptidyl or non-peptidyl structures.

Apart from the unpredictable saline forms for all three analyzed species, the occurrence of chemical modifications [mainly oxidation, but also, in the case of labeled species, elimination of H_2O or of an unascribed group ("Θ") displaying a monoisotopic mass very close to 16] contributed to the increase in the ion diversity and therefore the signal number in spectra. First, signals attributed to singly and doubly oxidized forms were likely due to monooxidation at thioether S atoms (commonly observed for methionine-containing peptides) of species enclosing either one thioether function (unlabeled heptapeptide and cysteinyl S-alkylated tripeptide) or two thioether functions (labeled heptapeptide). For the various species, the closeness of the oxygen monoisotopic mass (15.995 Da) and the difference in the K and Na monoisotopic masses (15.974 Da) cannot allow discrimination between the signal of a K form and that of a corresponding monooxidized Na form. Second, the presence in MS spectra of labeled species (mainly labeled tripeptide) of signals at $\Delta m/z$ −18 or −16, which were absent in the spectra of the unlabeled heptapeptide, suggested a role for the steroidal moiety in generating forms derived from $[M + H]^+$ ions by elimination of H_2O or the Θ group, respectively (e.g., the labile tertiary steroidal 17β-OH could be involved in H_2O elimination).

Note that (i) signals at $\Delta m/z$ −18 and −16 were also observed in MS analyses of the labeled Cys^{534}Lys^{535} LBD tryptic fragment, which displays some characteristics very similar to those of labeled tripeptide, i.e., short peptide moiety with S-alkylated cysteine in its N-terminus, (ii) the similarity of the apparent monoisotopic mass of the Θ group with that of O and with the difference in K and Na monoisotopic masses increases the uncertainty concerning the number and the relative involvement of differing forms in a given signal, and (iii) from one experiment to another slight variations in the relative intensities of signals were observed, which probably reflected either differing Na^+/K^+ ratios in samples or differing extents of species oxidation or dehydration.

REFERENCES

1. Green, S., Walter, P., Kumar, V., Krust, A., Bornet, J. M., Argos, P., and Chambon, P. (1986) Human oestrogen receptor cDNA: sequence, expression and homology to v-*erb-A*, *Nature 320*, 134−139.
2. Greene, G. L., Gilna, P., Waterfield, M., Baker, A., Hort, Y., and Shine, J. (1986) Sequence and expression of human estrogen receptor complementary DNA, *Science 231*, 1150−1154.
3. Mosselman, S., Polman, J., and Dijkema, R. (1996) ERβ: identification and characterization of a novel human estrogen receptor, *FEBS Lett. 392*, 49−53.
4. Kuiper, G. G. J. M., Enmark, E., Pelto-Huikko, M., Nilsson, S., and Gustafsson, J. Å. (1996) Cloning of a novel estrogen receptor expressed in rat prostate and ovary, *Proc. Natl. Acad. Sci. U.S.A. 93*, 5925−5930.
5. Tsai, M. J., and O'Malley, B. W. (1994) Molecular mechanisms of action of steroid/thyroid receptor superfamily members, *Annu. Rev. Biochem. 63*, 451−486.
6. Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schultz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., and Evans, R. M. (1995) The nuclear receptor superfamily: the second decade, *Cell 83*, 835−839.
7. Egea, P. F., Klaholz, B. P., and Moras, D. (2000) Ligand-protein interactions in nuclear receptors of hormones, *FEBS Lett. 476*, 62−67.
8. Brzozowski, A. M., Pike, A. C. W., Dauter, Z., Hubbard, R. E., Bonn, T., Engström, O., Öhman, L., Greene, G. L., Gustafsson, J. Å., and Carlquist, M. (1997) Molecular basis of agonism and antagonism in the oestrogen receptor, *Nature 389*, 753−758.
9. Tanenbaum, D. M., Wang, Y., Williams, S. P., and Sigler, P. B. (1998) Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains, *Proc. Natl. Acad. Sci. U.S.A. 95*, 5998−6003.
10. Shiau, A. K., Barstad, D., Loria, P. M., Cheng, L., Kushner, P. J., Agard, D. A., and Greene, G. L. (1998) The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen, *Cell 95*, 927−937.
11. Gangloff, M., Ruff, M., Eiler, S., Duclaud, S., Wurtz, J., and Moras, D. (2001) Crystal structure of a mutant hERα ligand-binding domain reveals key structural features for the mechanism of partial agonism, *J. Biol. Chem. 276*, 15059−15065.
12. Eiler, S., Gangloff, M., Duclaud, S., Moras, D., and Ruff, M. (2001) Overexpression, purification and crystal structure of native ERα LBD, *Protein Expression Purif. 22*, 165−173.
13. Pike, A. C. W., Brzozowski, A. M., Hubbard, R. E., Bonn, T., Thorsell, A. G., Engström, O., Ljunggren, J., Gustafsson, J. Å., and Carlquist, M. (1999) Structure of the ligand-binding domain of oestrogen receptor beta in the presence of partial agonist and a full antagonist, *EMBO J. 18*, 4608−4618.
14. Pike, A. C. W., Brzozowski, A. M., Walton, J., Hubbard, R. E., Thorsell, A. G., Li, Y. L., Gustafsson, J. Å., and Carlquist, M. (2001) Structural insights into the mode of action of a pure antiestrogen, *Structure 9*, 145−193.
15. Wurtz, J. M., Bourguet, W., Renaud, J. P., Vivat, V., Chambon, P., Moras, D., and Gronemeyer, H. (1996) A canonical structure for the ligand-binding domain of nuclear receptors, *Nat. Struct. Biol. 3*, 87−94.
16. Lees, J. A., Fawell, S. E., and Parker, M. G. (1989) Identification of two transactivation domains in the mouse estrogen receptor, *Nucleic Acids Res. 17*, 5477−5488.
17. Tora, L., White, J., Brou, C., Tasset, D., Webster, N., Scheer, E., and Chambon, P. (1989) The human estrogen receptor has two independent nonacidic transcriptional activation functions, *Cell 59*, 477−487.
18. Feng, W., Ribiero, R. C., Wagner, R. L., Nguyen, H., Apriletti, J. W., Fletterick, R. J., Baxter, J. D., Kushner, P. J., and West, B. L. (1998) Hormone-dependent coactivator binding to a hydrophobic cleft on nuclear receptors, *Science 280*, 1747−1749.
19. Grese, T. A., and Dodge, J. A. (1998) Selective estrogen receptor modulators (SERMs), *Curr. Pharmacol. Des. 4*, 71−92.
20. El Garrouj, D., Aliau, S., Aumelas, A., and Borgna, J. L. (1995) Steroidal affinity labels of the estrogen receptor. 2. 17α-(haloacetamidoalkyl)estradiols, *J. Med. Chem. 38*, 2339−2348.
21. Aliau, S., El Garrouj, D., Yasri, A., Katzenellenbogen, B. S., and Borgna, J. L. (1997) 17α-(Haloacetamidoalkyl)estradiols alkylate the human estrogen receptor at cysteine residues 417 and 530, *Biochemistry 36*, 5861−5867.

22. Cavaillès, V., Dauvois, S., Danielan, P. S., and Parker, M. G. (1994) Interactions of proteins with transcriptionally active estrogen receptors, *Proc. Natl. Acad. Sci. U.S.A. 91*, 10009−10013.
23. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding, *Anal. Biochem. 72*, 248−254.
24. Sherman, M. Y., and Goldberg, A. L. (1992) Involvement of the chaperonin dnaK in the rapid degradation of a mutant protein in *Escherichia coli*, *EMBO J. 11*, 71−77.
25. Seielstad, D. A., Carlson, K. E., Katzenellenbogen, J. A., Kushner, P. J., and Greene, G. L. (1995) Molecular characterization by mass spectrometry of the human estrogen receptor ligand-binding domain expressed in *Escherichia coli*, *Mol. Endocrinol. 9*, 647−658.
26. Reese, J. C., and Katzenellenbogen, B. S. (1991) Mutagenesis of cysteines in the hormone binding domain of the human estrogen receptor, *J. Biol. Chem. 266*, 10880−10887.
27. Harlow, K. W., Smith, D. N., Katzenellenbogen, J. A., Greene, G. L., and Katzenellenbogen, B. S. (1989) Identification of cysteine-530 as the covalent attachment site of an affinity labeling estrogen (ketononestrol aziridine) and antiestrogen (tamoxifen aziridine) in the human estrogen receptor, *J. Biol. Chem. 264*, 17476−17485.
28. Aliau, S., Mattras, H., Richard, E., and Borgna, J. L. (1999) Cysteine 530 of the human estrogen receptor α is the main covalent attachment site of 11β-(aziridinylalkoxyphenyl)estradiols, *Biochemistry 38*, 14752−14762.
29. Aliau, S., Delettre, G., Mattras, H., El Garrouj, D., Nique, F., Teutsch, G., and Borgna, J. L. (2000) Steroidal affinity labels of the estrogen receptor α. 4. Electrophilic 11β-aryl derivatives of estradiol, *J. Med. Chem. 43*, 613−628.
30. Greene, G. L., Nolan, C., Engler, J. P., and Jensen, E. V. (1980) Monoclonal antibodies to human estrogen receptor, *Proc. Natl. Acad. Sci. U.S.A. 77*, 5115−5119.
31. Hegy, G. B., Shackleton, C. H. L., Carlquist, M., Bonn, T., Engström, O., Sjöholm, P., and Witkowska, H. E. (1996) Carboxymethylation of the human estrogen receptor ligand-binding domain-estradiol complex: HPLC/ESMS peptide mapping shows that cysteine 447 does not react with iodoacetic acid, *Steroids 61*, 367−373.
32. Witkowska, H. E., Carlquist, M., Engström, O., Carlsson, B., Bonn, T., Gustafsson, J. Å., and Shackleton, C. H. L. (1997) Characterization of bacterially expressed rat estrogen receptor β ligand binding domain by mass spectrometry: Structural comparison with estrogen receptor α, *Steroids 62*, 621−631.
33. Goldstein, S. W., Bordner, J., Hoth, L. R., and Geoghegan, K. F. (2001) Chemical and biological issues related to X-ray crystallography of the ligand-binding domain of estrogen receptor α, *Bioconjugate Chem. 12*, 406−413.
34. Grese, T. A., et al. (1997) Structure−activity relationships of selective estrogen receptor modulators: modifications to the 2-arylbenzothiophene core of raloxifene, *J. Med. Chem. 40*, 146−167.
35. Bourguet, W., Germain, P., and Gronemeyer, H. (2000) Nuclear receptor ligand-binding domains: three-dimensional structures, molecular interactions and pharmacological implications, *Trends Pharm. Sci. 21*, 381−388.

BI0205092