Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Electronic Supplementary Information

Monitoring ligand-mediated helix 12 transitions within the human estrogen receptor α using bipartite tetracysteine display

Ranju Pokhrel[1], Tang Tang[1] and Justin M. Holub[1,2,3]*
[1]Department of Chemistry and Biochemistry, [2]Edison Biotechnology Institute, [3]Molecular and Cellular Biology Program, Ohio University, Athens, OH 45701

Reagents and chemicals. Tris, bisacrylamide, sodium dodecyl sulfate (SDS), tetramethyl-ethylenediamine (TEMED), imidazole, LB agar, sodium chloride, potassium chloride (KCl), sodium chloride (NaCl), ethylenediaminetetraacetic acid (EDTA), lysozyme and ampicillin were purchased from Thermo Fisher Scientific (Waltham, MA). LB-medium, β-mercaptoethanol (BME), isopropyl β-D-1-thiogalactopyranoside (IPTG), bovine serum albumin (BSA), 1-butanol and Coomassie Brilliant Blue R-250 were obtained from MP biomedicals (Santa Ana, CA). Fulvestrant was purchased from MedChem Express (Monmouth Junction, NJ). Ni-NTA agarose resin was obtained from Molecular Cloning Laboratories (San Francisco, CA). Precision Plus protein marker was obtained from Bio-Rad (Hercules, CA). Q5 Hot Start High-Fidelity 2× Master Mix, DpnI and DH5-α competent *E. coli* (high efficiency) cells were purchased from New England Biolabs (Ipswitch, MA). Bacterial protein extraction lysis buffer (PE LB) and yeast PE LB were obtained from G-Biosciences (St. Louis, MO). Tris(2-carboxyethlyl)phosphine (TCEP) was purchased from Gold Biotechnology (St Louis, MO). Steriflip vacuum filter and 0.22 μm sterile syringe filters were ordered from Millipore (Darmstadt, Germany). 17-β estradiol (E_2) and raloxifene were obtained from Cayman Chemical Co. (Ann Arbor, MI). Dithiothreitol (DTT) was purchased from Gold Biotechnologies (St. Louis, MO). PyClock was purchased from Millipore (Burlington, MA) Sodium orthovanadate (Na_3VO_4), glycerol, polyvinylpyrrolidone (PVP), ammonium persulfate (APS), DNAse, protease inhibitor cocktail, arsenic(III)chloride, N,N-diisopropylethylamine (DIEA), *N*-methyl-2-pyrrolidone (NMP), piperidine, 2-mercaptoethanol, triisopropylsilane (TIPS), N-methylmorpholine (NMM), acetic anhydride, palladium(II)acetate, phosphorous pentoxide, 2-mercaptoethane-sulfonate (MES), 4-morpholinepropanesulfonicacid (MOPS), 4-hydroxy-tamoxifen, bazedoxifene acetate and trichloroacetic acid were obtained from Sigma Aldrich (St. Louis, MO). BL21(DE3) cells were purchased from Agilent Life Technologies (Santa Clara, CA). All Fmoc-protected amino acids, Fmoc-PAL-AM resin and coupling reagents were obtained from Novabiochem (Billerica, MA). 1H-Benzotriazolium-1-[bis-(dimethylamino)-methylene]-5-chloro-hexafluorophosphate-(1-),3-oxide (HCTU) was obtained from Peptides International (Louisville,

KY). Glycine and chloramphenicol were obtained from EMD Millipore (Billerica, MA). Phenol was purchased from Corning Inc. (Corning, NY). Potassium chloride, 2-propanol, acetone, potassium phosphate monobasic, potassium phosphate dibasic, magnesium sulfate anhydrous, toluene, ethyl acetate, sand, silica gel, TLC plates and dichloromethane (DCM) were obtained from Fisher Scientific (Fair Lawn, NJ). Fluorescein, mercuric(II) oxide, and trifluoroacetic acid (TFA) were purchased from Acros Organics (Morris Plains, NJ). Acetonitrile (ACN) was purchased from NeoBits (Sunnyvale, CA). 1,2-ethanedithiol (EDT) was purchased from Fluka (Mexico City, Mexico). Chloroform was purchased form Pharmco (Brookfield, CT). Dimethyl sulfoxide (DMSO) and ethynyl estradiol were purchased from Santa Cruz Biotechnology (Dallas, TX). Sodium hydroxide (NaOH) and arsenic (III) chloride were purchased from VWR International (Radnor, PA). All primers used for site-directed mutagenesis were obtained from Integrated DNA Technologies (Coralville, IA). All other reagents were purchased from commercial sources and used without further purification unless otherwise stated.

Synthesis of 4',5'-bismercuric trifluoroacetate. The synthesis of 4',5'-bismercuric trifluoroacetate was performed using previously described protocols[1]. 500 mg (1.5 mmol) fluorescein was reacted with 2 equivalents of HgO in 7.5 ml of TFA. The resulting mixture was stirred at room temperature for 5 hours which resulted in an orange solution containing a yellow precipitate. The TFA was then evaporated under reduced pressure and the precipitate was washed with water, filtered and allowed to air dry. The obtained crude fluorescein 4',5'-bismercuric trifluoroacetate was stored in an evacuated desiccator overnight with 20 g of phosphorous pentoxide. The overall yield of 4',5'-bismercuric trifluoroacetate was 1102 mg (77%).

Synthesis of FlAsH-EDT$_2$. 480 mg (0.5 mmol) of 4',5' bismercuric trifluoroacetate was added to 20 equivalents of $AsCl_3$ and 8 equivalents of DIEA in dry NMP with a small amount of palladium acetate added as a catalyst. This reaction mixture was heated to 60 °C for 3 hours under an inert (N_2) atmosphere. The reaction was then quenched with 1:1 (v/v) 0.25 M potassium phosphate-acetone (pH 6.9) to give an orange solution. 2 ml of EDT was then transferred to the mixture to stabilize FlAsH-EDT$_2$. 20 ml of chloroform was added, and the organic layer was separated and dried over anhydrous magnesium sulfate. The drying agent was removed after 30 minutes by filtration and the chloroform was removed under reduced pressure. The resultant oil was dissolved in toluene and washed with water to remove traces of residual NMP. The organic layer was then separated and dried over anhydrous magnesium sulfate. Following drying, the solution was filtered and concentrated to give an orange oil. The crude product was purified by column

chromatography (silica gel) and eluted with 20% (v/v) ethyl acetate in toluene. Purified fractions were combined, and the solvent was removed by evaporation. The final product was washed with cold ethanol and filtered to give an off-white solid. Final yield of FlAsH-EDT_2 was 72 mg (9.5%).

NMR Analysis of FlAsH-EDT_2. 1H-NMR: (500 MHz, $CDCl_3$) 8.00 (d, 1H, H-4), 7.69 (t, 1H, H-5), 7.62 (t, 1H, H-6), 7.21 (d, 1H, H-7), 6.61 (d, 2H, H-1'), 6.52 (d, 2H, H-2') 3.6 (m, 8H, SCH2CH2S). The appearance of a multiplet around 3.6 ppm indicated the successful addition of two EDT molecules per FlAsH-EDT_2 molecule. All NMR experiments were carried out on a 500 MHz Bruker Avance spectrometer and processed using Bruker Topspin version 3.2.

Synthesis of control peptides. All control peptides used for FlAsH-EDT_2 binding assays were synthesized on Fmoc-PAL-AM resin (Novabiochem, Billerica, MA) using standard Fmoc-based solid-phase peptide synthesis procedures[2, 3]. Peptides were synthesized on a 50 μmol scale that was based on the resin loading level. All coupling and deprotection reactions were performed in a microwave-accelerated reaction system (CEM, Mathews, NC) using software programs written in-house. Fmoc deprotections were performed by adding 3 ml of 25% piperidine in NMP supplemented with 0.1 M hydroxybenzotriazole (HOBt)[4] to the resin-bound peptides. All amino acids were coupled through the addition of 5 equivalents (eq) of amino acid, 5 eq of PyClock and 10 eq of DIEA in NMP to the deprotected peptides. Washing cycles of 3 ml NMP (5×) were conducted between each deprotection and coupling step described above. Once a complete peptide sequence was obtained, the resin-bound peptides were acetylated at their N-terminus by adding 6% (v/v) NMM and 6% (v/v) acetic anhydride in NMP and the mixture was incubated at room temperature for 15 minutes with continuous stirring. Following acetylation, the resin-bound peptides were washed 2× with NMP and 3× with DCM and dried under vacuum. The peptides were then globally deprotected and cleaved from the resin using a cleavage cocktail composed of 88% TFA, 5% water, 5% phenol and 2% TIPS (v/v/v/v) for 30 minutes at 37 °C in a microwave reactor. Following cleavage, the peptides were precipitated in cold diethyl ether and pelleted by centrifugation. Precipitated peptides were then resuspended in 15% (v/v) ACN in water, frozen and lyophilized to dryness. Crude peptide powders were stored at -20°C until purification.

Purification of peptides by RP-HPLC. For purification, crude peptide powders were re-suspended in 5% aqueous ACN (v/v) and purified using a ProStar HPLC system (Agilent) across a reversed-phase semi-preparative C18 column (Grace, 10 μm, 250 mm × 10 mm). A linear gradient of 15-50% solvent B (0.1% TFA in ACN) over solvent A (0.1% TFA in water) at a flow

rate of 4 mL/min was used for all purifications. All major peaks were collected and analyzed by electrospray ionization (ESI) mass spectrometry to confirm product identity. Purified peaks containing the desired peptide were combined, frozen and lyophilized twice. All peptides were stored as dry powders at -20°C until further use.

Peptide characterization The purity of the peptides was determined by analytical RP-HPLC. Briefly, peptide products were dissolved in HPLC grade water to a final concentration of 2.5 μM. The dissolved peptides were then analyzed across a reversed-phase analytical C18 column (Grace, 5 μm, 50 × 2.1 mm) and eluted over 20 min using a linear gradient of 5-95% solvent B (0.1% TFA in ACN) over solvent A (0.1% TFA in water). All peptides were purified to >95% purity as determined by product peak integration of analytical HPLC chromatograms (Figure S1). Peptide identities were confirmed using a Q Exactive Plus Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Scientific) in the *m/z* range of 500-2200. For mass analysis, peptides were dissolved in 500 μL 10% ACN (v/v) in water and injected to the system at a flow rate of 10 μL/min. Mass data were collected and processed using Xcalibur v3.0 (Thermo Scientific) and MagTran v1.0 (Amgen). Calculated and observed masses of the purified peptides are outlined in Table S1.

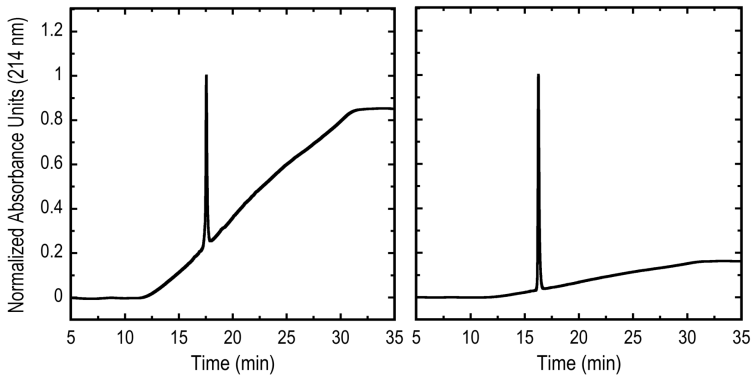

Figure S1. Analytical HPLC chromatograms of control peptides used in this work. Chromatogram on the left represents the positive control peptide and the chromatogram on the right represents the negative control peptide. All spectra were monitored at 214 nm. See text for details of analytical parameters.

Table S1. Calculated and observed masses for control peptides used in this work.

Peptide	**Sequence**	**Calc. mass (m/z)**	**Obs. mass (m/z)**
Positive Control	Ac-FLNCCPGCCMEP-NH_2	1358.69	1358.49
Negative Control	Ac-FLNAAPGAAMEP-NH_2	1230.43	1229.62

FlAsH-EDT$_2$ binding assays using control peptides. Fluorescent binding experiments were performed using the positive (Ac-FLNCCPGCCMEP-NH2) and negative (Ac-FLNAAPGAAMEP-NH2) control peptides to test profluorescent nature of FlAsH-EDT$_2$. Here, the kinetics of FlAsH-EDT$_2$ binding to the control peptides was monitored as described previously[5]. Briefly, a buffer solution composed 10 mM MES and 100 mM MOPS pH 7.2 was prepared. 10 mM stock solution of EDT and 1 mM stock solution of FlAsH-EDT$_2$ were prepared in DMSO and diluted in the MOPS buffer. EDT and FlAsH-EDT$_2$ were added to the buffer at final concentrations of 10 μM and 1 μM respectively. The binding of FlAsH-EDT$_2$ was initiated by adding the peptides to the previously prepared solution at a final concentration of 40 μM. Fluorescence emission of FlAsH-peptide complex was monitored using a Fluoromax 2 fluorimeter (Horiba Scientific, CA). Fluorescence counts were then collected for 1 hour (with a sampling interval of one minute) at 530 nm with an excitation at 508 nm (Figure S2). Data were processed using Kaleidagraph version 4.5 (Synergy Software) and the data were fit to a single exponential (association kinetics for one concentration of ligand) using GraphPad Prism version 8 (GraphPad Software). For FlAsH complexed with the positive control peptide, the association rate (k_{on}) was determined to be 147,000 M^{-1} min^{-1}. Association rate was not determined for the negative control peptide.

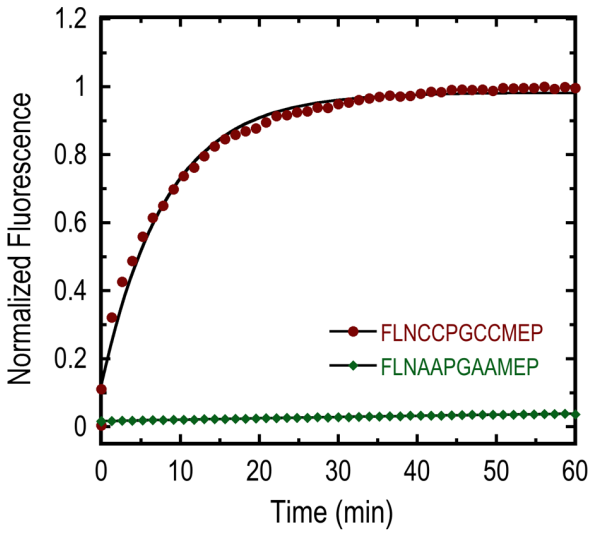

Figure S2. Kinetics of association of FlAsH-EDT$_2$ and control peptides. Graph shows time course of the reaction of 1 μM FlAsH-EDT$_2$ with 40 μM peptide in binding buffer (10 mM MES, 100 mM MOPS, pH 7.2). The increase in fluorescence intensity at 530 nm (excitation at 508 nm) is shown for a solution of peptide upon treatment with 1 μM FlAsH-EDT$_2$ at time zero. The experimental fits shown are to single exponentials (association kinetics for one concentration of ligand).

Plasmid preparation and transformation. Plasmid vectors (pMCSG7) coding for His_6-tagged wild-type estrogen receptor α ligand binding domain (wtERα-LBD) were generous gifts provided by Dr. Kendall Nettles (Scripps Research Institute, Jupiter, FL). Immediately upon arrival in our laboratory, the plasmids were sequenced to determine the primary structure of the wtERα-LBD (Table S2). For storage and expression, all plasmids were transformed into BL21(DE3) chemically-competent cells according to the manufacturer's guidelines. Following transformation, the cells were plated on LB agar plates containing 1× ampicillin and grown overnight at 37 °C. A single isolated colony was then picked and grown in 5 ml LB medium containing 1× ampicillin at 37 °C overnight with constant shaking at 225 rpm. Bacterial transformants were stored as glycerol stocks prepared from 500 μl inoculum, 250 μl glycerol and 250 μl water. All glycerol stocks were stored at -80°C until further use.

Site-directed mutagenesis. Plasmids coding for ERα-LBD mutants used in this work were generated through site-directed mutagenesis using a protocol provided along with Q5 Hot Start High-Fidelity Master Mix (New England Biolabs). Briefly, 1.25 μl of forward primer (10 ng/ml), 1.25 μl of reverse primer (10 ng/ml), 2 μl of template DNA (50 ng/ml), 12.5 μl master mix and 8 μl nuclease-free water were subjected to 35 cycles of PCR amplification using the following reaction conditions: denaturation at 92 °C (30 seconds), annealing at 72 °C (30 seconds), and primer extension at 72 °C (20-30 seconds Kb^{-1}). Following the PCR reaction, 1 μl of DpnI was added to the mixture to digest parent DNA and the sample was incubated at 37 °C for 1 hour. The amplified products were then transformed into DH5-α competent cells (New England Biolabs) using a transformation protocol provided by the manufacturer. Following transformation, the cells were plated on LB agar plates containing 1× ampicillin and incubated at 37 °C overnight. A single isolated colony was then picked and grown in 25 ml LB medium supplemented with 1× ampicillin. The next day, cell pellets were harvested from the overnight inoculum and plasmids were purified using a Midi Prep Plasmid Extraction Kit (Qiagen) according to the manufacturer's guidelines. All plasmids were sequenced, and the resulting DNA sequences were translated using the web-based Expasy translate tool (http://web.expasy.org/translate). Positive mutations were identified (Table S2) and the plasmids were subsequently transformed into chemically competent BL21(DE3) cells according to the manufacturer's instructions. All transformants were stored as glycerol stocks at -80 °C as described above. ERα-LBD mutants were generated following the flow chart shown in Figure S3 using primers outlined in Table S3.

Table S2. Sequence and nomenclature of ERα-LBD mutants; mutated residues are shown in red.

Protein	Mutations	Protein Sequence (mutation sites are shown in red)
wtERα-LBD	None	MHHHHHHSSG VDLGTENLYF QSNAMKRSKK NSLALSLTAD QMVSALLDAE PPILYSEYDP TRPFSEASMM GLLTNLADRE LVHMINWAKR VPGFVDLTLH DQVHLLECAW LEILMIGLVW RSMEHPGKLL FAPNLLLDRN QGKCVEGMVE IFDMLLATSS RFRMMNLQGE EFVCLKSIIL LNSGVYTFLS STLKSLEEKD HIHRVLDKIT DTLIHLMAKA GLTLQQQHQR LAQLLLILSH IRHMSNKGME HLYSMKCKNV VPLYDLLLEM LDAHRLHAPT
ERα-LBD-ΔC$_0$	C381S; C417S; C530A	MHHHHHHSSG VDLGTENLYF QSNAMKRSKK NSLALSLTAD QMVSALLDAE PPILYSEYDP TRPFSEASMM GLLTNLADRE LVHMINWAKR VPGFVDLTLH DQVHLLESAW LEILMIGLVW RSMEHPGKLL FAPNLLLDRN QGKSVEGMVE IFDMLLATSS RFRMMNLQGE EFVCLKSIIL LNSGVYTFLS STLKSLEEKD HIHRVLDKIT DTLIHLMAKA GLTLQQQHQR LAQLLLILSH IRHMSNKGME HLYSMKAKNV VPLYDLLLEM LDAHRLHAPT
ERα-LBD-ΔCC$_N$	E380C; C417S; C530A	MHHHHHHSSG VDLGTENLYF QSNAMKRSKK NSLALSLTAD QMVSALLDAE PPILYSEYDP TRPFSEASMM GLLTNLADRE LVHMINWAKR VPGFVDLTLH DQVHLLCCAW LEILMIGLVW RSMEHPGKLL FAPNLLLDRN QGKSVEGMVE IFDMLLATSS RFRMMNLQGE EFVCLKSIIL LNSGVYTFLS STLKSLEEKD HIHRVLDKIT DTLIHLMAKA GLTLQQQHQR LAQLLLILSH IRHMSNKGME HLYSMKAKNV VPLYDLLLEM LDAHRLHAPT
ERα-LBD-ΔCC$_C$	C381S; C417S; C530A; A546C; H547C	MHHHHHHSSG VDLGTENLYF QSNAMKRSKK NSLALSLTAD QMVSALLDAE PPILYSEYDP TRPFSEASMM GLLTNLADRE LVHMINWAKR VPGFVDLTLH DQVHLLESAW LEILMIGLVW RSMEHPGKLL FAPNLLLDRN QGKSVEGMVE IFDMLLATSS RFRMMNLQGE EFVCLKSIIL LNSGVYTFLS STLKSLEEKD HIHRVLDKIT DTLIHLMAKA GLTLQQQHQR LAQLLLILSH IRHMSNKGME HLYSMKAKNV VPLYDLLLEM LDCCRLHAPT
ERα-LBD-ΔC$_4$	E380C; C417S; C530A; A546C; H547C	MHHHHHHSSG VDLGTENLYF QSNAMKRSKK NSLALSLTAD QMVSALLDAE PPILYSEYDP TRPFSEASMM GLLTNLADRE LVHMINWAKR VPGFVDLTLH DQVHLLCCAW LEILMIGLVW RSMEHPGKLL FAPNLLLDRN QGKSVEGMVE IFDMLLATSS RFRMMNLQGE EFVCLKSIIL LNSGVYTFLS STLKSLEEKD HIHRVLDKIT DTLIHLMAKA GLTLQQQHQR LAQLLLILSH IRHMSNKGME HLYSMKAKNV VPLYDLLLEM LDCCRLHAPT

Table S3. Primers used for the generation of ERα-LBD mutants used in this work.

Set	Direction	Primers
1	Forward	5′-atctgtacagcatgaaggccaagaacgtggtgcccc-3′
	Reverse	3′-tagacatgtcgtacttccggttcttgcaccacgggg-5′
2	Forward	5′-ggacaggaaccagggaaaaagtgtagagggc-3′
	Reverse	3′- cctgtccttggtccctttttcacatctcccg-5′
3	Forward	5′-tgatcaggtccaccttctagaaagtgcctggctag-3′
	Reverse	3′-actagtccaggtggaagatctttcacggaccgatc-5′
4	Forward	5′-ccctccatgatcaggtccaccttctatgctgtgcctggctaga-3′
	Reverse	3′-gggaggtactagtccaggtggaagatacgacacggaccgatct-5′
5	Forward	5′-gctggagatgctggactgctgccgcctacatgcgccc-3′
	Reverse	3′-cgacctctacgacctgacgacggcggatgtacgcggg-5′

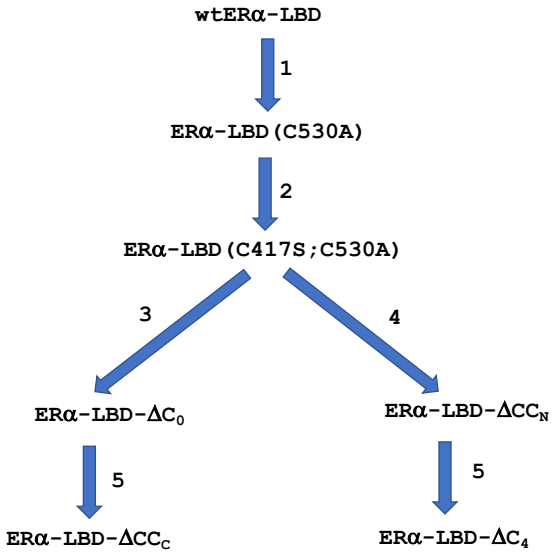

Figure S3. Workflow used to generate ERα-LBD mutants. Plasmids coding for wtERα-LBD were first subjected to site-directed mutagenesis to mutate the native Cys residue on helix 11 (C530) to an Ala using primer set 1. The native C417 of ERα-LBD(C530A) was then mutated to a Ser using primer set 2, generating ERα-LBD(C417S;C530A). Primer set 3 was used to mutate the final exposed native Cys (C318) of ERα-LBD(C417S;C530A) to a Ser, generating ERα-LBD-ΔC_0. Primer set 5 was reacted with ERα-LBD-ΔC_0 to generate the ERα-LBD-ΔCC_C mutant, which contains two Cys residues within helix 12 at positions 546 and 547. ERα-LBD(C417S;C530A) was also mutated with primer set 4 to generate the ERα-LBD-ΔCC_N mutant, which contains two Cys residues at positions 380 and 381. ERα-LBD-ΔC_4 was generated from ERα-LBD-ΔCC_N using primer set 5 and contains Cys residues at 380, 381, 546 and 547 to monitor H12 transitions.

Protein expression and purification. All ERα-LBD proteins used herein were expressed and purified using a modification of the technique reported by Eiler, *et al*[6]. Protein expression and purification was initiated by streaking BL21(DE3) transformants from glycerol stocks on an LB agar plate containing 1× ampicillin and incubating the plate at 37 °C overnight. Following overnight incubation, a single isolated colony was picked and used to inoculate 5 ml LB media containing 1× ampicillin. The inoculum was incubated at 37 °C overnight with constant shaking at 225 rpm. This starter culture was then used to inoculate 1 L LB media containing 1× ampicillin and was incubated at 37 °C with constant shaking at 225 rpm. Bacterial growth was monitored by measuring the optical density (OD) at 600 nm on a UV/Vis spectrophotometer (Cary). Once the

OD reached 0.6, the bacterial culture was cooled to 16 °C with constant shaking at 225 rpm for 20 minutes. Protein expression was then induced by adding IPTG to the LB medium at a final concentration of 0.1 mM. The culture was then incubated at 16 °C with constant shaking at 225 rpm for 18 hours. Following incubation, the cells were harvested by centrifugation and the pellets from 250 ml culture volumes were stored at -80 °C until further use.

All storage, transfer and centrifuge tubes used during the purification process of ERα-LBD were coated with a 1% polyvinylpyrrolidone (w/v) solution overnight. Bacterial pellets from 250 mL volumes of induced culture were thawed on ice for 30 minutes and 8.75 mL of lysis buffer (3:1 Bacterial PE:Yeast PE (v/v), 50 U DNAse I, 2.5 mg lysozyme, 10 mM imidazole, 2 M urea, 10 mM 2-mercaptoethanol, 25 μL protease inhibitor cocktail and 10% glycerol, pH 8.0) was added to resuspend the bacterial pellets. The resuspended cells were then allowed to incubate in lysis buffer at room temperature for 20 minutes. Following lysis, the samples were centrifuged at 40,000 $\times$ *g* for 45 minutes. The supernatant was then transferred to a clean centrifuge tube and centrifuged again at 40,000 $\times$ *g* for an additional 30 minutes. Following centrifugation, the supernatant was filtered using a 0.22 μm vacuum filtration unit (Steriflip). The His_6 affinity column was prepared by loading 2 mL Ni-NTA resin into a fritted 10 ml PPE column (Thermo) and washing the resin 2$\times$ with 10 ml loading buffer (3:1 Bacterial PE:Yeast PE (v/v), 10 mM imidazole, 2 M urea, 1 mM BME, pH 8.0). Following washing, the cleared supernatant was added to the prepared column and incubated with end-over-end rotation for 45 minutes at 4 °C. Following incubation, the proteins were washed 5$\times$ with 10 ml of wash buffer (50 mM Tris base, 500 mM NaCl, 20 mM imidazole, 15 mM BME and 10% glycerol, pH 8.0). The proteins were then eluted from the column by adding elution buffer (50 mM Tris, 500 mM NaCl, 750 mM imidazole, 15 mM BME and 10% glycerol, pH 8.0) and collecting 3 ml fractions (5$\times$). Purified fractions were combined and immediately dialyzed into either binding buffer (50 mM Tris base, 500 mM KCl, 2 mM DTT, 1 mM EDTA, 1 mM Na_3VO_4 and 10% glycerol pH 8.0) for FlAsH binding assays or folding buffer (20 mM Tris, 150 mM NaCl, pH 8.0) for structural analysis. All proteins were dialyzed for 48 hours at 4°C with one buffer change after 24 hours. Following dialysis, the proteins were concentrated using centrifugal filtration units (Amicon). Purified proteins were concentrated to approximately 15 μM as quantified using a Bradford assay[7]. The purity of proteins was accessed by separating them on a 14% polyacrylamide gel by SDS-PAGE and staining with Coomassie blue (Figure S4, insets). The concentrated proteins were then aliquoted into prechilled microfuge tubes, rapidly frozen and stored at -80°C until further use.

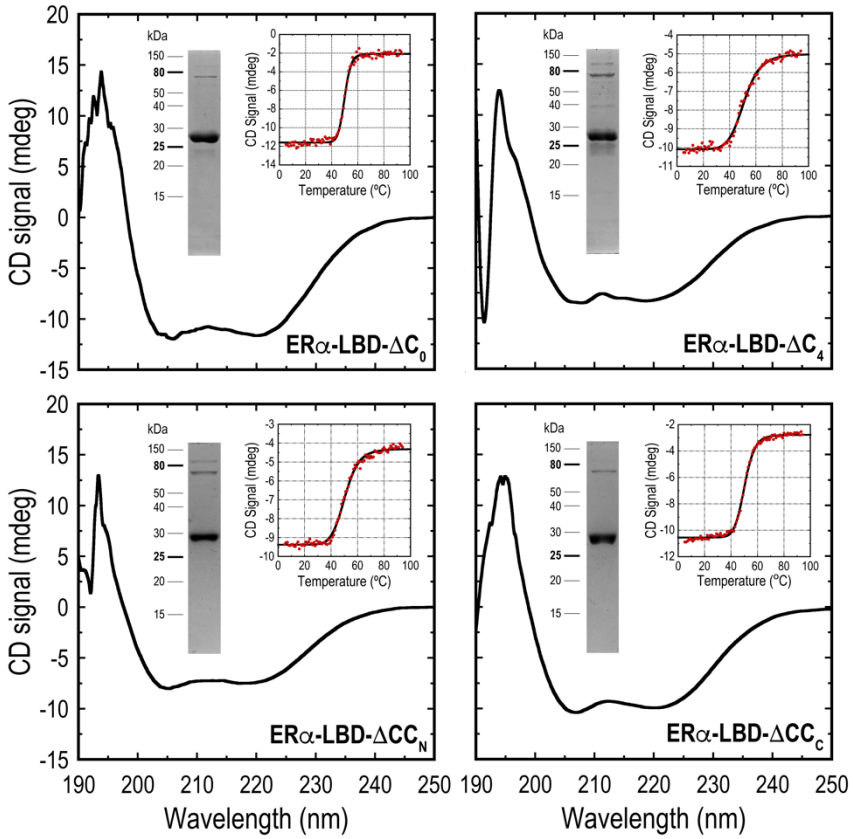

Figure S4. Circular dichroism spectra and SDS-PAGE of recombinant ERα-LBD proteins used in this work. All spectra show far-UV CD signatures typical of proteins with predominantly α-helical character. Graphical insets show thermal unfolding of proteins monitored at 222 nm plotted as a function of increasing temperature. SDS-PAGE gels (insets) show a single band visible at the expected molecular weight of the recombinant ERα-LBD proteins.

Circular dichroism spectropolarimetry. Wavelength-dependent circular dichroism (CD) spectra for all proteins were recorded using a Jasco J-715 circular dichroism spectropolarimeter interfaced with a Peltier temperature control unit. All CD data were obtained from solutions containing 3.5 μM protein in folding buffer (20 mM Tris and 150 mM NaCl, pH 8.0) at 20 °C. Protein solutions were allowed to incubate at 20 °C for 10 minutes before being analyzed. Scans were performed from 190-250 nm and spectra represent the averaged accumulation of four scans (Figure S4). Temperature-dependent circular dichroism (CD) spectra for all proteins were recorded using a Jasco J-715 CD spectropolarimeter interfaced with a Peltier temperature control unit. Melt data were obtained from solutions containing 3.5 μM protein in folding buffer (20 mM Tris and 150 mM NaCl, pH 8.0) from 5-95 °C. Protein solutions were allowed to incubate at 5 °C

for 10 minutes before being analyzed. Thermal melts were generated by monitoring the CD of each protein at 222 nm as the temperature was increased from 5 to 95 °C at 1 °C min^{-1} intervals (Figure S4, insets). Melting temperatures (Table S4) were determined using equation 1 below. All CD spectra were generated using J-700 Software version 1.5 (Jasco) and data was processed using Kaleidagraph version 4.5 (Synergy Software).

$$(1) \qquad CD_{obs} = \frac{(CDmax - CDmin)}{\left[1+\left(\frac{T}{Tm}\right)^{n}\right]} + CD_{min}$$

where CD_{obs} is the observed CD signal in mdeg, CD_{max} is the maximum CD value, CD_{min} is the minimum CD value, T is the temperature and Tm is the temperature at half maximal CD signal. The steepness of the linear part of the curve is described by the slope factor, n.

Competitive binding assay. The binding affinity of the ERα-LBD for its natural ligand 17-β estradiol (E2) was evaluated using a PolarScreen™ Estrogen Receptor Alpha Competitor Assay, Red (Life Technologies). Fluorescence polarization (FP) experiments were performed in 384-well plates (#3363, Corning) using a Cytation3 plate reader (BioTek Instruments Inc.) outfitted with a red FP filter cube (530 nm/590 nm). For this assay, ERα-LBD proteins were diluted in ERα screening buffer to a final concentration of 77 nM and pre-incubated with Fluormone™ tracer (1.5 nM) for 45 minutes. Following the incubation, E2 was serially diluted and added to the pre-bound protein-Fluormone complex. The reaction mixture was then allowed to incubate for an additional 2 hours at room temperature. Data were collected on Gen5 software (BioTek) and processed using Kaleidagraph version 4.5 (Figure S5). The IC_{50} values were determined using equation 2:

$$(2) \qquad FP_{obs} = \frac{FP_{max} - FP_{min}}{\left[1+\left(\frac{L}{IC_{50}}\right)^{m}\right]} + FP_{min}$$

where FP_{obs} is the observed polarization value, FP_{max} is the maximum polarization value, FP_{min} is the minimum polarization value, L is the ligand concentration and IC_{50} is the half maximal inhibitory concentration. The steepness of the linear part of the curve is described by the slope factor, m.

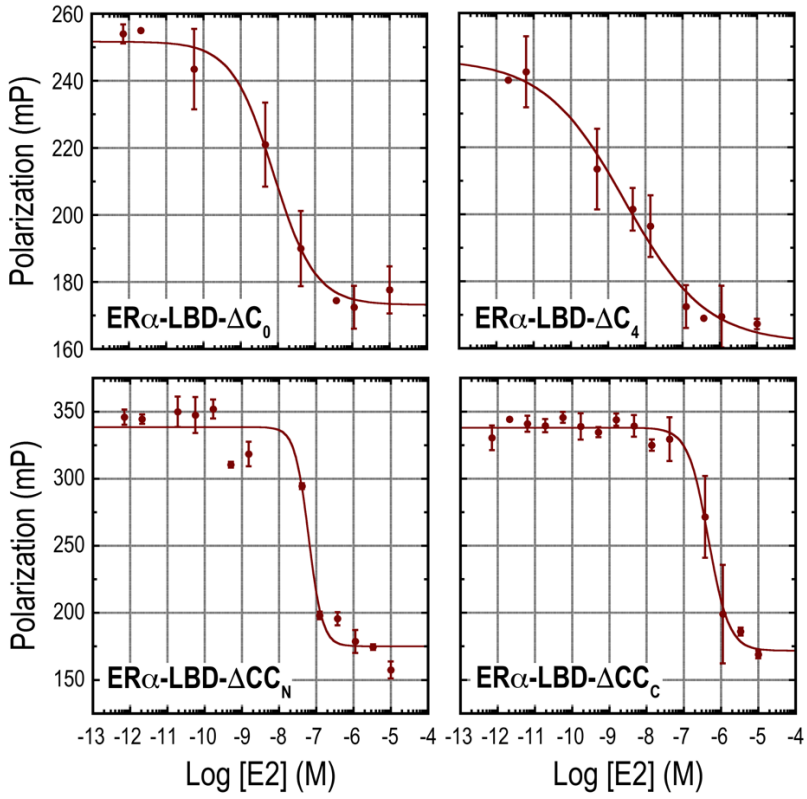

Figure S5. Competitive binding assays for ERα-LBD mutants used in this work. Results show fluorescence polarization in millipolarization units (mP) plotted against concentration of natural ligand (E2). Data are shown for each concentration as a mean of three independent experiments. Error bars represent standard deviation.

Table S4. Thermal stability and binding data for ERα-LBD proteins.

Protein	Tm (°C)	IC_{50} (nM)
ERα-LBD-ΔC_0	49.8	7.56
ERα-LBD-ΔCC_N	50.4	68.1
ERα-LBD-ΔCC_C	50.2	456.3
ERα-LBD-ΔC_4	51.0	2.80

Fluorescence endpoint experiments. The ability for our ERα-LBD proteins to bind FlAsH-EDT_2 in the presence or absence of ligand was evaluated using fluorescence intensity measurements. For these experiments, 25 μL volumes of 10 μM receptor were incubated in binding buffer supplemented with 1 mM TCEP overnight at room temperature in individual wells of a 384-well black, flat-bottom plate (#3575, Corning) to ensure full reduction of Cys thiols[8]. Following

overnight incubation, binding buffer alone (no ligand) or binding buffer containing ligand (10 μM) was added to the reduced proteins in triplicate and the samples were allowed to incubate at room temperature in the dark for two hours. Prior to addition, each ligand was dissolved in 100% DMSO and further diluted in 20% DMSO (v/v) in binding buffer. Aliquots of EDT and FlAsH-EDT_2 in binding buffer were then added to the samples at final respective concentrations of 10 μM and 1 μM and the samples were allowed to incubate in the dark for 24 hours. The fluorescence intensity of each solution was then measured using a SpectraMax M5e multi-mode plate reader (Molecular Devices), with an excitation wavelength of 508 nm and an emission wavelength of 530 nm set to an automatic cutoff of 530 nm. Data were processed using Kaleidagraph version 4.5 (Synergy Software).

SDS-PAGE of FlAsH-protein complexes. The ability for FlAsH to label and crosslink ERα-LBD-ΔC_4 proteins was assessed using in-gel fluorescence (Figures S6 and S7). For these studies, 25 μL volumes of 10 μM ERα-LBD-ΔC_4 proteins were pre-incubated in binding buffer supplemented with 1 mM TCEP overnight at room temperature in individual wells of a 384-well black, flat-bottom plate (#3575, Corning) to ensure full reduction of Cys thiols. Following overnight incubation, ligand in binding buffer (10 μM) or binding buffer only (no ligand) was added to each sample and the mixtures were allowed to incubate at room temperature for two hours. Prior to addition, each ligand was dissolved in 100% DMSO and further diluted in 20% DMSO (v/v) in binding buffer. Following complexation, freshly-prepared solutions of EDT and FlAsH-EDT_2 in binding buffer were added to the wells at a final concentrations of 10 μM and 5 μM respectively. The reactions were allowed to incubate in the dark for 24 hours at room temperature. Following incubation, the samples were mixed with modified 5× SDS sample buffer (0.312 M Tris, 10% (w/v) SDS, 20 mM TCEP, 0.05 % (w/v) bromophenol blue and 40% (v/v) glycerol, pH 6.8) and loaded onto a 14% polyacrylamide gel for separation by SDS-PAGE. The separated proteins were then visualized under UV illumination on a Bio-Rad ECL/fluorescence gel imaging cabinet (Hercules, CA). Images were captured using the Pro-Q Emerald 300 protocol provided with Bio-Rad Image Lab software version 5.2.1. Staining of the gel with Coomassie blue indicated the presence of ERα-LBD proteins. Quantification of band intensity was performed using ImageJ software.

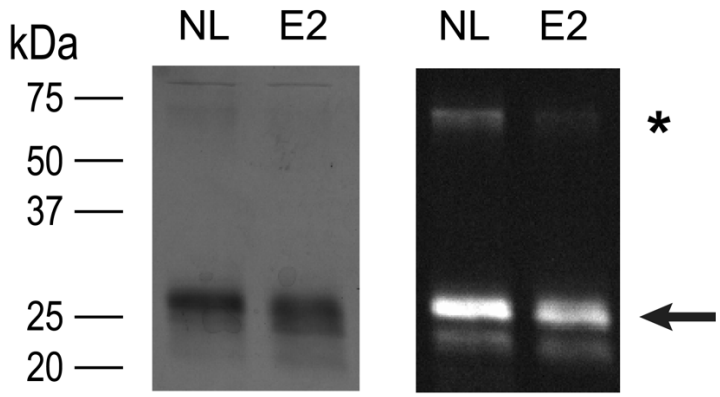

Figure S6. FlAsH complexed with ERα-LBD-ΔC_4 proteins visualized using in-gel fluorescence. Fully-reduced ERα-LBD proteins were incubated with or without ligand and complexed with FlAsH before being separated by SDS-PAGE. Protein-FlAsH complexes were visualized by staining with Coomassie blue (left images) or in-gel fluorescence (right images). Monomeric ERα-LBDs are indicated with an arrow and the position of crosslinked (dimeric) receptors are indicated with an asterisk (*). Size marker (kDa) is shown to the left of the gels. Treatments are indicated above each lane, NL: no ligand; E2: 17β-estradiol.

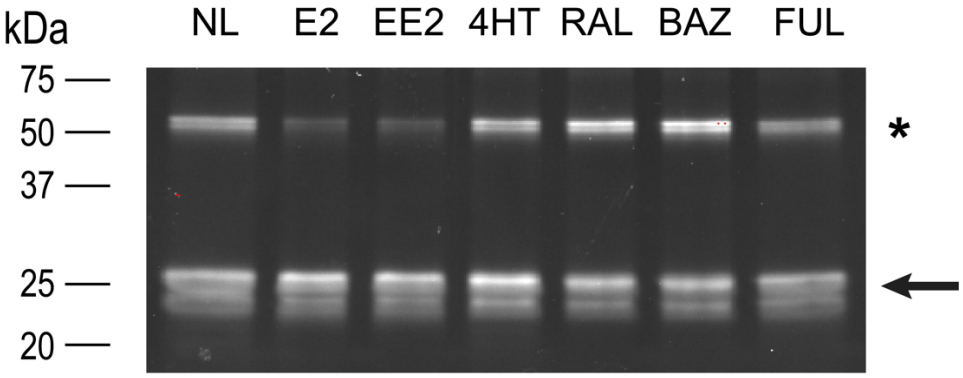

Figure S7. In-gel fluorescence of ERα-LBD-ΔC_4 proteins complexed with various ligands and FlAsH. Fully-reduced ERα-LBD proteins were incubated with or without ligand and complexed with FlAsH before being separated by SDS-PAGE. Protein-FlAsH complexes were visualized by in-gel fluorescence. Monomeric LBDs are indicated with an arrow and the position of crosslinked (dimeric) receptors are indicated with an asterisk (*). Size marker (kDa) is to the left of the gel. Treatments are indicated above each lane, NL: no ligand; E2: 17β-estradiol; EE2: ethynyl estradiol; 4HT: 4-hydroxytamoxifen; RAL: raloxifene; BAZ: bazedoxifene; FUL: fulvestrant.

Binding kinetics of FlAsH-EDT$_2$ to ERα-LBD-C$_4$. The association rates (K_{on}) of FlAsH-EDT$_2$ to ERα-LBD-ΔC$_4$ proteins was evaluated using fluorescence intensity measurements in the presence or absence of ligand. For these experiments, 25 μL volumes of 10 μM ERα-LBD-ΔC$_4$ were incubated in binding buffer supplemented with 1 mM TCEP overnight at room temperature in individual wells of a 384-well black, flat-bottom plate (#3575, Corning) to ensure full reduction of Cys thiols. Following overnight incubation, buffer alone (no ligand) or buffer containing (10 μM) ligand was added to the reduced proteins and the samples were allowed to incubate at room temperature in the dark for two hours. Prior to addition, each ligand was dissolved in 100% DMSO and further diluted in 20% DMSO (v/v) in binding buffer. Aliquots of EDT and FlAsH-EDT$_2$ in binding buffer were then added to the samples at final respective concentrations of 10 μM and 1 μM. The change in fluorescence intensity as a function of time of each solution was then immediately measured using a SpectraMax M5e multi-mode plate reader (Molecular Devices), with an excitation wavelength of 508 nm and an emission wavelength of 530 nm set to an automatic cutoff of 530 nm. Fluorescence counts were collected at room temperature over two hours at one-minute intervals. Data were processed using Kaleidagraph version 4.5 (Synergy Software) and the data was fit to a single exponential (association kinetics for one concentration of ligand) using GraphPad Prism version 8 (GraphPad Software).

Determination of FlAsH-EDT$_2$ apparent equilibrium dissociation constant (K_{app}). Equilibrium dissociation constants for FlAsH complexed with ERα-LBD-ΔC$_4$ in the presence and absence of ligands were determined by monitoring fluorescence intensity as a function of protein concentration. For these studies, all liganded ERα-LBD-ΔC$_4$ complexes were prepared as described for the endpoint experiments (*vida supra*). Following incubation with ligand, 50 μL of the protein alone or protein-ligand complex was transferred to individual wells of a polystyrene-coated 384-well black, flat-bottom plate (#3575, Corning, NY) and serially diluted by half into fresh binding buffer. Freshly-prepared solutions of EDT and FlAsH-EDT$_2$ in binding buffer were immediately added to the wells at final concentrations of 10 μM and 1 μM respectively. The reactions were then allowed to incubate in the dark for 24 hours at room temperature. Following incubation, the fluorescence of each sample was measured using a SpectraMax M5e multi-mode plate reader (Molecular Devices, Sunnyvale, CA), with an excitation wavelength of 508 nm and an emission wavelength of 530 nm set to an automatic cutoff of 530 nm. Data was processed using SoftmaxPro version 6.4 and graphs were generated using Kaleidagraph V.4.5 (Synergy Software) by plotting fluorescence intensities of each reaction as a function of protein

concentration. K_d values were determined from curve fits obtained from equation 3, the observed apparent K_d values represent an average ± standard deviation of the calculated apparent K_d of three or more independent trials.

$$(3) \qquad F_{obs} = F_{min} + (((F_{max} - F_{min})/(2L))(L + P + K_d - ((L + P + K_d)^2 - (4LP))^{0.5}))$$

where, F_{obs} represents the observed fluorescence at any total protein concentration P, L is the total concentration of FlAsH-EDT_2, F_{min} and F_{max} are the maximum and minimum fluorescence values, respectively. K_d is the apparent equilibrium dissociation constant.

Statistical analysis. For all data points, at least three independent measurements were taken and averaged for each receptor class in the presence or absence of ligand. All data points represent average of three independent experiments; error bars are plotted as standard deviations. A two-tailed Student's t-test was used to analyze statistical significance before normalizing the data and a 95% confidence interval was used to denote the significance.

References

1. Adams, S. R.; Tsien, R. Y., Preparation of the membrane-permeant biarsenicals FlAsH-EDT2 and ReAsH-EDT2 for fluorescent labeling of tetracysteine-tagged proteins. *Nat Protoc* **2008,** *3*, 1527-34.

2. Chang, C. D.; Meienhofer, J., Solid-phase peptide synthesis using mild base cleavage of N alpha-fluorenylmethyloxycarbonylamino acids, exemplified by a synthesis of dihydrosomatostatin. *Int J Pept Protein Res* **1978,** *11*, 246-9.

3. Fields, G. B.; Noble, R. L., Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int J Pept Protein Res* **1990,** *35*, 161-214.

4. Palasek, S. A.; Cox, Z. J.; Collins, J. M., Limiting racemization and aspartimide formation in microwave-enhanced Fmoc solid phase peptide synthesis. *J Pept Sci* **2007,** *13*, 143-8.

5. Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y., New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications. *J Am Chem Soc* **2002,** *124*, 6063-76.

6. Eiler, S.; Gangloff, M.; Duclaud, S.; Moras, D.; Ruff, M., Overexpression, purification, and crystal structure of native ER alpha LBD. *Protein Expr Purif* **2001,** *22*, 165-73.

7. Bradford, M. M., A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **1976,** *72*, 248-54.

8. Krishnan, B.; Gierasch, L. M., Cross-strand split tetra-Cys motifs as structure sensors in a beta-sheet protein. *Chem Biol* **2008,** *15*, 1104-15.