Structural Determinants of the Binding and Activation of Estrogen Receptor α by Phenolic Thieno[2,3-d]pyrimidines

Vamshikrishna Reddy Sammeta,[a] Brian Anderson,[a] John D. Norris,[b] Chad D. Torrice,[b] Carstyn Joiner,[c] Shubin Liu,[d] Haoxi Li,[e] Konstantin I. Popov,[f] Sean W. Fanning,[c] Donald P. McDonnell,[b] and Timothy M. Willson.*[a]

[a] *Structural Genomics Consortium, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC 27599, USA.*
[b] *Department of Pharmacology and Cancer Biology, Duke University School of Medicine, Durham, NC 27710, USA.*
[c] *Department of Cancer Biology, Loyola University of Chicago Stritch School of Medicine, Maywood, IL 60611, USA.*
[d] *Research Computing Center, University of North Carolina, Chapel Hill, NC 27599, USA.*
[e] *Laboratory for Molecular Modeling, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC 27599, USA.*
[f] *Center of Integrative Chemical Biology and Drug Discovery, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC 27599, USA.*

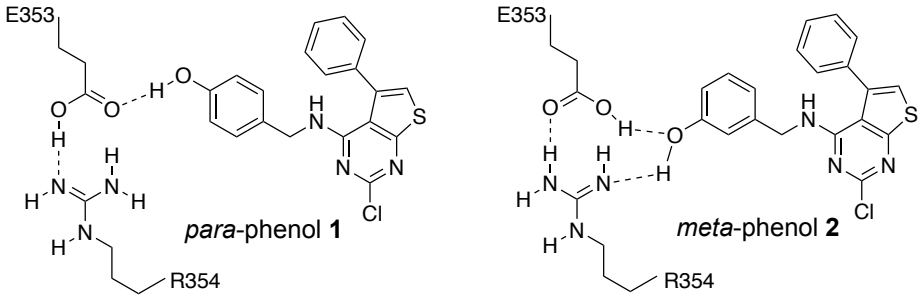
* tim.willson@unc.edu

Dedicated to an inspirational teacher and mentor *Scott E. Denmark* on the occasion of his 70th birthday and award of the Paracelsus prize.

Abstract

Synthetic, structural, and computational approaches were used to solve the puzzle as to how a phenolic nonsteroidal estrogen **1** with only a single H-bond to its receptor was more potent than an isomer **2** which formed an intricate network of H-bonds. Synthesis of a series of substituted phenols revealed that pK_a was not a determinant of estrogenic activity. First-principles calculation also failed to explain the difference in activity of **1** and **2**. Molecular dynamics revealed that **1** formed a more stable receptor complex compared to **2**, which may explain its increased activity despite forming fewer apparent H-bonds with the protein.

Graphical Abstract



Puzzle: **1** forms fewer H-bonds with ERα than **2**, but is a better estrogen

Keywords: estrogen receptor, density functional calculations, ligand effects, molecular dynamics, structure-activity relationships.

Introduction

We recently described a series of phenolic thieno[2,3-*d*]pyrimidines, such as **1** (Figure 1), that bind and activate the estrogen receptor α (ERα). These nonsteroidal estrogens were able to mimic the natural steroidal hormone 17β-estradiol (E2) with nanomolar affinities for ERα and potent agonist activity in reporter gene assays.[1] Surprisingly, the X-ray cocrystal structure of thieno[2,3-d]pyrimidine (**1**) bound to ERα showed only a single hydrogen bond between the phenolic oxygen and E353 of the receptor (see Figure 2A).[1] In contrast the phenolic oxygen of E2 forms an intricate network of hydrogen bonds with E353, R394, and a water molecule in the ERα ligand pocket that is essential for molecular recognition of steroidal estrogens.[2] Initial analysis of the binding site suggested that the isomeric *meta*-phenol (**2**) would be an improved estrogen mimetic, however it was shown to be much weaker for both receptor binding and agonist activity.[1] The result was especially surprising since the 2-isomer of the steroidal estrogen (Figure 1) that places a phenol in the same position in the ligand binding pocket as *para*-phenol (**1**) has been shown to have a 5-fold weaker affinity for ERα.[3] To resolve this puzzle and to further our understanding of how thieno[2,3-d]pyrimidines function as nonsteroidal estrogens we performed a series of structure-function, physiochemical, and computational studies to probe the interaction of the phenol pharmacophore with the receptor.

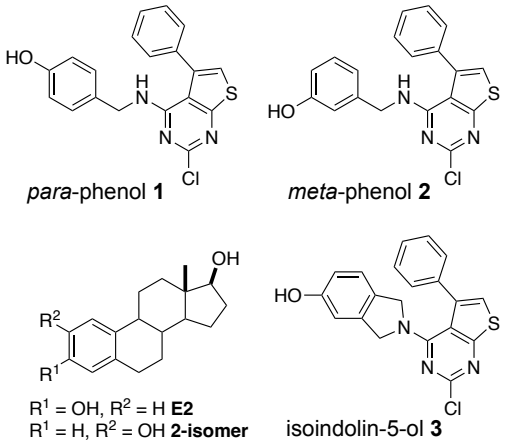


Figure 1. Thieno[2,3-d]pyrimidine ERα ligands

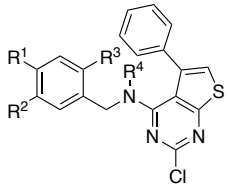
Results and Discussion

Differences in ERα activity of para- and meta-phenols

To confirm the differences in ERα activity, freshly prepared samples of **1** and **2** were tested in a two-hybrid ligand sensing assay (LiSA) that utilizes the recruitment of a specific peptide to the surface of the receptor to measure the ability of steroidal and nonsteroidal estrogens to interact

with the ligand-binding domain of ERα.[1, 4, 5] The assay confirmed our previous observation that *meta*-phenol (**2**) was >15-fold weaker than *para*-phenol (**1**) as an ERα ligand in the LiSA (Table 1), a result that was in part due to differences in binding affinity as shown in a radioligand competition binding assay using [^{3}H]-E2 (Figure S1). This difference in ERα binding activity in the LiSA translated to ~100-fold weaker agonist activity measured using a standard reporter gene assay in HepG2 cells, with *meta*-phenol (**2**) being a much poorer estrogen than *para*-phenol (**1**) (Table 1). A cocrystal structure of *meta*-phenol (**2**) with the ERα ligand binding domain (LBD) revealed that, despite its weaker estrogenic activity, the phenolic oxygen did indeed form hydrogen bonds with E353, R394, and a bound water molecule while the hydrophobic thieno[2,3-d]pyrimidine core occupied the same region of the ligand binding pocket as **1** (Figure 2B). Thus, although *meta*-phenol (**2**) was able to form an intricate network of hydrogen bonds with E353 and R394 that mimicked the binding of the natural steroid hormone E2 within the ERα ligand binding pocket, it was surprisingly less effective as a nonsteroidal estrogen than *para*-phenol (**1**) in both the LiSA and reporter gene assays.

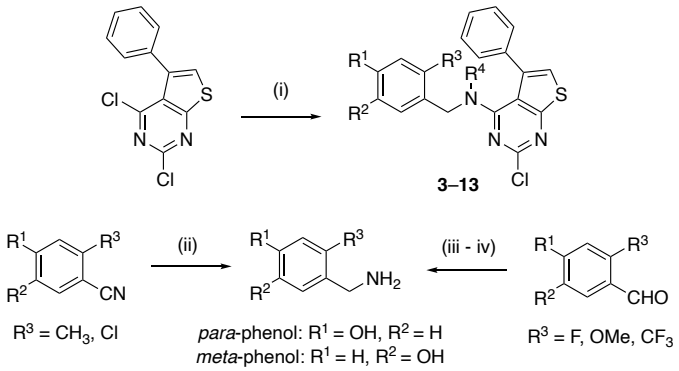
Table 1. Physiochemical properties and bioactivity of phenolic thieno[2,3-d]pyrimdines



Compound	R^1	R^2	R^3	R^4	δ_{OH} (ppm)[a]	pK_a (H_2O)[b]	σ^- (R^3)[c]	ERα (pIC_{50})[d]	ERα (pEC_{50})[e]
1	OH	H	H	H	9.33	9.73	0	7.96	7.08
2	H	OH	H	H	9.34	9.72	0	6.72	5.14
3	OH	H	–CH_2–		–	–	–	7.00	5.83
4	OH	H	CH_3	H	9.23	9.89	-0.07	5.15	5.08
5	OH	H	OCH_3	H	9.43	9.58	+0.12	5.59	<5.0
6	OH	H	F	H	9.85	8.92	+0.34	8.62	7.12
7	OH	H	Cl	H	9.86	8.90	+0.37	7.16	5.89
8	OH	H	CF_3	H	10.06	8.59	+0.43	6.24	5.13
9	H	OH	CH_3	H	9.07	10.14	-0.17	i.a.[f]	i.a.[f]
10	H	OH	OCH_3	H	8.89	10.43	-0.27	4.92	<5.0
11	H	OH	F	H	9.32	9.75	+0.06	6.82	5.56
12	H	OH	Cl	H	9.63	9.26	+0.23	5.96	<5.0
13	H	OH	CF_3	H	10.24	8.30	+0.54	4.92	<5.0

[a] ^{1}H NMR chemical shift of the phenol OH in d_6-DMSO. [b] Calculated pK_a (H_2O) = (-1.573 x δ_{OH}) + 24.41 (ref. 8). [c] Hammett substituent constant for R^3 (ref. 9). [d] ERα binding from the LiSA assay. [e] ERα agonism using an ERE reporter assay. [f] No activity at concentrations up to 30 μM. – not determined.

To shed light on the origin of the differences in ERα binding and activity of the *para*-phenol (**1**) and *meta*-phenol (**2**), several new thieno[2,3-d]pyrimidine analogs were synthesized. An S_NAr reaction of 2,4-dichloro-5-phenylthieno[2,3-d]pyrimidine[1] with various benzylamines was used to produce the phenol analogs (Scheme 1).



Scheme 1. Synthesis of thieno[2,3-d]pyrimidines (**3**–**13**). (i) Amine, $CHCl_3$, 70 ºC, Et_3N. (ii) $NaBH_4$, $CoCl_2$•$6H_2O$, THF/H_2O (2:1). (iii) $NH_2OH.HCl$, H_2O/EtOH (4:1). (iv) AcOH, Zn dust, Δ.

The bicyclic isoindolin-5-ol (**3**) (Figure 1) was the first analog to be prepared since it combined, in a single compound, the architecture of both the *para*- and *meta*-phenols using a simple atom count from the nitrogen to the oxygen (Figure S2). In the ERα assays, isoindolin-5-ol (**3**) showed activity intermediate between **1** and **2** (Table 1). In the LiSA, isoindolin-5-ol (**3**) was ~10-fold weaker than the *para*-phenol (**1**) but was 2-fold better than the *meta*-phenol (**2**). The cocrystal structure of **3** with ERα LBD (Figure 2C) revealed a binding mode in which the isoindolin-5-ol formed the full network of hydrogen bonds with E353, R394, and the water molecule that mimicked the steroidal hormone E2. An overlay of the X-ray structures of *para*-phenol (**1**), *meta*-phenol (**2**), and isoindolin-5-ol (**3**) (Figure 2D) showed that three thieno[2,3-*d*]pyrimidines had an RMSD = 0.183 Å with the only difference in the ligands being the positioning of the OH of the *para*-phenol (**1**). For isoindolin-5-ol (**3**) there appeared to be no steric penalty to adoption of a rotamer that placed its OH in the same orientation as the *para*-phenol (**1**). Thus, the preference for isoindolin-5-ol (**3**) to bind the receptor in its crystallographically observed orientation must be due to some other factor.

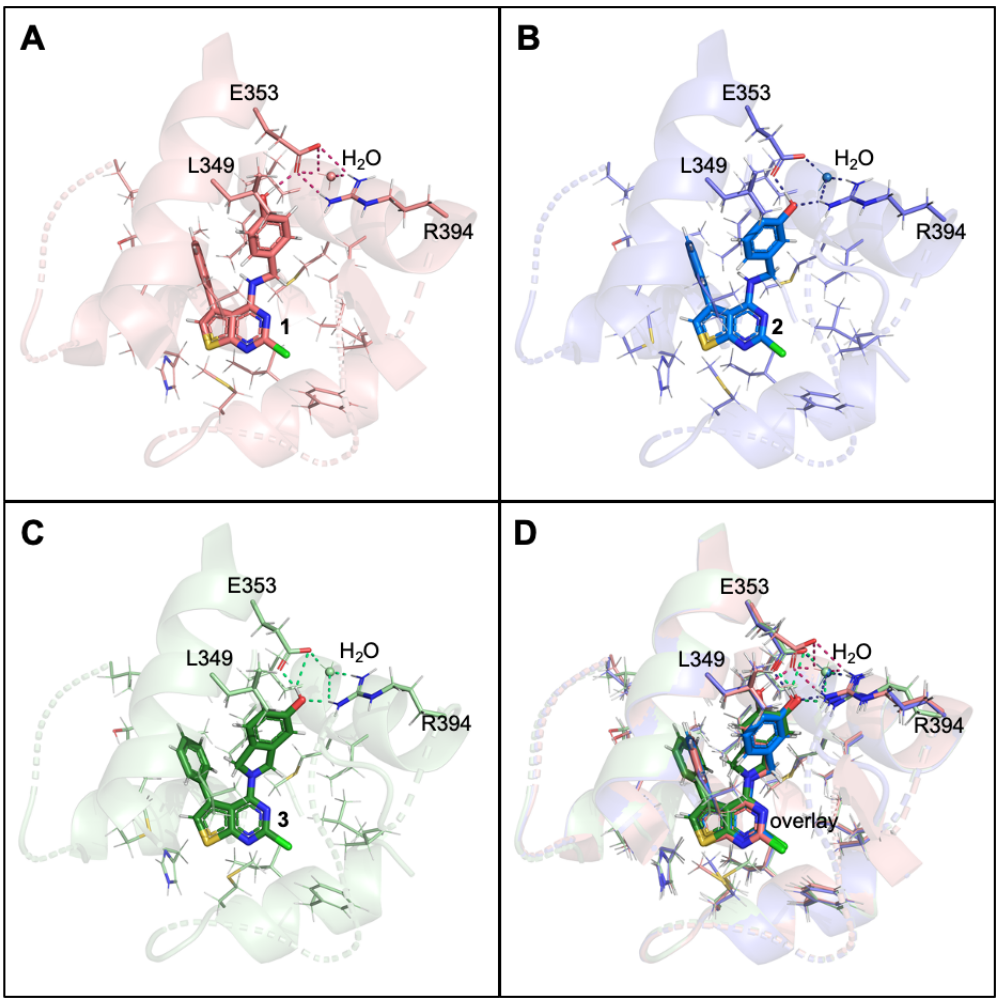


Figure 2. ERα LBD cocrystal structures. A. *para*-Phenol (**1**, PDB: 7RKE). B. *meta*-Phenol (**2**, PDB: 7SFO). C. Isoindolin-5-ol (**3**, PDB: 7RMN). D. Overlay of the three structures. E353, R394, L349, and a bound water molecule that form the phenol binding pocket are highlighted in each structure with key H-bonds shown as dashed lines.

Effect of phenol pK_a on ERα activity

Phenols are unique among aromatic substituents in their potential hydrogen bonding interactions with proteins since the oxygen atom can function as both a proton donor and acceptor. The acidity of the phenol in the aqueous environment is a key determinant of its ability to transfer or accept a proton from the receptor protein. To probe the potential role of phenol acidity in influencing the binding mode of the thieno[2,3-d]pyrimidines two parallel series of analogs were synthesized in the *para*-phenol (**4**–**8**) and *meta*-phenol (**9**–**13**) series (Scheme 1). The analogs were designed to place the R^3 substituent at a position on the phenol ring that would not be affected by potential steric clashes based on analysis of the cocrystal structures (Figure 2D). The benzyl amines required for the S_NAr reaction were synthesized by two routes. The methyl- and chloro-substituted phenols were accessed from commercially available nitriles

by cobalt (II) facilitated borahydride reduction,[6] while the fluoro-, methoxy-, and trifluromethyl-substituted phenols required a two-step synthesis from the aldehyde by hydroxamate formation and reduction using zinc dust[7] (Scheme 1).

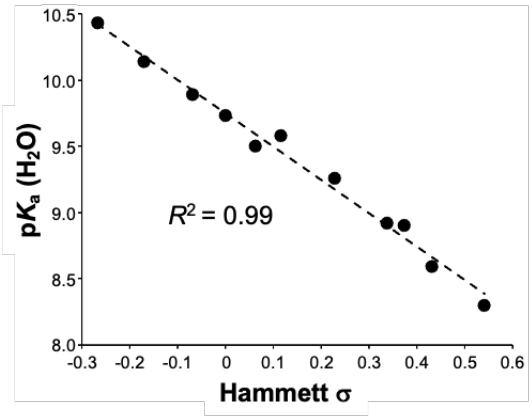


Figure 3. Correlation of phenol pK_a and Hammett σ^- parameter

The pK_as of the *para*-phenols (**1**, **4**–**8**) and *meta*-phenols (**2**, **9**–**13**) were determined from their ^{1}H-NMR spectra (Figure S3). It has been previously determined that the pK_a of a wide range of substituted phenols in water can be accurately calculated from the chemical shift of the phenolic proton in d_6-DMSO.[8] Using this relationship, the pK_a (H_2O) of the phenolic thieno[2,3-*d*]pyrimidines (**1**–**2**, **4**–**13**) varied over 2 log units from 8.30 to 10.43 (Table 1). Notably, the *para*-phenol (**1**) and *meta*-phenol (**2**) had nearly identical pK_a indicating that effect of the benzylamine substituent was essentially neutral. A plot of the Hammett substituent constant against pK_a (H_2O) showed a correlation with R^2 = 0.99 (Figure 3), indicating that the inductive and mesomeric effect of the R^3 substituent was the primary regulator of phenol acidity in both series.

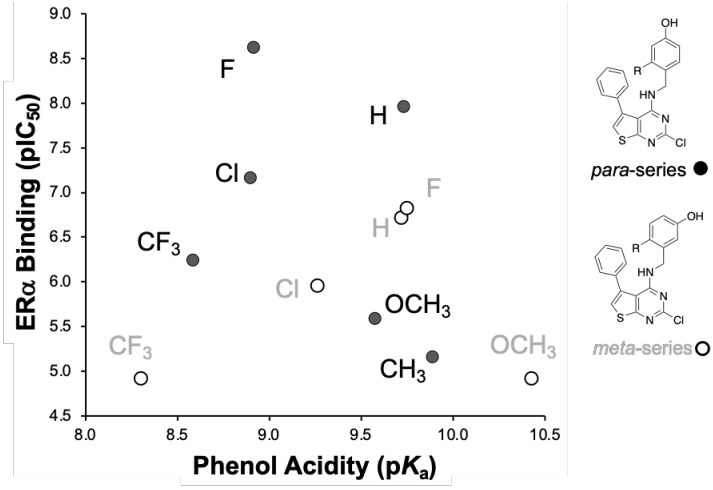


Figure 4. Relationship between phenol acidity and activity in ERα LiSA

In the *para*-phenol series (**4**–**8**) the fluoro-substituted analog (**6**) showed ~5-fold improved activity in the ERα LiSA compared to the unsubstituted phenol (**1**) (Table 1). However, the other analogs had weaker activity, with the overall rank order being F > H > Cl > CF_3 > OCH_3 > CH_3. In the *meta*-phenol series (**9**–**13**) the fluoro-substituted analog (**11**) had only slightly higher activity than parent phenol (**2**), with the other analogs showing weaker apparent binding affinity. In this series the rank order of ERα LiSA activity was F ~ H > Cl > CF_3 ~ OCH_3 > CH_3. For both series, activity in the ERα reporter gene assay matched the apparent binding affinity in the LiSA, although the weaker ligands produced incomplete dose-response curves due to the limitation in maximum concentration tested (10 μM). Examination of the substituent effects on ERα activity showed that, although there was a similar rank order in the *para*-phenol or *meta*-phenol series, there was no apparent relationship with phenol pK_a (H_2O) when considering them independently or in combination (Figure 4). Thus, phenol pK_a alone was not the driver of ERα activity in the LiSA or reporter gene assays for these thieno[2,3-*d*]pyrimidine ligands. This result contrasts with the structure activity of simple nonsteroidal estrogens *para*-hexestrol and *meta*-hexestrol, where relative binding affinity and agonist activity matched the electron-donating properties of the substituents.[9, 10]

First-principles calculation of ERα binding

To gain further insight into how the differences in H-bonding of the phenols in **1**, **2**, and **3** with the ERα LBD could affect their activity, we performed a series of density functional theory (DFT) calculations.[11] From the X-ray co-crystal structures, three models were constructed using the ligands **1**–**3** and the atoms from the amino acids in the first shell of the binding pocket, comprising >500 atoms in each case. All hydrogen atoms of the models were optimized at the DFT M062X/6-31G(d) level of theory[12] and basis set superposition error (BSSE) corrected[13] calculations performed at two theoretical levels using 6-311+G(d) basis set[14] for O, N, Cl, and S elements and 6-311G(d) basis set for C and H elements using Gaussian.[15] As shown in Table 2, the *meta*-phenol (**2**) showed the highest binding energy in both DFT models, with the *para*-phenol (**1**) producing values ~10 kcal/mol lower energy and the isoindolin-5-ol (**3**) returning intermediate values. The calculations appeared to reflect the additional H-bonding interactions that were observed in the *meta*-phenol (**2**) and isoindolin-5-ol (**3**) X-ray crystal structures and did not explain the improved activity of *para*-phenol (**1**) as an ERα ligand.

PDB Code (compound)	DFT Model 1[a] binding energy (kcal/mol)	DFT Model 2[b] binding energy (kcal/mol)
7RKE (**1**)	-42.1	-50.7
7SFO (**2**)	-52.0	-59.6
7RMN (**3**)	-47.7	-54.3

[a] DFT M06-2X/6-31G(d). [b] DFT M06-2X/6-311+G(d) for O/N/Cl/S elements and 6-311G(d) for C/H elements.

Table 2. DFT calculations of ERα binding affinity

MD simulations of ERα binding

A series of molecular dynamics (MD) simulations was performed to explore the idea that activity of the phenolic thieno[2,3-d]pyrimidines in the ERα LiSA and reporter gene assays was the result of a complex dynamic of receptor binding and stabilization of the hormone receptor in an active conformation. MD simulations were performed over 100 ps using structures prepared from the X-ray coordinates of **1**, **2**, and **3** with ERα LBD (PDB: 7RKE, 7SFO, and 7RNM, respectively) that included the structural water molecules within 8 Å of E353, R394, and the ligand. The MD simulations revealed some significant differences in the ligand-residue distances and ligand-induced protein conformation changes (Figure 5). The distribution of distances of the phenol to each of three key residues (E353, L349 and R394) in the ligand binding pocket was determined. The results indicated that *para*-phenol (**1**) (Figure 5A, pink) was located closer to L349 during the time of the simulations compared to *meta*-phenol (**2**) (Figure 5A, blue) and isoindolin-5-ol (**3**) (Figure 5A, green). The *para*-phenol (**1**) trajectory also showed a tighter unimodal distribution of the ligand in its interaction distances with E353 and R394, consistent with its higher potency. Analysis of binding site residues root-mean-square fluctuations (RMSF) (Figure 5B) revealed greater stabilization of R394 by *para*-phenol (**1**) compared to *meta*-phenol (**2**) with isoindolin-5-ol (**3**) producing an intermediate stabilization. This rank order matched the relative potency of **1**, **2**, and **3** in the LiSA and reporter gene assays (Table 1). No significant differences in the stabilization of E353, L349 were observed. Among the other residues of the ligand binding pocket there was a specific stabilization of F404 by *para*-phenol (**1**). In sum, the MD simulations showed that the binding of *para*-phenol (**1**) to the ERα LBD resulted in a relatively more stable ligand-receptor complex compared to *meta*-phenol (**2**) or isoindolin-5-ol (**3**) despite the absence of a full complement of H-bonding interactions between the phenol and the ligand binding pocket.

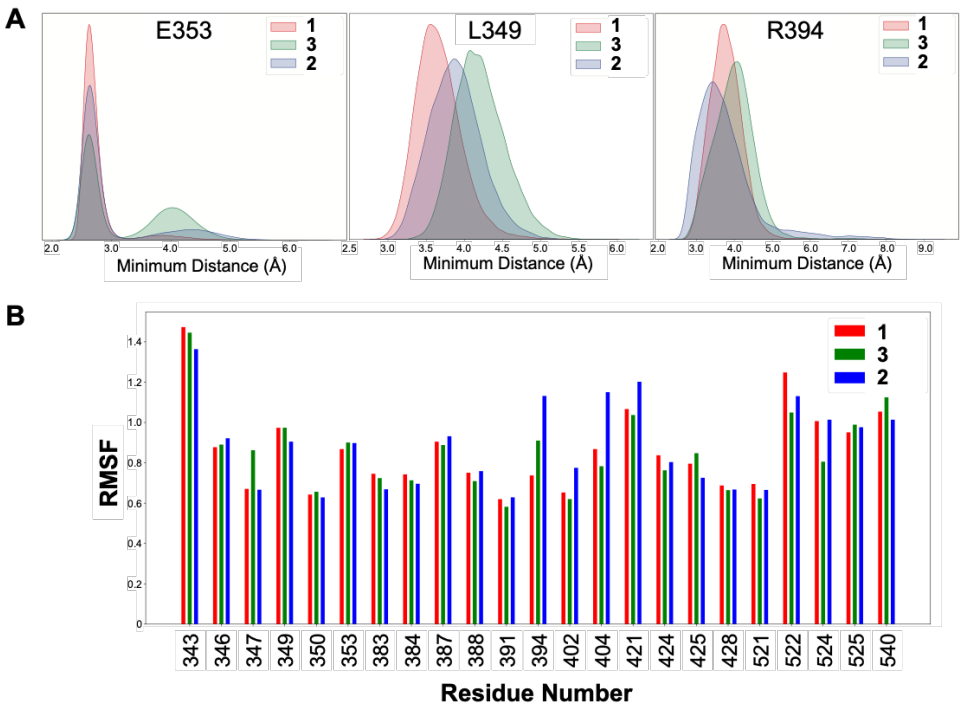


Figure 5. Results of MD Simulations. A. Ligand-residue distance distributions for *para*-phenol (**1:** 7RKE), *meta*-phenol (**2**: 7SFO) and isoindolin-5-ol (**3**:7RNM) with E353, L349, and R394. Based on the shape of the distribution and the peak value, the *para*-phenol (**1**, pink) spent more time in close proximity to the residues. B. RMSF values for protein residues in the ERα binding site calculated from MD trajectories. All values in the figure were calculated from combining three 1 μs MD trajectories within each complex.

Conclusions

The phenolic thieno[2,3-d]pyrimidines are a new chemotype of non-steroidal estrogens that bind and activate ERα. Surprisingly the *para*-phenol (**1**) was more a potent estrogen than the *meta*-phenol (**2**), despite the latter appearing to be a better mimic of A-ring of the natural hormone E2. The difference in activity could not be explained by the number and strength of H-bonds between the receptor and the phenolic ligands, as demonstrated by a series of analogs in which the pK_a of the phenols was systematically varied. First-principles calculation of the binding energy also failed to explain the greater activity of the *para*-phenol (**1**). However, MD simulations of the ligand-receptor complex indicated that the *para*-phenol (**1**) produced a more stable ensemble of conformations. Our results support a hypothesis that the ligand binding site of ERα did not evolve to simply a maximize the H-bond interactions with its phenolic hormone. Instead, the molecular recognition of estrogens (C3 phenol) over progestins (C3 ketone) uses a network of H-bonds with a phenol that cannot be achieved by a ketone. Nonsteroidal estrogens such as the phenolic thieno[2,3-d]pyrimidines have more options in how they bind and stabilize

the receptor and a single H-bond is apparently sufficient to produce a stable ligand-receptor complex.

Experimental Section

General Procedures.

All the reactions were performed in oven-dried glassware. Thin-layer chromatography (TLC) was performed using aluminum backed silica coated plates. TLC plates were visualized under Ultraviolet light. The ^{1}H and ^{13}C NMRs were recorded on Agilent 400MR INOVA NMR, [^{1}H (400 MHz), ^{13}C (101 MHz)] and Bruker Avance III HD 850 MHz four channel spectrometer equipped with a TCI H-C/N-D 5 mm CryoProbe [^{1}H (850 MHz), ^{13}C (214 MHz)] with complete proton decoupling for ^{13}C. Chemical shifts were analyzed on MestReNova software. Chemical shifts are reported in parts per million with the solvent resonance as the internal standard (DMSO-d_6, ^{1}H: δ2.50 ppm, ^{13}C: δ39.52 ppm). Coupling constants are reported in Hertz (Hz). Abbreviations are used as follows: s = singlet, d = doublet, t = triplet, q = quartet, hept = heptet, m = multiplet, dd=doublet of doublet. Samples were analyzed with a ThermoFisher Q Exactive HF-X (ThermoFisher, Bremen, Germany) mass spectrometer (electrospray ionization) coupled with a Waters Acquity H-class liquid chromatograph system.

Substituted amines: Commercially available phenolic benzylamines were purchased from Combi-Blocks Inc. Non-commercially available phenolic benzylamines (Scheme 1, R_3 = Me, Cl) were synthesized by reduction of the nitrile with $NaBH_4$ and $CoCl_2$•$6H_2O$ in aqueous THF.[6] Non-commercially available phenolic benzylamines (Scheme 1, R_3 = F, OMe, CF_3) were synthesized by amination of the aldehyde with $NH_2OH.HCl$ in aqueous EtOH followed by reduction of the imine with Zn dust in AcOH.[7]

5-Phenylthieno[2,3-d]pyrimidines 1–13: To a reaction vial containing a substituted amine (0.25 mmol, 1.25 eq.) was added triethylamine (5 eq.) dropwise in one portion and the contents were stirred at 75°C until a clear solution was obtained. Methanol (0.5 mL) was added to the mixture in those cases where a clear solution was not observed after 5 min. 2,4-Dichloro-5-phenylthieno[2,3-d]pyrimidine[1] (0.2 mmol, 1 eq.) was added in one portion and the solution stirred at 75°C and the reaction was monitored by TLC over 12–24 h for consumption of thienopyrimidine. The reaction contents were cooled to room temperature, transferred into a separatory funnel containing water (30 mL) and CH_2Cl_2 (15 mL). The organic layer was extracted, washed with brine, dried over anhydrous sodium sulfate, and evaporated to yield a

crude product that was purified by flash chromatography using a gradient elution of ethyl acetate/hexane.

4-(((2-Chloro-5-phenylthieno[2,3-d]pyrimidin-4-yl)amino)methyl)phenol (1).[1] White solid; yield 66%; ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.33 (s, 1H), 7.51 (s, 1H), 7.48 – 7.37 (m, 5H), 6.98 (d, *J* = 8.5 Hz, 2H), 6.65 (d, *J* = 8.6 Hz, 2H), 5.81 (t, *J* = 5.2 Hz, 1H), 4.39 (d, *J* = 5.2 Hz, 2H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 167.7, 157.9, 157.0, 155.4, 135.2, 134.7, 129.5, 129.4, 129.2, 129.0, 127.9, 122.1, 115.6, 112.7, 44.4; HRMS (ES+) m/z calc. for [$C_{19}H_{14}ClN_3OS$ + H^+]: 368.0619; found: 368.0614.

3-(((2-Chloro-5-phenylthieno[2,3-d]pyrimidin-4-yl)amino)methyl)phenol (2).[1] White solid; yield 58%; ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.34 (s, 1H), 7.53 (s, 1H), 7.51 – 7.38 (m, 5H), 7.07 (t, *J* = 7.67 Hz, 1H), 6.59 (t, *J* = 11.81 Hz, 3H), 5.98 (t, *J* = 5.39 Hz, 1H), 4.46 (d, *J* = 5.53 Hz, 2H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 167.9, 158.1, 157.9, 155.4, 139.5, 135.2, 134.8, 129.9, 129.5, 129.3, 129.1, 122.2, 118.1, 114.6, 114.5, 112.7, 44.6; HRMS (ES+) m/z calc. for [$C_{19}H_{14}ClN_3OS$ + H^+]: 368.0619; found: 368.0615.

2-(2-Chloro-5-phenylthieno[2,3-d]pyrimidin-4-yl)isoindolin-5-ol (3). Crème color solid; yield 47%; ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.32 (s, 1H), 7.60 (s, 1H), 7.51 – 7.37 (m, 5H), 6.84 (d, *J* = 8.85 Hz, 1H), 6.56 (dd, *J* = 8.28, 1.85 Hz, 1H), 6.43 (s, 1H), 4.33 (d, *J* = 17.62 Hz, 4H); ^{13}C NMR (214 MHz, DMSO-d_6) δ 173.2, 162.3, 160.0, 156.2, 140.8, 140.1, 138.9, 132.1, 131.1, 131.0, 128.8, 125.9, 124.2, 117.8, 116.2, 111.7, 58.3, 57.8; HRMS (ES+) m/z calc. for [$C_{20}H_{14}ClN_3OS$ + H^+]: 379.0546; found: 380.06145.

4-(((2-Chloro-5-phenylthieno[2,3-d]pyrimidin-4-yl)amino)methyl)-3-methylphenol (4). White solid; yield 82%; ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.23 (s, 1H), 7.50 (s, 1H), 7.46 – 7.32 (m, 5H), 6.87 (d, *J* = 8.30 Hz, 1H), 6.52 (d, *J* = 2.02 Hz, 1H), 6.46 (dd, *J* = 8.21, 2.43 Hz, 1H), 5.59 (t, *J* = 5.20 Hz, 1H), 4.35 (d, *J* = 5.01 Hz, 2H), 2.04 (s, 3H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 167.6, 157.8, 157.1, 155.4, 137.6, 135.2, 134.6, 130.1, 129.4, 129.3, 129.1, 125.9, 122.1, 117.4, 113.0, 112.8, 42.5, 19.1; HRMS (ES+) m/z calc. for [$C_{20}H_{16}ClN_3OS$ + H^+]: 381.0703; found: 382.0767.

4-(((2-Chloro-5-phenylthieno[2,3-d]pyrimidin-4-yl)amino)methyl)-3-methoxyphenol (5). Crème solid; yield 51%; ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.43 (s, 1H), 7.48 (s, 1H), 7.47 – 7.38 (m, 5H), 6.90 (d, *J* = 8.13 Hz, 1H), 6.29 (d, *J* = 2.09 Hz, 1H), 6.22 (dd, *J* = 8.11, 2.16 Hz, 1H), 5.76 (t, *J* = 5.36 Hz, 1H), 4.33 (d, *J* = 5.27 Hz, 2H), 3.50 (s, 3H); ^{13}C NMR (101 MHz,

DMSO-d_6) δ 167.6, 158.9, 158.6, 157.8, 155.4, 135.3, 134.6, 130.5, 129.4, 129.3, 129.1, 122.1, 115.5, 112.8, 106.9, 99.2, 55.6, 40.8; HRMS (ES+) m/z calc. for [$C_{20}H_{16}ClN_3O_2S + H^+$]: 397.0652; found: 398.0716.

4-(((2-Chloro-5-phenylthieno[2,3-d]pyrimidin-4-yl)amino)methyl)-3-fluorophenol (6): White solid; yield 53%; ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.85 (s, 1H), 7.52 (s, 1H), 7.42 (s, 5H), 7.05 (t, *J* = 8.89 Hz, 1H), 6.55 – 6.46 (m, 2H), 5.83 (t, *J* = 5.40 Hz, 1H), 4.42 (d, *J* = 5.26 Hz, 2H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 167.8, 162.5, 160.1, 157.8, 155.3, 135.1, 134.7, 131.1, 131.1, 129.4 (d, *J* = 14.50 Hz), 129.1, 122.2, 114.6 (d, *J* = 15.58 Hz), 112.8, 111.7, 102.9 (d, *J* = 24.01 Hz), 38.6; HRMS (ES+) m/z calc. for [$C_{19}H_{13}ClFN_3OS + H^+$]: 385.0452; found: 386.0514.

3-Chloro-4-(((2-chloro-5-phenylthieno[2,3-d]pyrimidin-4-yl)amino)methyl)phenol (7). White solid; yield 56%; ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.86 (s, 1H), 7.52 (s, 1H), 7.48 – 7.37 (m, 5H), 7.13 (d, *J* = 8.38 Hz, 1H), 6.77 (d, *J* = 2.44 Hz, 1H), 6.65 (dd, *J* = 8.38, 2.40 Hz, 1H), 5.89 (t, *J* = 5.50 Hz, 1H), 4.45 (d, *J* = 5.52 Hz, 2H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 167.8, 158.2, 157.9, 155.3, 135.2, 134.7, 133.4, 131.6, 129.5, 129.4, 129.1, 125.2, 122.3, 116.3, 114.7, 112.9, 42.4; HRMS (ES+) m/z calc. for [$C_{19}H_{13}Cl_2N_3OS + H^+$]: 401.0156; found: 402.0224.

4-(((2-Chloro-5-phenylthieno[2,3-d]pyrimidin-4-yl)amino)methyl)-3-(trifluoromethyl)phenol (8). White solid; yield 42%; ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.06 (s, 1H), 7.52 (s, 1H), 7.49 – 7.36 (m, 5H), 7.29 (d, *J* = 8.43 Hz, 1H), 7.00 (d, *J* = 2.39 Hz, 1H), 6.96 (dd, *J* = 8.41, 2.19 Hz, 1H), 5.83 (t, *J* = 5.32 Hz, 1H), 4.57 (d, *J* = 5.06 Hz, 2H); ^{13}C NMR (214 MHz, DMSO-d_6) δ 170.6, 160.5, 159.9, 157.9, 137.8, 137.4, 135.1, 132.1, 132.0, 131.7, 130.8 (q, *J* = 29.58 Hz), 128.5, 127.8, 125.0, 122.3, 115.7 (q, *J* = 5.45 Hz), 115.6, 43.8; HRMS (ES+) m/z calc. for [$C_{20}H_{13}ClF_3N_3OS + H^+$]: 435.0420; found: 436.0482.

3-(((2-Chloro-5-phenylthieno[2,3-d]pyrimidin-4-yl)amino)methyl)-4-methylphenol (9) White solid; yield 55%; ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.07 (s, 1H), 7.53 (s, 1H), 7.49 – 7.36 (m, 5H), 6.89 (d, *J* = 9.39 Hz, 1H), 6.52 (dd, *J* = 4.30, 1.96 Hz, 2H), 5.84 (t, *J* = 5.43 Hz, 1H), 4.40 (d, *J* = 5.67 Hz, 2H), 2.02 (s, 3H); ^{13}C NMR (214 MHz, DMSO-d_6) δ 170.5, 160.8, 158.5, 158.1, 139.5, 137.9, 137.4, 134.0, 132.2, 132.0, 131.8, 128.4, 124.9, 117.6, 117.1, 115.5, 45.5, 20.7; HRMS (ES+) m/z calc. for [$C_{20}H_{16}ClN_3OS + H^+$]: 381.0703; found: 382.0769.

3-(((2-Chloro-5-phenylthieno[2,3-d]pyrimidin-4-yl)amino)methyl)-4-methoxyphenol (10). White solid; yield 49%; ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.89 (s, 1H), 7.50 (s, 1H), 7.49 – 7.41 (m, 5H), 6.70 (d, *J* = 9.42 Hz, 1H), 6.58 (dd, *J* = 6.57, 2.90 Hz, 2H), 5.94 (t, *J* = 5.79 Hz,

1H), 4.40 (d, *J* = 5.71 Hz, 2H), 3.48 (s, 3H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 167.8, 158.0, 155.4, 151.1, 150.3, 135.2, 134.7, 129.5, 129.3, 129.1, 126.1, 122.2, 116.4, 114.7, 112.8, 112.0, 56.0, 40.8; HRMS (ES+) m/z calc. for [$C_{20}H_{16}ClN_3O_2S$ + H^+]: 397.0652; found: 398.0715.

3-(((2-Chloro-5-phenylthieno[2,3-d]pyrimidin-4-yl)amino)methyl)-4-fluorophenol (11). White solid; yield 51%; ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.32 (s, 1H), 7.54 (s, 1H), 7.50 – 7.37 (m, 5H), 6.93 (t, *J* = 9.93 Hz, 1H), 6.60 (dd, *J* = 5.89, 3.27 Hz, 2H), 6.05 (t, *J* = 5.67 Hz, 1H), 4.49 (d, *J* = 5.97 Hz, 2H); ^{13}C NMR (214 MHz, DMSO-d_6) δ 170.7, 160.7, 158.0, 157.1, 156.6, 156.0, 137.6 (d, *J* = 77.85 Hz), 132.2, 132.0, 131.8, 128.2 (d, *J* = 16.35 Hz), 125.1, 118.8 (d, *J* = 22.58 Hz), 118.5 (d, *J* = 3.89 Hz), 118.1 (d, *J* = 7.79 Hz), 115.4, 41.6 (d, *J* = 3.89 Hz); HRMS (ES+) m/z calc. for [$C_{19}H_{13}ClFN_3OS$ + H^+]: 385.0452; found: 386.0514.

4-Chloro-3-(((2-chloro-5-phenylthieno[2,3-d]pyrimidin-4-yl)amino)methyl)phenol (12) White solid; yield 55%; ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.63 (s, 1H), 7.54 (s, 1H), 7.52 – 7.38 (m, 5H), 7.16 (d, *J* = 8.58 Hz, 1H), 6.70 (d, *J* = 5.96 Hz, 1H), 6.64 (dd, *J* = 8.55, 2.95 Hz, 1H), 6.12 (t, *J* = 6.01 Hz, 1H), 4.50 (d, *J* = 5.68 Hz, 2H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 168.0, 158.0, 156.8, 155.3, 136.2, 135.1, 134.8, 130.4, 129.5, 129.3, 129.1, 122.4, 121.9, 116.6, 116.2, 112.8, 42.8; HRMS (ES+) m/z calc. for [$C_{19}H_{13}Cl_2N_3OS$ + H^+]: 401.0156; found: 402.0222.

3-(((2-Chloro-5-phenylthieno[2,3-d]pyrimidin-4-yl)amino)methyl)-4-(trifluoromethyl)phenol (13). White solid; yield 57%; ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.24 (s, 1H), 7.55 (s, 1H), 7.52 – 7.36 (m, 6H), 6.81 (s, 0H), 6.75 (dd, *J* = 8.59, 1.78 Hz, 1H), 6.15 (t, *J* = 5.98 Hz, 1H), 4.64 (d, *J* = 5.33 Hz, 2H); ^{13}C NMR (214 MHz, DMSO-d_6) δ 170.8, 163.9, 160.8, 157.9, 141.4, 137.8, 137.5, 132.2, 132.0, 131.8, 131.1 (q, *J* = 5.30 Hz), 128.0 (q, *J* = 272.09 Hz), 125.2, 119.9 (q, *J* = 30.36 Hz), 118.8, 116.9, 115.5, 44.1; HRMS (ES+) m/z calc. for [$C_{20}H_{13}ClF_3N_3OS$ + H^+]: 435.0420; found: 436.0485.

ERα LiSA. HepG2 cells were maintained in Basal Medium Eagles supplemented with 8% fetal bovine serum (FBS) and plated (2.5 x 10^6 cells/10 mL) in T75 flasks for 24 h prior to transfection. Cells were transfected with VP16-ERα, 5XGalLuc3, pM-alpha II, and Renilla luciferase at a 1:30:1:5 ratio using Lipofectin (Thermo Fisher) according to manufacturer's protocol. Following 24 h transfection, cells were plated in 96-well plates and treated for 24 h with with test compounds (1.0 x 10^{-11}–1.0 x 10^{-5} M). Dual luciferase assays were performed from cell extracts. Renilla luciferase served as control for cellular toxicity. Data was fit in Prism Software to determine IC_{50} values.

ERE reporter gene assay. HepG2 cells were plated in T75 flasks (1.5×10^{-6} cells/10 mL) and sub-cultured for 24 h. Cells were transfected with pRST7-ERα, Renilla luciferase, and a 7X-ERE-TK-luc reporter gene at a 1:4:25 ratio using Lipofectin transfection reagent. Following 24 h transfection, cells were treated with test compounds (1.0×10^{-11}–1.0×10^{-5} M) for 24 h. Cells were lysed, and extracts were assayed for Firefly and Renilla luciferase activity using dual luciferase reagent. Data was fit in Prism Software to determine EC_{50} values.

ER whole cell competition binding assay. 1.0×10^5 MCF7 cells were plated in DMEM/F12 with 8% charcoal-stripped fetal bovine serum in 24 well plates for 24 h. Following incubation, cells were treated with test compounds (1.0×10^{-10}–1.0×10^{-6} M) in the presence of 0.1 nM ^{3}H-E2. To determine background levels of radioactivity, control wells were treated with excess cold E2 (50 nM). After 2 h incubation, cells were treated with lysis buffer (200 μL; 2% SDS, 10% Glycerol, 10 mM Tris-HCl pH 6.8), and then volumes were increased to 500 μL using 10 mM Tris-HCl (pH 8.0). 300 μL of the lysates were added to 3 mL of Cytoscint (MP Biomedicals) and analyzed by scintillation counting (Beckman LS 6000SC).

Protein Expression and Purification. A His_6-TEV-tagged ERα LBD (305–555) with a Y537S mutation in pET21a was expressed in *E.coli* BL21(DE3) using the previously reported methods.[16] Cells were grown at 37 ºC with shaking in LB broth, supplemented with ampicillin, until they reached mid-log phase growth (OD_{600} = 0.6). The temperature was reduced to 16 ºC and protein expression was induced with 0.4 M IPTG for 16 h. Cells were harvested by centrifugation at 4,000 *xg* for 15 min then resuspended with 20 mL/g cell paste with 50 mM HEPES pH 8.0, 500 mM NaCl, 0.5 mM TCEP, 20 mM imidazole pH 8.0, 5% glycerol, and EDTA-free PIC tabs (Pierce). Cells were lysed by sonication and cellular debris was cleared by centrifugation at 18,000 *xg* for 45 min. A BioRad NGC Quest FPLC was used for protein purification. The supernatant containing the His_6-TEV-ERα was captured using a BioRad Nuvia IMAC column, washed with lysis buffer until baseline was reached, and eluted with 50 mM HEPES pH 8.0, 500 mM NaCl, 0.5 mM TCEP, 500 mM imidazole pH 8.0, and 5% glycerol. 1: 5,000 mol:mol His-TEV protease was added to the protein and the solution was dialyzed against 4 L of 50 mM HEPES pH 8.0, 500 mM NaCl, 0.5 mM TCEP, pH 8.0, 5% glycerol overnight at 4ºC with stirring for 16 hours. The mixture was placed back over the IMAC column to remove the His-tag and His-TEV protease. ERα LBD Y537S was purified a final time using a GE Superdex 16/600 size exclusion column. A single peak corresponding to approximately 50 kDa MW corresponding to the ERα LBD Y537S dimer was concentrated to 10-15 mg/mL, flash frozen, and stored at -80ºC.

Crystallization, X-Ray Data Collection, and Structure Solution. Each small molecule ligand was added to 10 mg/mL ERα LBD Y537S at 1 mM alongside GRIP peptide at 3 mM overnight at 4°C. Hanging drop vapor diffusion was used to crystallize the protein using Hampton VDX plates. Clear rectangular crystals emerged after 16 h at room temperature in 15–25% PEG 3,350, 100 mM $MgCl_2$, 100 mM Tris pH 8.0 at 5 mg/mL. Paratone-N was used as a the cryoprotectant. X-ray diffraction data were collected at the Structural Biology Center 19-BM beamline at the Advanced Photon Source, Argonne National Laboratories at 0.97 Å. Data were indexed and scaled using HKL 3,000.[17] Phenix[18] was used for molecular replacement using PDB 6CBZ as the input model.[19] The structures were refined using iterative rounds of Phenix Refine and manual inspection/structure editing using Coot. Significant difference density was observed in the ligand binding pockets with one round of structure refinement that resolved after placing the compounds in the ligand binding pocket and further refinement. X-ray crystal structure collection and refinement statistics are recorded in Table S3. The coordinates and map were deposited in the protein databank (PDB) with accession codes 7RKE, 7SFO, and 7RMN.

Computational Methods

DFT Calculations. Three structural models were constructed for density functional theory (DFT)[11] calculations using the ligands **1**–**3** and the atoms from the amino acids in the first shell of the binding pocket; defined as all atoms of the receptor located within 8 Å of any atom of the respective ligand. Each model had a total number of atoms in excess of 500. All hydrogen atoms of the models were optimized at the DFT M062X/6-31G(d) level of theory[12] with all positions of non-hydrogen elements fixed. After geometrical optimizations, BSSE (basis set superposition error) corrected[13] calculations were carried out to obtain the interaction energy between the ligand and the receptor. The BSSE-corrected calculations were performed at two theoretical levels, one with at the M06-2X/6-31G(d) level of theory and the other with M06-2X functional and 6-311+G(d) basis set[14] for O, N, Cl, and S elements and 6-311G(d) basis set for C and H elements. All calculations were performed with the Gaussian package, version C01.[15]

MD Simulations. The initial complex structures were prepared from the X-ray coordinates (PDB: 7RKE, 7RNM and 7SFO Chain B) using Protein Preparation Wizard within Maestro.[20] The structural water molecules located within 8 Å from the binding site residues E353, R394, and the ligand in each structure were included. The initial systems for MD simulations were set

up using antechamber[21] and tleap programs in the AmberTools.[22] The prepared structures were solvated in a rectangular box with a minimum 12 Å distance between the complex and box edges, and sodium ions were added to neutralize the system charges. MD simulations were performed with pmemd.cuda in Amber20[23-25] using Amber forcefields (*ff14SB*, *tip3p* and *gaff2*).[26-29] All systems were equilibrated in NVT ensembles where they were heated to 310 K over 100 ps with 10.0 kcal/mol. The $Å^{-2}$ harmonic restraints on the heavy atoms of the protein, ligand, and structural water molecules were retained from the initial PDB structure. Then, the restraints were gradually relieved over 1 ns in NPT ensembles. Three replicate simulations of 1 μs each were carried out for each protein-ligand complex. These simulations were performed using 2.0 fs timestep under constant pressure (1 bar) regulated by Monte Carlo barostat and constant temperature (310 K) regulated by Langevin thermostat with 2.0 ps^{-1} collision frequency. The bonds involving hydrogen atoms were constrained by the SHAKE algorithm.[30] Particle mesh Ewald (PME) was used to calculate electrostatic interactions with 10 Å cutoff for non-bonded interactions[31] The simulation trajectories were analyzed with cpptraj package[32] in the AmberTools and visualized using PyMol 2.1 (Schrödinger, LLC) software.

Supplementary Material

Supporting information is available. Figure S1 – Competition binding data for **1** and **2**. Figure S2 – Atom count from the amine N to phenol O for **1**, **2**, **3**. Figure S3 – ^{1}H NMR chemical shift of the phenol in **1**, **2**, and **4**–**13**. Table S1 – X-ray crystallographic refinement data. NMR spectra of final compounds **1**–**13**.

Acknowledgments

The authors thank Alex Tropsha (UNC) for discussion of computational methodologies and critical reading of the manuscript. The Structural Genomics Consortium is a registered charity (no: 1097737) that receives funds from Bayer AG, Boehringer Ingelheim, Bristol Myers Squibb, Genentech, Genome Canada through Ontario Genomics Institute [OGI-196], EU/EFPIA/OICR/McGill/KTH/Diamond Innovative Medicines Initiative 2 Joint Undertaking [EUbOPEN grant 875510], Janssen, Merck KGaA (aka EMD in Canada and US), Pfizer and Takeda. This work was supported by the Department of Defense through the FY 17 BRCP Innovator Award under Award No. BC170954. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the Department of Defense. Research reported in this publication was supported in part by NIH 1R01GM140154 and NC Biotech Center Institutional Support Grant 2018-IDG-1030, with additional funding from

Loyola University Chicago Startup Funds and Susan G. Komen CCR 19608597. Results shown in this report are derived from work performed at Argonne National Laboratory (ANL), Structural Biology Center (SBC) at the Advanced Photon Source (APS), under U.S. Department of Energy, Office of Biological and Environmental Research contract DE-AC02-06CH11257.

Author Contribution Statement

T.M.W., *S.W.F.*, and *D.P.M* conceived and designed the experiments. *V.K.S.*, *B.A.*, *J.D.N.*, *C.D.T.*, and *C.J.* performed the experiments and analyzed the data. *S.L.* conducted DFT calculations. *H.L.* and *K.I.P.* conducted MD simulations. *T.M.W.* wrote the article.

Accession Codes

Atomic coordinates for the X-ray structure of ERα LBD Y537S in complex with **1**, **2**, and **3** (PDB: 7RKE, 7SFO, and 7RNM) are available from the RCSB Protein Data Bank (www.rscb.org).

References

[1] V. R. Sammeta, J. D. Norris, S. Artham, C. D. Torrice, J. Byemerwa, C. Joiner, S. W. Fanning, D. P. McDonnell, T. M. Willson, 'A New Chemotype of Chemically Tractable Nonsteroidal Estrogens Based on a Thieno[2,3-d]pyrimidine Core', *ACS Med. Chem. Lett.* **2022**, *13*, 1151-1158.

[2] D. M. Tanenbaum, Y. Wang, S. P. Williams, P. B. Sigler, 'Crystallographic Comparison of the Estrogen and Progesterone Receptor's Ligand Binding Domains', *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5998-6003.

[3] E. Palomino, M. J. Heeg, J. P. Horwitz, S. C. Brooks, 'Binding, X-ray and NMR Studies of the Three A-ring Isomers of Natural Estradiol', *J. Steroid. Biochem.* **1990**, *35*, 219-229.

[4] J. D. Norris, L. A. Paige, D. J. Christensen, C. Y. Chang, M. R. Huacani, D. Fan, P. T. Hamilton, D. M. Fowlkes, D. P. McDonnell, 'Peptide Antagonists of the Human Estrogen Receptor', *Science* **1999**, *285*, 744-746.

[5] L. A. Paige, D. J. Christensen, H. Gron, J. D. Norris, E. B. Gottlin, K. M. Padilla, C. Y. Chang, L. M. Ballas, P. T. Hamilton, D. P. McDonnell, D. M. Fowlkes, 'Estrogen Receptor (ER) Modulators Each Induce Distinct Conformational Changes in ER alpha and ER beta', *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3999-4004.

[6] J. W. Heinzman, B. Ganem, 'The Mechanism of Sodium Borohydride-Cobaltous Chloride Reductions', *J. Am. Chem. Soc.* **1982**, *104*, 6801-6802.

[7] U. Karmakar, R. Samanta, 'Pd(II)-Catalyzed Direct Sulfonylation of Benzylamines Using Sodium Sulfinates', *J. Org. Chem.* **2019**, *84*, 2850-2861.

[8] M. Penhoat, 'Scope and Limitations of a 1H NMR Method for the Prediction of Substituted Phenols pKa Values in Water, CH_3CN, DMF, DMSO and i-PrOH', *Tetrahedron Lett.* **2013**, *54*, 2571-2574.

[9] H. Gao, J. A. Katzenellenbogen, R. Garg, C. Hansch, 'Comparative QSAR Analysis of Estrogen Receptor Ligands', *Chem. Rev.* **1999**, *99*, 723-744.

[10] R. W. Hartmann, A. Heindl, H. Schonenberger, 'Ring-substituted 1,2-Dialkylated 1,2-Bis(hydroxyphenyl)ethanes. 2. Synthesis and Estrogen Receptor Binding Affinity of 4,4'-, 5,5'-, and 6,6'-Disubstituted Metahexestrols', *J. Med. Chem.* **1984**, *27*, 577-585.

[11] R. G. Parr, Y. Weitao, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, 1994.

[12] Y. Zhao, D. G. Truhlar, 'The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-class Functionals and 12 Other Functionals.', *Theor. Chem. Account* **2008**, *120*, 215–241.

[13] S. Simon, M. Duran, J. J. Dannenberg, 'How Does Basis Set Superposition Error Change the Potential Surfaces for Hydrogen Bonded Dimers?', *J. Chem. Phys.* **1996**, *105*, 11024-11031.

[14] K. Raghavachari, J. S. Binkley, R. Seeger, J. A. Pople, 'Self-Consistent Molecular Orbital Methods. 20. Basis Set for Correlated Wave-functions', *J. Chem. Phys.* **1980**, *72*, 650-654.

[15] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W.

Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.
[16] S. W. Fanning, R. Jeselsohn, V. Dharmarajan, C. G. Mayne, M. Karimi, G. Buchwalter, R. Houtman, W. Toy, C. E. Fowler, R. Han, M. Lainé, K. E. Carlson, T. A. Martin, J. Nowak, J. C. Nwachukwu, D. J. Hosfield, S. Chandarlapaty, E. Tajkhorshid, K. W. Nettles, P. R. Griffin, Y. Shen, J. A. Katzenellenbogen, M. Brown, G. L. Greene, 'The SERM/SERD Bazedoxifene Disrupts ESR1 Helix 12 to Overcome Acquired Hormone Resistance in Breast Cancer Cells', *eLife* **2018**, *7*, e37161.
[17] W. Minor, M. Cymborowski, Z. Otwinowski, M. Chruszcz, 'HKL-3000: the Integration of Data Reduction and Structure Solution - from Diffraction Images to an Initial Model in Minutes', *Acta Crystallogr. D* **2006**, *62*, 859-866.
[18] D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkoczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, P. D. Adams, 'Macromolecular Structure Determination using X-rays, Neutrons and Electrons: Recent Developments in Phenix', *Acta Crystallogr. D* **2019**, *75*, 861-877.
[19] P. Y. Maximov, B. Abderrahman, S. W. Fanning, S. Sengupta, P. Fan, R. F. Curpan, D. M. Q. Rincon, J. A. Greenland, S. S. Rajan, G. L. Greene, V. C. Jordan, 'Endoxifen, 4-Hydroxytamoxifen and an Estrogenic Derivative Modulate Estrogen Receptor Complex Mediated Apoptosis in Breast Cancer', *Mol. Pharmacol.* **2018**, *94*, 812-822.
[20] G. M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, 'Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments', *J. Comput. Aided Mol. Des.* **2013**, *27*, 221-234.
[21] J. Wang, W. Wang, P. A. Kollman, D. A. Case, 'Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations', *J. Mol. Graph. Model.* **2006**, *25*, 247-260.
[22] D. A. Case, H. M. Aktulga, K. Belfon, I. Y. Ben-Shalom, J. T. Berryman, S. R. Brozell, D. S. Cerutti, I. T.E. Cheatham, G. A. Cisneros, V. W. D. Cruzeiro, T. A. Darden, N. Forouzesh, G. Giambaşu, T. Giese, M. K. Gilson, H. Gohlke, A. W. Goetz, J. Harris, S. Izadi, S. A. Izmailov, K. Kasavajhala, M. C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T. S. Lee, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K. M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K. A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, A. Shajan, J. Shen, C. L. Simmerling, N. R. Skrynnikov, J. Smith, J. Swails, R. C. Walker, J. Wang, J. Wang, H. Wei, X. Wu, Y. Wu, Y. Xiong, Y. Xue, D. M. York, S. Zhao, Q. Zhu, P. A. Kollman, *Amber 2023*, University of California, San Francisco, 2023.
[23] R. Salomon-Ferrer, A. W. Gotz, D. Poole, S. Le Grand, R. C. Walker, 'Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald', *J. Chem. Theory Comput.* **2013**, *9*, 3878-3888.
[24] S. L. Grand, A. W. Götz, R. C. Walker, 'SPFP: Speed without Compromise—A Mixed Precision Model for GPU Accelerated Molecular Dynamics Simulations', *Comput. Phys. Commun.* **2013**, *184*, 374-380.
[25] A. W. Gotz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand, R. C. Walker, 'Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born', *J. Chem. Theory Comput.* **2012**, *8*, 1542-1555.
[26] X. He, V. H. Man, W. Yang, T. S. Lee, J. Wang, 'A Fast and High-quality Charge Model for the Next Generation General AMBER Force Field', *J. Chem. Phys.* **2020**, *153*, 114502.
[27] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, 'ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB', *J. Chem. Theory Comput.* **2015**, *11*, 3696-3713.

[28] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, 'Development and Testing of a General Amber Force Field', *J. Comput. Chem.* **2004**, *25*, 1157-1174.
[29] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, 'Comparison of Simple Potential Functions for Simulating Liquid Water', *J. Chem. Phys.* **1983**, *79*, 926-935.
[30] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, 'Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes', *J. Comput. Phys.* **1977**, *23*, 327-341.
[31] T. Darden, D. York, L. Pedersen, 'Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems', *J. Chem. Phys.* **1993**, *98*, 10089-10092.
[32] D. R. Roe, T. E. Cheatham, 3rd, 'PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data', *J. Chem. Theory Comput.* **2013**, *9*, 3084-3095.