

Title Page:

A computational based approach to identify estrogen receptor alpha/beta heterodimer selective ligands

Carlos G. Coriano, Fabao Liu, Chelsie K. Sievers, Muxuan Liang, Yidan Wang, Yoongho Lim, Menggang Yu and Wei Xu

Molecular & Environmental Toxicology Center, Department of Oncology, University of Wisconsin-Madison, 7459 WIMR II, 1111 Highland Avenue, Madison, WI, 53705-2275, USA. (CC and WX)

Department of Oncology, University of Wisconsin-Madison, 7459 WIMR II, 1111 Highland Avenue, Madison, WI, 53705-2275, USA. (CC, FL, CS, YW and WX)

Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison, Wisconsin, U.S.A. Clinical Sciences Center 600 Highland Avenue Madison, WI 53792-4675 (ML and MY)

Division of Bioscience and Biotechnology, BMIC, Konkuk University, Seoul 05029, Republic of Korea (YL)

Running Title Page:

A) Running title: Identification of ER α / β heterodimer selective ligands

B) Corresponding author: Wei Xu, Ph.D.

Email: wxu@oncology.wisc.edu

Telephone: 608-265-5540

Address: 7459 WIMR II, 1111 Highland Ave. McArdle Laboratory, Madison, WI 53705-2275

C) Number of pages: 38

Number of tables: 2

Number of figures: 5

Number of references: 29

Total number of words: 7020

Abstract: 190

Introduction: 725

Discussion: 753

D) Abbreviations:

BRET- Bioluminescent Resonance Energy Transfer Assay

ER-Estrogen receptors

DBD- DNA binding domain

DMSO- Dimethylsulfoxide

DNA- Deoxyribose nucleic acid

DPN- Diarylpropionitrile

E2- 17 β -estradiol

ERE- Estrogen response element

FBS- Fetal bovine serum

ICI- ICI 182,780

LBD- Ligand binding domain

PPT- Propylpyrazole triol

RLU- Raw luciferase units

RLuc- Renilla luciferase

SFS- Stripped FBS

YFP- Yellow fluorescent protein

MOL #108696

Abstract:

The biological effects of estrogens are transduced by two estrogen receptors (ERs), ER α and ER β , which function in dimer forms. The ER α/α homodimer promotes and the ER β/β inhibits estrogen dependent growth of mammary epithelial cells, the functions of ER α/β heterodimers remain elusive. Using compounds that promote ER α/β heterodimerization, we have shown that ER α/β heterodimers appeared to inhibit tumor cell growth and migration *in vitro*. Further dissection of ER α/β heterodimer functions was hampered by the lack of ER α/β heterodimer specific ligands. Herein we reported a multistep workflow to identify the selective ER α/β heterodimer-inducing compound. Phytoestrogenic compounds were first screened for ER transcriptional activity using reporter assays and ER dimerization preference using a Bioluminescent Resonance Energy Transfer Assay (BRET). The top hits were subjected to *in silico* modeling to identify the pharmacophore that confers ER α/β heterodimer specificity. The pharmacophore encompassing seven features that are potentially important for the formation of the ER α/β heterodimer was retrieved and subsequently used for virtual screening of large chemical libraries. Four chemical compounds were identified that selectively induce ER α/β heterodimers over their respective homodimers. Such ligands will become unique tools to reveal the functional insights of ER α/β heterodimers.

MOL #108696

Introduction:

The biological effects of estrogenic compounds are mediated by two estrogen receptors (ERs), namely ER α and ER β . These receptors are expressed in a cell-type and tissue-specific manner, yet they can also co-localize within the same cell and their presence varies based on different disease state (Lau et al., 1999; Leygue et al., 1998; Nilsson and Gustafsson, 2013; Powell et al., 2012; Weihua et al., 2003). Both ERs share a conserved nuclear receptor domain structure that encompass a DNA binding domain (DBD), ligand binding domain (LBD), a central hinge region and two activation functional domains (AF). The ligand binding to ER α or ER β induces a conformational change that leads to receptor dimerization, where either homodimers (ER α/α or ER β/β) or heterodimers (ER α/β) can be formed.

The existence of the ER α/β heterodimer was first described 20 years ago using *in vitro* translated receptors and an estrogen response element (ERE) in a gel shift assay, Cowley and colleagues showed that ER heterodimers could bind to a consensus ERE and recruit coactivators *in vitro* (Cowley et al., 1997). Similar observations were made by others (Pace et al., 1997; Tremblay et al., 1999). Pettersson et al. showed direct interaction between ER β and ER α in a GST pull-down assay and binding of the heterodimer to DNA (Pettersson et al., 1997). Two dimerization domains were mapped to DBD and LBD, respectively (Brzozowski et al., 1997; Pace et al., 1997). ER heterodimers were shown to form in a ligand dependent and independent manner *in vitro* (Pace et al., 1997). Recent technical advances confirmed the formation of ER α/β heterodimer *in vivo*. Our lab developed a BRET¹ (Bioluminescence Resonance Energy

MOL #108696

Transfer) assay to monitor ER dimerization in live cells (Powell and Xu, 2008). BRET assays revealed that the types of ER dimer pair being formed depend on the chemical characteristics of the ligand and its concentration (Powell and Xu, 2008). Moreover, the ER α / β heterodimers have been detected *in vivo* using molecular imaging techniques (Paulmurugan et al., 2011) and in breast cancer tissues using proximity ligation assay (Iwabuchi et al., 2017). Evidence also shows that the ER α / β heterodimer is transcriptionally active and may regulate a distinct set of genes from their respective homodimers (Tremblay et al., 1999).

In contrast to the established role that the ER α / α homodimer is a driver of estrogen mediated cellular proliferation and ER β / β homodimers elicit an anti-proliferative and pro-apoptotic effect, the function of ER α / β heterodimers in the biological processes is the least understood. Unlike the ER α / α and ER β / β homodimers where subtype-specific ligands for ER α and ER β aided in elucidating their function (Lindberg et al., 2003; Weihua et al., 2003), ligands that specifically induce ER α / β heterodimer have not been identified, largely due to the absence of a full-length crystal structure for the ER α / β heterodimer.

Indirect evidence suggesting that ER α / β heterodimers might have an anti-proliferative role in breast cancer cells have previously been reported (Hall and McDonnell, 1999; Powell et al., 2012). Endoxifen, the primary metabolite of tamoxifen with growth inhibitory effects, stabilizes ER β and induces the formation of ER α / β heterodimers in cells expressing both ERs (Wu et al., 2011). Furthermore, high-throughput BRET assays identified a phytoestrogen, cosmosiin, that favors ER α / β heterodimer formation (Powell et al., 2012). This ER α / β heterodimer-inducing compound elicited anti-proliferative effects

MOL #108696

in prostate and breast cancer cells. Although cosmosiin induces the formation of ER α / β heterodimers but not the pro-proliferative ER α / α homodimers, it is only effective at high concentrations (e.g. 10 μ M) and also slightly induces ER β / β homodimers (Powell et al., 2012). More potent and selective ER α / β heterodimer-inducing ligands are needed to elucidate the biological functions of heterodimers.

Herein we describe a multistep screening strategy (i.e., cell based assays and *in silico* modeling) to identify ER α / β heterodimer selective ligands. Reporter assays and BRET assays were employed to screen a small library of flavonoid type phytoestrogenic compounds, from which a pharmacophore model was generated using SYBYL-GALAHAD (Genetic Algorithm with Linear Assignment of Hypermolecular Alignment of Database). The pharmacophore model was subsequently used in a 3D search query of two commercial chemical databases to identify new active structures. Four compounds were identified from the *in silico* screen that selectively induce ER α / β heterodimers. We showed that the representative compounds induce expression of putative ER α / β target genes by co-recruiting ER α and ER β to the target gene promoter. Such ER α / β -selective compounds will be exploited for determining the biological functions of ER α / β heterodimers, their downstream effectors and target genes.

MOL #108696

Materials and methods:

Cell culture and Chemicals

Cell culture media was obtained from Invitrogen (Carlsbad, CA). HEK293 cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% Gibco Fetal Bovine Serum (FBS; Invitrogen) at 37°C and 5% CO₂. T47D-KBLuc cells were routinely maintained in RPMI 1640 medium supplemented with 10% FBS and 100 units/mL (1%) penicillin/streptomycin. Experiments were conducted in phenol free media and dextran charcoal-stripped fetal bovine serum purchased from Hyclone (Logan, Utah, USA).

Compounds were dissolved in 100% dimethyl sulfoxide (DMSO) and finally diluted in culture medium prior to the assay. 17 β -estradiol (E2) and ICI 182,780 were purchased from Sigma. Thirty-one compounds used in our initial screening were a gift from the Lim Lab and have been previously described (Hwang et al., 2011; Hyun et al., 2010; Shin et al., 2011). They were chosen for screening based on their structural similarity the lead compound cosmosiin which was previously identified to induce ER α / β heterodimerization (Powell et al., 2012). Test compounds were purchased from ChemBridge (<http://www.chembridge.com>) and Maybridge (<http://www.maybridge.com>).

BRET assays

Dimerization of ERs was measured by BRET assays as previously described in [3]. Briefly, HEK293 cells were transfected with either a single BRET fusion plasmid (pCMX-ER α -RLuc or pCMX-RLuc-ER β) or co-transfected with RLuc and YFP BRET fusions

MOL #108696

(pCMX-ER α -RLuc+pCMX-YFP-ER β for ER α /ER β heterodimers; pCMX-ER α -RLuc+pCMX-ER α -YFP for ER α homodimers; or pCMX-RLuc-ER β +pCMX-YFP-ER β for ER β homodimers). 24 hours post-transfection, cells were trypsinized and plated in a 96-well white-bottom microplate and incubated with ligands for 1 hour. Coelenterazine h (Promega, Madison, WI) was added in PBS at a final concentration of 5 μ M, and 460 nm and 530 nm emission detection measurements were immediately taken at 0.1 second per wavelength read per well on a Perkin Elmer Victor 3-V plate reader. Similar assays were done using E2-binding defective mutants of the LBDs of ER α and ER β , ER α G521R-RLuc and YFP-ER β G491R. Each compound was an independent experiment tested in a dose response with three biological replicates per dose. For each condition (ER α / α , ER β / β and ER α / β), a two-way ANOVA with random effect was conducted to obtain *P* values for each comparison of the individual compounds with DMSO controls. Then these *p* values were adjusted by multiple comparison analysis to control false discovery rate (FDR) less than 0.05.

ER Luciferase reporter assays using T47D-KBLuc cells

T47D-KBluc is a well characterized cell line for the screening of estrogenic compounds (Wilson et al., 2004). These cells express both ER α and ER β and have been stably transfected with pGL2.TATA.lnr.luc.neo which contains three estrogen responsive elements upstream of a luciferase reporter gene. Cells were seeded in 96-well plates at an initial concentration of 1×10^4 cells per well in RPMI 1640 phenol free medium supplemented with 10% charcoal stripped FBS for 24 hours at 5% CO₂ atmosphere at 37 °C. Cells were allowed to attach overnight and media was removed and replaced with

MOL #108696

media containing 10 μ M compound. 10 nM E2 and 1% DMSO were used as positive and negative controls, respectively. The potent ER antagonist ICI 182,780, was used for counter-screen to determine ER specificity. Cells were incubated with compound for 18-24 hrs at 37°C in 5% CO₂. Following incubation with compounds, luciferase was measured using the Bright-Glo™ Luciferase Assay System (Promega, cat#E2620) on a Perkin Elmer Victor 3-V plate reader. Luciferase activity was normalized according to protein concentration. Values were expressed as fold change over DMSO (mean value of induction as a multiple of the value of vehicle controls) and error bars represent standard deviation.

Quantitative Real-time PCR

Total RNA was extracted from the cells using Trizol reagent (Invitrogen, Carlsbad, CA). The first-strand cDNA was synthesized by RevertAid First Strand cDNA Synthesis kit (Thermo) according to the manufacturer's instructions. Q-PCR was conducted using SYBR Green dye (Roche Scientific, Basel Switzerland) and a CFX96 instrument (BioRad, Hercules, CA). Primer sequences (IDT, Coralville, IA) used in this study were as follows:

BAG1-qRT-F: GCCCAAGGATTTGCAAGCTG, BAG1-qRT-R:

CTGTGTCACACTCGGCTAGG; ATP6V0E1-qRT-F: CCTCACTGTGCCTCTCATTGT,

ATP6V0E1-qRT-R: AGCAAACACTGAACAGGTCACCA. BAG-ChIP-F:

AGGAAGCTCTGATAGAAGGCAGA, BAG-ChIP-R:

AGAACAGTCCACAGGAAGAGGT; ATP6V0E1-ChIP-F: CCCCTGGCAGTTTCGTAC,

ATP6V0E1-ChIP-R: TCTTGTTTCATAATTTGACTTTGGAG.

MOL #108696

Chromatin immunoprecipitation (ChIP)

Flag-tagged ER β was stably expressed in MCF7 cells by retroviral induction. MCF7-ER β cells were cultured in 10-cm dish and cross-linked with 1% Formaldehyde for 10 min at room temperature. Crosslink was quenched for 5 min at room temperature by the addition of glycine to a final concentration of 0.125 M. Anti-Flag antibody (Sigma-Aldrich) and anti-ER α (Santa Cruz, HC-20) were used for ChIP assays. ChIP assays were performed as described previously (Zeng et al., 2016; Zeng and Xu, 2015). The experiment was done in triplicate samples of biological replicates. Statistical testing was performed using the unpaired two-tailed Student's t-test analysis. Experiments were repeated at least twice. *P* value < 0.05 (*) was considered statistically significant.

Fluorescence Polarization Competition Ligand Binding assays

The binding affinity of ligands for ER α and ER β were measured using PolarScreen ER Competitive Binding Assay Kit (Invitrogen). Purified ER α and ER β (30 nM and 20 nM), were incubated with serial dilutions of test compounds (1 mM to 10 nM) and fluorescein-labeled E2. Fluorescence polarization was measured using a Victor X5 microplate reader (PerkinElmer). Approximate IC₅₀ values were determined by GraphPad Prism Software (Graph-Pad Software Inc.) from competitive binding curves.

Preparation of the initial ligands

All computational studies were done using SYBYL molecular modeling package (Tripos, St. Louis, MO) in a Stereo 3D Dell T5500 molecular graphics computer (Intel dual quad, Nvidia FX 4800 graphics). All of the structures used were built using the Sketch module,

MOL #108696

energy minimized and prepared using SYBYL's Ligand Preparation module. The Quick 3D parameter was used where 3D coordinates were generated and charges were neutralized.

Generation of Ligand Based Pharmacophore Hypothesis and Virtual Screening

Pharmacophore hypothesis was generated using Genetic Algorithm with Linear Assignment of Hypermolecular Alignment of Database (GALAHAD) module of SYBYL software suite. There were seven compounds in the training set to generate the pharmacophore hypothesis. The GALAHAD module was run for 100 generations with a population size of 45, at least 5 molecules were required to hit for the program to consider it a pharmacophoric feature. Default values were used for all other settings. Between all of the models, the one with the best Energy, Sterics and Pharmacoric Similarity values based on Pareto ranking was selected as the best model. For 3D virtual screening, the generated pharmacophore hypothesis model was converted into a 3D search query using the Unity-3D module.

3D Virtual Screening of two commercial databases

The selected pharmacophore model was validated and converted into a UNITY query for pharmacophore guided virtual screening studies. The query was then used for screening two commercial chemical databases Maybridge (<http://www.maybridge.com>), and Chembrige (<http://chembridge.com>) which were obtained from the ZINC public database. A flexible 3D search was executed and no filters or restrictions were applied. UNITY module uses a conformationally flexible 3D-searching algorithm to result in rapid

MOL #108696

identification of molecules that match with the given pharmacophore. Compounds which had their chemical groups spatially overlap with the features of the pharmacophore model were captured as hits. Subsequent hits were then confirmed to match all 7 key pharmacophoric features by visual analysis. Hits were then ranked by Qfit and by SYBYLs integrated ranking features.

Results

Identification of ER agonists using T47D-KBluc reporter cell line

The T47D-KBluc breast cancer cell line is a well-characterized cell line that has an ERE-driven luciferase reporter stably integrated. It is considered a versatile cell system for screening estrogenic compounds because it expresses both ER α and ER β (Wilson et al., 2004). Using this cell line, 37 flavonoid subclass chemical compound library (Table 1) were screened due to their structural similarity to cosmosiin, a previously identified ER α / β heterodimer inducing compound (Powell et al., 2012). All compounds were tested at one final concentration of 10 μ M, because at this concentration even weak estrogenic compounds are able to activate ER transcriptional activity in T47D-KBLuc (Powell et al., 2012). With a two-fold cutoff, 13 of 37 compounds from three out of the four subclasses were identified as hits (Figure 1 and supplemental Figure 1). Seven compounds in the flavone subset were identified as hits. Compared to the DMSO control, compounds **3**, **5**, **6**, and **7** showed the moderate activation and compounds **15**, **16** elicited almost 10-fold induction and **17** nearly a 20-fold induction of ERE reporter. Three compounds **18**, **23**,

MOL #108696

and 24 in the flavanone subset and two compounds **26 and 29** in the isoflavone subset were retained as hits. Compound **28** is genistein, an isoflavone known to be an ER agonist that has been shown to induce all three ER dimer pairs (Powell and Xu, 2008). The 12 compounds were then subjected to a counter-screen in the presence of ER antagonist ICI 182,780. Co-treatment of ICI 182,780 completely ablated ERE-luciferase activation, demonstrating that their transcriptional response is ER dependent (Supplemental Figure 2).

In vitro fluorescence polarization assays were used to determine the relative binding affinities of the phytochemicals that activated the ER in reporter assays. Fluorescence polarization assays are a competitive ligand binding assay that measures the replacement of a fluorescein-labeled E2 by unliganded compounds from the ligand binding domains of ER α and ER β . The dose-response curves for representative ligands and the relative binding affinities are shown in Figure 1B with half maximal inhibitory concentration (IC₅₀) ranged from 1.45 μ M to 721 μ M.

Ligand-induced ER dimerization measured by Bioluminescence Resonance Energy Transfer (BRET) assay

The 12 compounds were subsequently tested for their ability to induce ER α/α , ER β/β and ER α/β ER dimers in BRET assay (Figure 2A-C). Of the flavones, compounds **3, 5, 6,** and **7** induced dimerization of ER α/α homodimers and ER α/β heterodimers at 10 μ M. Compounds **15** induced all three dimer pairs at 1 μ M. Compound **16** induced ER α/α and ER β/β at 1 μ M and all dimer pairs at 10 μ M. Compound **17** induced ER β/β and ER α/β dimerization while restricting the induction of ER α/α homodimers at 1 μ M but only induced

MOL #108696

the formation of ER α / β at 10 μ M (Figure 2A-C). From the flavanones, ER α / β heterodimers were selectively induced by compounds **23** and **24** at 10 μ M, whereas compound **18** induced ER β / β and ER α / β dimerization at 1 μ M (Figure 2A-C). Of the isoflavone subclass (Figure 2A-C), compound **26** induced ER α / β and ER β / β dimerization at 10 μ M, whereas compound **29** induced ER α / β and ER β / β dimerization at 1 μ M.

To determine if newly identified compounds from the reporter assay and BRET assay indeed activate ER target gene expression, we measured mRNA levels of two ER target genes BAG1 and ATP6V0E1 after compound 29 treatment, using compound 28 (genistein) as a positive control. These compounds were selected as they exhibited the highest activities in the reporter assay (Figure 1A). Because most breast cancer cell lines do not express ER β , we constructed MCF7 cells that stably express Flag-tagged ER β (Figure 2D). Treatment of Flag-ER β MCF7 cells with compounds 28 and 29 significantly increased the mRNA levels of BAG1 and ATP6V0E1 as compared to DMSO control (Figure 2E, F). BAG1 has been implicated to be an ER β / β specific target gene and ATP6V0E1 as an ER α / β specific target gene, respectively (Grober et al., 2011). Because compound 29 was found to induce ER α / β and ER β / β dimerization but not ER α / α homodimers at 1 μ M, we went on to examine if compound 29 differentially recruited ER α and ER β to the target gene promoters at a dose (1 μ M) that elicits dimer specificity. ChIP-qPCR analysis showed that compound 29 treatment increased ER β association at the promoters of both BAG1 and ATP6V0E1 genes as compared to DMSO control. In contrast, compound 29 was only able to increase ER α recruitment to ATP6V0E1 but not to BAG1 promoter. This result is consistent with the classification of compound 29 as an ER β / β and ER α / β dimer inducer by BRET assay and that ATP6V0E1 is likely regulated

MOL #108696

by ER α / β heterodimer vs. BAG1 which is likely regulated by ER β / β homodimer (Figure 2G-H).

Of the tested compounds only three selectively induced ER α / β heterodimerization at select concentrations (compounds **17**, **23** and **24**), but three other compounds **18**, **26** and **29** preferentially induced ER α / β and ER β / β dimers over ER α / α homodimers. Interestingly, compounds **17**, **18**, **23**, **24**, **26** and **29** that induced ER α / β and ER β / β dimers showed higher binding affinity for ER β than for ER α (Figure 1B). These six compounds, together with ER α / β heterodimer inducing compound *cosmosiin*, constitute a lead heterodimer selective compound library for pharmacophore development.

Pharmacophore development using the GALAHAD module in SYBYL

The structures of 37 compounds from the initial dataset (Table 1) were built into the SYBYL software platform using the Sketch module where hydrogens were added to every structure and energy minimized and saved as Mol2 files. All structures were then converted into a 3D conformation for each input structure.

Compounds **17**, **18**, **23**, **24**, **26** and **29** plus *cosmosiin* were used as the training set (Figure 3) to build a pharmacophore model in the GALAHAD module. Ligands were flexibly aligned to each other completely independent of a template. This generates a molecular alignment based on the pharmacophoric features of the final conformations of the training set. Twenty pharmacophore models were generated by GALAHAD; each of the models represents a different trade off among competing criteria (Supplemental Table 1). These models contained the same number of features and specificity. The most

MOL #108696

significant pharmacophore hypothesis models are characterized by assessing the relation between maximizing pharmacophore consensus, maximizing steric consensus, and minimizing conformer potential energy (Caballero, 2010). Within each set of hypotheses, models were first ranked by Pareto surface score (Sterics vs. Energy), of where the best model has the lowest energy and the highest steric score, as illustrated in the upper left-hand corner of Figure 3B. Concerning Energy and Pharmacophoric similarity criteria, the best model with low energy and high H-Bond score lies in the upper left hand corner of the graph in Figure 3B. Finally, the best model judged by the Pharmacophoric similarity and Sterics scores lies at the upper-right corner (Figure 3B bottom). In Figure 3B, the ideal model in each ranking is depicted by a blue circle. Taking all models into consideration, Model_006 (represented with a red diamond in Fig. 3) had a balanced consensus ranking in all three criteria thus was chosen as the best GALAHAD model (Figure 3C).

Model_006 is comprised of one conformer for each molecule in the training set. All conformers aligned represent low-energy conformations of the molecules, and the final alignment shows a satisfactory superimposition of the pharmacophoric points. Model_006 contains 7 key features including 3 hydrophobes, 3 acceptor atoms and 1 donor atom. The pharmacophore model clearly shows the importance of the hydrophobic center that is essential in ER pharmacophores for ER α and ER β selective ligands (Anstead et al., 1997; Brzozowski et al., 1997). The pharmacophore model was validated for its ability to identify ER α / β heterodimer selective compounds from the full data set (Supplemental Table 2).

MOL #108696

3D Virtual Screening of the ChemBridge and Maybridge databases identified 167 compounds as potential hits

The pharmacophore model was converted into a 3D search query using SYBYL's Unity 3D module. The search query was then used to screen the commercial chemical databases from ChemBridge and Maybridge. Both chemical libraries were retrieved from the Zinc Database (<http://zinc.docking.org/>), a free database of commercially-available libraries for virtual screening. Flexible 3D screening with no restrictions of both databases was performed using the UNITY tool (Figure 4A). A total number of 900 initial molecules were identified as hits, many of which contained different chemical scaffolds.

The hits were manually inspected to ensure all chemical groups from the compounds spatially overlapped with the corresponding features of the pharmacophore model. After visual inspection, 81% of the hits failed to match all 7 pharmacophoric features thus were discarded. The 167 remaining hits contained 19 different core scaffolds that match the spatial arrangements of our pharmacophore hypothesis. Subsequently, the hits were ranked using SYBYL's integrated ranking features (Supplemental Table 3 & 4), among which the top 22 hits were purchased and further characterized (Supplemental Table 5).

Validation and Characterization of hits

The "hits" were confirmed to activate ER transcription in T47D-KBLuc reporter assay at 10 μ M final concentration (Supplemental Figure 3) and to induce dimerization in BRET assays (Supplemental Table 6). The ability of compounds to induce all three ER dimer pairs were tested at increasing doses between 1 and 20 μ M in BRET assays (data not

MOL #108696

shown). Four compounds selectively induced the ER α / β heterodimer but not ER α / α nor ER β / β homodimers at specific concentrations (Figure 4B). The lowest concentrations at which these four compounds selectively induce ER α / β heterodimer are 1 μ M.

The binding affinity of compounds **4**, **6**, **9**, and **10** to ER α and ER β were measured by *in vitro* fluorescence polarization assay (Figure 4C). The relative binding affinities are calculated as IC₅₀. Compounds **9** and **10** elicit the highest binding affinity. The IC₅₀ values for compound **9** to ER α and ER β were 1.4 μ M and 2.0 μ M, respectively. The IC₅₀ values for compound **10** to ER α and ER β were 1.9 μ M and 3.2 μ M, respectively. Although like compound **29** that was used to build pharmacophore model, compound **9** and **10** induced ER α / β heterodimer at 1 μ M, but they elicited improved binding affinity to ER α and ER β (Figure 4C), suggesting that *in silico* modeling expedites identification of stronger ERs agonists with similar heterodimer specificity. Thus compound **9** and **10** would be better compounds to pursue for probing ER α / β heterodimer functions.

ER α is the dominant heterodimeric partner in the presence of selective ER α / β heterodimer compounds

We previously reported that E2 induces heterodimer formation by binding to ER α (Powell and Xu, 2008). To examine whether the selective heterodimer inducing compounds also induce heterodimer via binding to ER α , BRET assays were performed with a combination of wild type and mutant ER α and ER β constructs (Powell and Xu, 2008). The expressed mutant proteins contained a single mutation in the LBD (ER α G521R and ER β G491R) of receptors which ablate ligand binding (Powell and Xu, 2008; Tremblay et al., 1999). A combination of wild type and mutant ER α and ER β fusion proteins were used:

MOL #108696

ER α G521R-RLuc with wild-type YFP-ER β , YFP-ER β G491R with wild-type ER α -RLuc, wild-type YFP-ER β with wild-type ER α -RLuc, and ER α G521R-RLuc with YFP-ER β G491R. As has been previously reported for E2 (Powell and Xu, 2008), ligand-binding competent ER α LBD but not ER β LBD is required for heterodimerization in the presence of compounds, reinforcing that ER α is the dominant partner for heterodimerization (Figure 5).

2.5 Discussion

Current ER-positive breast cancer therapies target ER α , either using selective estrogen receptor modulators (SERMs) to inhibit ER α transcriptional activity or using selective estrogen receptor degrader (SERD) to reduce ER α protein levels. However, the therapeutic potential of ER β in breast cancer has been poorly investigated. Our previous studies using ER α / β heterodimer-selective ligands show that ER β , via heterodimerization with ER α , can antagonize the pro-proliferative effects of ER α , rendering the heterodimer as a novel preventive or therapeutic target for hormone-dependent diseases. However, few ER α / β heterodimer-inducing selective compounds have been discovered and they generally elicit only weak binding affinities to ERs. Therefore, the goal of this study was to combine computational and experimental approaches to identify compounds with improved binding affinity and dimerization specificity.

Emerging biochemical evidence support the formation of ER α / β heterodimers when two receptors are co-expressed (Cowley et al., 1997; Pettersson et al., 1997), in particular, ER α / β heterodimers were recently detected in breast tissues using proximity ligation assay (PLA) (Iwabuchi et al., 2017). However, the functions of ER α / β heterodimer remain

MOL #108696

elusive due to the lack of a crystal structure and heterodimer specific compounds. Uncovering the biological function of the ER α / β heterodimer is important for understanding ER signaling and designing ER targeted therapeutics based on receptor dimerization status. The main distinction of heterodimer inducing compounds from the existing SERM and SERDs is that they target different steps in ER activation. ER heterodimer compounds target ER dimerization, a step between the ligand binding and the receptor association with chromatin. In our previous published reporting, we have shown that SERMs such as tamoxifen, raloxifene, and the full ER antagonist ICI 182,780 do not interfere with the formation of all three dimer pairs (Powell and Xu, 2008). Although more studies are needed to demonstrate that ER α / β heterodimer indeed serves as a therapeutic target, the concept of inducing ER β to pair with ER α , thus antagonizing ER α 's proliferative function, is distinct from the existing breast cancer therapies to target ER α alone.

We reason that identifying and improving chemical probes would be the essential step towards understanding the biological role of the ER α / β heterodimers. Natural phytoestrogens often elicit higher binding affinity to ER β than to ER α (Kuiper et al., 1997; Kuiper et al., 1998). Several phytoestrogens showed slight selectivity for ER α / β heterodimer were found to be anti-proliferative in cancer cell lines co-expressing ER α and ER β (Powell et al., 2010; Powell et al., 2012; Powell and Xu, 2008). However, the slight selectivity and low potency of these compounds prevent definitive elucidation of the functions of ER α / β heterodimers. Pharmacophore based techniques and virtual screening have successfully been employed for the discovery of ER subtype selective ligands (Huang et al., 2015). Herein, using a combination of cell-based assays (i.e. reporter assay

MOL #108696

and BRET) and pharmacophore modeling and virtual screening, we identified four ER α / β heterodimer selective ligands (Table 2) with improved efficacy than cosmisiin, a previously identified compound with slight preference for ER α / β heterodimer. The main hurdle for identifying ER α / β heterodimer selective ligands lies in the lack of a crystal structure. Ligand binding is necessary but insufficient for the formation of ER α / β heterodimer. Previous studies suggest that ligand binding is essential to induce a conformational change of ER to accommodate helix 12 in the functional dimers. In this process, ER α and ER β appear to play separate roles such that ligand-bound ER α is the dominant partner for heterodimer formation. We have shown that ER β -subtype specific ligands promote the formation of ER β / β homodimers but ER α subtype specific ligands could induce both ER α / α homodimers and ER α / β heterodimers (Powell and Xu, 2008). Because of the lack of protein crystal structures to be able to build a structure-based pharmacophore model for a virtual ligand screen, we combined a multistep screening strategy with a ligand based pharmacophore model to identify ER α / β heterodimer selective ligands. We confirmed that ligand binding is necessary but insufficient for inducing ER dimerization. Furthermore, the formation of ER homo- vs. heterodimers appears to be ligand concentration dependent (Figure 2 and 4). Our results also showed that a ligand must induce conformational change of ER α in a manner that it preferentially selects the other ER subtype as a partner (Figure 5). Finally, we characterized the estrogenic activity and dimerization ability of 59 compounds, leading to the identification of four ER α / β heterodimer selective ligands. To our knowledge, building a pharmacophore model to identify the chemical features responsible for induction of ER α / β heterodimer is unprecedented. Thus, the more selective and potent compounds identified

MOL #108696

in this study will serve as useful probes to elucidate ER α / β heterodimer functions *in vitro* and *in vivo*.

MOL #108696

Acknowledgments:

The authors would like to thank the Small Molecule Screening Facility at the University of Wisconsin-Madison for providing access to computational hardware. We would also like to thank Dr. William Ricke (University of Wisconsin-Madison) for generously sharing cell lines. We also thank Carol J. Diaz-Diaz for helpful discussions.

MOL #108696

Authorship Contributions:

Participated in research design: Coriano, Liu and Xu

Conducted experiments: Coriano, Liu, Wang and Sievers

Performed data analysis: Coriano, Liu and Wang

Performed statistical analysis: Liang and Yu

Wrote or contributed to the writing of the manuscript: Coriano, Liu and Xu

References:

- Anstead GM, Carlson KE and Katzenellenbogen JA (1997) The estradiol pharmacophore: ligand structure-estrogen receptor binding affinity relationships and a model for the receptor binding site. *Steroids* **62**(3): 268-303.
- Brzozowski AM, Pike AC, Dauter Z, Hubbard RE, Bonn T, Engström O, Ohman L, Greene GL, Gustafsson JA and Carlquist M (1997) Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **389**(6652): 753-758.
- Caballero J (2010) 3D-QSAR (CoMFA and CoMSIA) and pharmacophore (GALAHAD) studies on the differential inhibition of aldose reductase by flavonoid compounds. *Journal of molecular graphics & modelling* **29**(3): 363-371.
- Cowley SM, Hoare S, Mosselman S and Parker MG (1997) Estrogen receptors alpha and beta form heterodimers on DNA. *The Journal of biological chemistry* **272**(32): 19858-19862.
- Grober OMV, Mutarelli M, Giurato G, Ravo M, Cicatiello L, De Filippo MR, Ferraro L, Nassa G, Papa MF, Paris O, Tarallo R, Luo SJ, Schroth GP, Benes V and Weisz A (2011) Global analysis of estrogen receptor beta binding to breast cancer cell genome reveals an extensive interplay with estrogen receptor alpha for target gene regulation. *Bmc Genomics* **12**.
- Hall JM and McDonnell DP (1999) The estrogen receptor beta-isoform (ERbeta) of the human estrogen receptor modulates ERalpha transcriptional activity and is a key regulator of the cellular response to estrogens and antiestrogens. *Endocrinology* **140**(12): 5566-5578.
- Huang W, Wei W, Yang Y, Zhang T and Shen Z (2015) Discovery of Novel Selective ER α /ER β Ligands by Multi-pharmacophore Modeling and Virtual Screening. *Chem Pharm Bull (Tokyo)* **63**(10): 780-791.
- Hwang D, Hyun J, Jo G, Koh D and Lim Y (2011) Synthesis and complete assignment of NMR data of 20 chalcones. *Magnetic resonance in chemistry* **49**(1): 41-45.
- Hyun J, Woo Y, Hwang DS, Jo G, Eom S, Lee Y, Park JC and Lim Y (2010) Relationships between structures of hydroxyflavones and their antioxidative effects. *Bioorg Med Chem Lett* **20**(18): 5510-5513.
- Iwabuchi E, Miki Y, Ono K, Onodera Y, Suzuki T, Hirakawa H, Ishida T, Ohuchi N and Sasano H (2017) In situ detection of estrogen receptor dimers in breast carcinoma cells in archival materials using proximity ligation assay (PLA). *The Journal of Steroid Biochemistry and Molecular Biology* **165**, Part B: 159-169.
- Kuiper GG, Carlsson B, Grandien K, Enmark E, Häggblad J, Nilsson S and Gustafsson JA (1997) Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and beta. *Endocrinology* **138**(3): 863-870.
- Kuiper GG, Lemmen JG, Carlsson B, Corton JC, Safe SH, van der Saag PT, van der Burg B and Gustafsson JA (1998) Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta. *Endocrinology* **139**(10): 4252-4263.
- Lau KM, Mok SC and Ho SM (1999) Expression of human estrogen receptor-alpha and -beta, progesterone receptor, and androgen receptor mRNA in normal and malignant ovarian epithelial cells. *Proc Natl Acad Sci U S A* **96**(10): 5722-5727.
- Leygue E, Dotzlaw H, Watson PH and Murphy LC (1998) Altered estrogen receptor alpha and beta messenger RNA expression during human breast tumorigenesis. *Cancer research* **58**(15): 3197-3201.
- Lindberg MK, Movérare S, Skrtic S, Gao H, Dahlman-Wright K, Gustafsson JA and Ohlsson C (2003) Estrogen receptor (ER)-beta reduces ERalpha-regulated gene transcription, supporting a "ying yang" relationship between ERalpha and ERbeta in mice. *Molecular endocrinology* **17**(2): 203-208.

MOL #108696

- Nilsson S and Gustafsson J (2013) Estrogen Receptors: Their Actions and Functional Roles in Health and Disease, in *Nuclear Receptors, Proteins and Cell Regulation* pp 91-141.
- Pace P, Taylor J, Suntharalingam S, Coombes RC and Ali S (1997) Human estrogen receptor beta binds DNA in a manner similar to and dimerizes with estrogen receptor alpha. *The Journal of biological chemistry* **272**(41): 25832-25838.
- Paulmurugan R, Tamrazi A, Massoud TF, Katzenellenbogen JA and Gambhir SS (2011) In vitro and in vivo molecular imaging of estrogen receptor α and β homo- and heterodimerization: exploration of new modes of receptor regulation. *Molecular endocrinology* **25**(12): 2029-2040.
- Pettersson K, Grandien K, Kuiper GG and Gustafsson JA (1997) Mouse estrogen receptor beta forms estrogen response element-binding heterodimers with estrogen receptor alpha. *Molecular endocrinology* **11**(10): 1486-1496.
- Powell E, Huang SX, Xu Y, Rajski SR, Wang Y, Peters N, Guo S, Xu HE, Hoffmann FM, Shen B and Xu W (2010) Identification and characterization of a novel estrogenic ligand actinopolymorphol A. *Biochemical pharmacology* **80**(8): 1221-1229.
- Powell E, Shanle E, Brinkman A, Li J, Keles S, Wisinski KB and Huang W (2012) Identification of estrogen receptor dimer selective ligands reveals growth-inhibitory effects on cells that co-express ER α and ER β . *PLoS one* **7**(2): e30993.
- Powell E and Xu W (2008) Intermolecular interactions identify ligand-selective activity of estrogen receptor alpha/beta dimers. *Proceedings of the National Academy of Sciences of the United States of America* **105**(48): 19012-19017.
- Shin SY, Woo Y, Hyun J, Yong Y, Koh D, Lee YH and Lim Y (2011) Relationship between the structures of flavonoids and their NF- κ B-dependent transcriptional activities. *Bioorganic & medicinal chemistry letters* **21**(20): 6036-6041.
- Tremblay GB, Tremblay A, Labrie F and Giguère V (1999) Dominant activity of activation function 1 (AF-1) and differential stoichiometric requirements for AF-1 and -2 in the estrogen receptor alpha-beta heterodimeric complex. *Molecular and cellular biology* **19**(3): 1919-1927.
- Weihua Z, Andersson S, Cheng G, Simpson ER, Warner M and Gustafsson JA (2003) Update on estrogen signaling. *FEBS letters* **546**(1): 17-24.
- Wilson VS, Bobseine K and Gray LE (2004) Development and characterization of a cell line that stably expresses an estrogen-responsive luciferase reporter for the detection of estrogen receptor agonist and antagonists. *Toxicol Sci* **81**(1): 69-77.
- Wu X, Subramaniam M, Grygo SB, Sun Z, Negron V, Lingle WL, Goetz MP, Ingle JN, Spelsberg TC and Hawse JR (2011) Estrogen receptor-beta sensitizes breast cancer cells to the anti-estrogenic actions of endoxifen. *Breast cancer research* **13**(2): R27.
- Zeng H, Lu L, Chan NT, Horswill M, Ahlquist P, Zhong X and Xu W (2016) Systematic identification of Ctr9 regulome in ERalpha-positive breast cancer. *Bmc Genomics* **17**(1): 902.
- Zeng H and Xu W (2015) Ctr9, a key subunit of PAFc, affects global estrogen signaling and drives ERalpha-positive breast tumorigenesis. *Genes Dev* **29**(20): 2153-2167.

MOL #108696

Footnotes

This work was supported by the National Institutes of Health National Cancer Institute (NIH/NCI) [R01 CA213293] to W.X.; National Institute of Health Institute National Institute of Environmental Health Sciences (NIH/NIEHS) [T32ES007015] to trainee C.C.; partial support of this work by the University of Wisconsin Comprehensive Cancer Center Support Grant [P30CA014520]; and National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) [U54DK104310].

CC and FL contributed equally.

MOL #108696

Legends for figures

Figure 1: Transcriptional and ligand binding assays of 37 flavonoid compounds. (A) T47D-KBLuc transcriptional assays showing ERE-luciferase reporter activity of 13 of 37 flavonoid compounds from 4 different subclasses, revealed 13 phytoestrogenic compounds able to transcriptionally activate ER in a dose dependent manner. Red line represents a two-fold cutoff for positive hits. RLU, relative luciferase units, normalized to β -gal control. Data are shown as Mean \pm SD. (B) Relative ligand binding affinity of 12 compounds to ER α or ER β .

Figure 2: BRET assays in HEK293 cells show dimer selectivity of different flavonoid subclasses. A-C) Fold change of BRET ratios when cells were treated with indicated compounds. A) ER α/α , B) ER β/β , C) ER α/β . 10nM E2 was used as a positive control. Each compound represents an individual experiment, those that induced dimer interaction at a threshold of *p* value of <0.05 were considered statistically significant. Fold change is relative to the negative control DMSO. Data are shown as Mean \pm SD of three biological replicates. * Indicates statistically significant. D) Western blot analysis of Flag-tagged ER β in MCF7-Flag-ER β cells. E, F) Relative ATP6V0E1 and BAG1 mRNAs levels in MCF7-Flag-ER β cells treated with indicated compounds. G) Compound 29-induced recruitment of ER β to the BAG1 and ATP6V0E1 promoters in MCF7-Flag-ER β cells shown by ChIP assays. H) The enrichment of ER α on the BAG1 and ATP6V0E1 promoters in MCF7-Flag-ER β cells after compound 29 treatment shown by ChIP assays.

Figure 3: Generation and selection of a pharmacophore hypothesis model of ER α/β heterodimer inducing ligands. A ligand based pharmacophore hypothesis was generated

MOL #108696

using GALAHAD. A) Structures and bioactivity values of the training set chemicals used to generate ligand based pharmacophore. The structures of the six lead compounds (cosmosiin, two isoflavones, four flavanones and a flavone) identified from the cell based assays. B) Plot of the different criteria used to select the best model. Plot of the Energy, Sterics, Mol_QRY and H_Bond values for GALAHAD models with selected four ligands that contribute to the consensus feature. A) Sterics vs Energy B) Pharmacophore similarity vs Energy C) Pharmacophore similarity vs Sterics. The blue circle represents the ideal best scoring for each condition. The red diamond represents Model 6. C) Selected Pharmacophore hypothesis GALAHAD Model. GALAHAD assumes pharmacophore/shape and alignments from sets of ligand molecules, to generate a pharmacophore hypothesis that can be used for a 3D search query. GALAHAD models were derived by using the ligands in the training set. It contains 7 features identified by GALAHAD represented by blue, green and purple spheres. The three hydrophobes are centered in the benzopyran and phenyl rings. The three acceptor atoms are in green and a donor atom is in purple.

Figure 4: 3D search query of two commercially available databases, the Chembridge and the Maybridge databases, which together have over a million chemicals. Resulted in a refined hit list of 167 compounds. A) Represents a schematic of the 3D Virtual Screening of the ChemBridge and Maybridge databases. B) Dose-response data of BRET assays in HEK293 cells, illustrating dimerization profile of selected hits. Data is shown as Mean \pm SD of three biological replicates. Data are normalized to DMSO control. * Indicates compounds that significantly induced dimerization as determined by two-way ANOVA. C)

MOL #108696

Measurement of compound binding to ER α and ER β using In Vitro Fluorescence Polarization Competition Binding Assays.

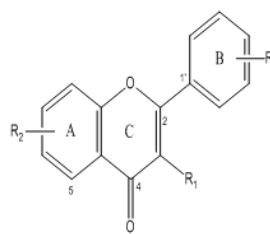
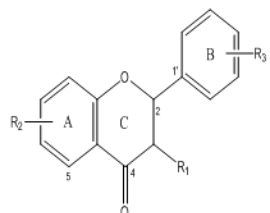
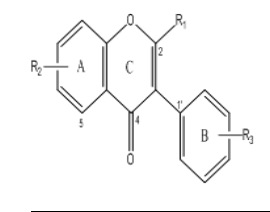
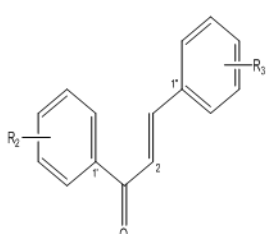
Figure 5: Mutant ER α and ER β LBDs reveal ER α as the dominant heterodimeric partner in the presence of selective ER α / β heterodimer compounds. A) heterodimerization of the wild type ER α and ER β , B) mutation in the ER α ligand binding domain ablates heterodimerization with ER β , C) heterodimerization of mutant ER α with mutant ER β , D) no dimerization is observed between mutant ER α and wild type ER β . Data are shown as Mean \pm SD. * Indicates statistically significant <0.05 .

MOL #108696

Table 1: Core structures and names of flavonoid compounds used in this study.

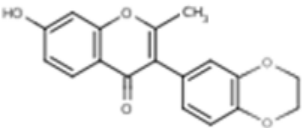
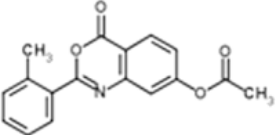
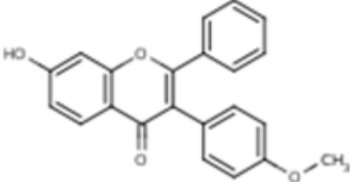
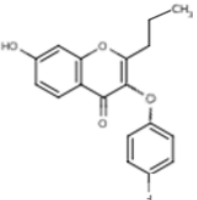
Thirty-one flavonoid compounds from 4 different subclasses: flavones, flavanones, isoflavones and chalcones were previously tested to

MOL #108696

	No.	Nomenclature	R1	R2	R3	M.W.
 <p>Flavone</p>	1	Flavone	H	H	H	222.24
	2	5-Hydroxyflavone	H	5-OH	H	238.24
	3	2',3'-Dihydroxyflavone	H	H	2',3'-Di-OH	254.24
	4	2',3'-Dimethoxyflavone	H	H	2',3'-Di-OCH3	282.29
	5	4'-Hydroxy-3'-methoxyflavone	H	H	4'-OH-3'-OCH3	268.26
	6	3',5-Dihydroxyflavone	H	5-OH	3'-OH	254.24
	7	4'-Hydroxy-5-methoxyflavone	H	5-OCH3	3'-OH	268.26
	8	5-Methoxyflavone	H	5-OCH3	H	252.26
	9	4',5,7-Trimethoxyflavone	H	5,7-Di-OCH	4'-OCH3	312.32
	10	8-Carboxyl-3-methylflavone	CH3	8-COOH	H	280.27
	11	5,6-Benzoflavone	H	5,6-Benzo	H	272.3
	12	7,8-Benzoflavone	H	7,8-Benzo	H	272.3
	13	2'-Methoxy- α -naphthoflavone	H	7,8-Benzo	2'-OCH3	302.32
	14	3,3',4',5,7-Pentahydroxyflavone-8-O-glucoside (Gossypin)	OH	5,7-Di-OH-8-O-Glucoside	3',4'-Di-OH	480.38
	15	3,4',5,7-Tetrahydroxyflavone (Kaempferol)	OH	5,7-Di-OH	4'-OH	286.24
	16	5,7-Dihydroxyflavone (Luteolin)	H	5,7-Di-OH	H	286.24
	17	3,7-Dihydroxyflavone	H	3,7-Di-OH	H	254.24
 <p>Flavanone</p>	18	4',5,7'-Trihydroxyflavanone (Naringenin)	H	5,7-Di-OH	4'-OH	272.25
	19	5-Methoxyflavanone	H	5-OCH3	H	254.28
	20	5,7-Dimethoxyflavanone	H	5,7-OCH3	H	284.31
	21	3',4',5',7-Tetramethoxyflavanone	H	7-OCH3	3',4',5'-Tri-OCH3	344.36
	22	Flavanone	H	H	H	224.08
	23	7-Hydroxyflavanone	H	7-OH	H	240.25
	24	4-Hydroxyflavanone	H	4-OH	H	254.24
	 <p>Isoflavone</p>	25	7-Methoxyisoflavone	H	7-OCH3	H
26		4',6,7-Trihydroxyisoflavone (Demethyltaxasin)	H	6,7-OH	4'-OH	270.24
27		4'-Hydroxy-6-methoxyisoflavone-7-D-guloside (Glycitin)	H	6-OCH3-7-D-Glucoside	4'-OH	446.4
28		5,7,4'-Trihydroxyisoflavone (Genistein)	H	5,7-Di-OH	4'-OH	270.24
29		5,7-Dihydroxy-4'-methoxyisoflavone (Biochanin A)	H	5,7-Di-OH	4'-OCH3	284.26
30		5,7,3',4'-Tetramethoxyisoflavone (Orobol)	H	5,7-Di-OCH3	3',4'-Di-OCH3	342.34
 <p>Chalcone</p>	31	2',3'',5''-Trimethoxychalcone	-	2'-OCH3	3'',5''-Di-OCH3	298.33
	32	3',3'',5''-Trimethoxychalcone	-	3'-OCH3	3'',5''-Di-OCH3	298.33
	33	3',3'',5'-Trimethoxychalcone	-	3'5'-Di-OCH3	3''-OCH3	298.33
	34	2'',3',5'-Trimethoxychalcone	-	3'5'-Di-OCH3	2'OCH3	314.33
	35	3',3'',5',5''-Tetramethoxychalcone	-	3'5'-Di-OCH3	3'',5''-Di-OCH3	328.36
	36	2'-Hydroxy-4',4''-dimethoxychalcone	-	2'-OH-4'-OCH3	4''-OCH3	284.31
	37	2'-Hydroxy-3'',4',5''-trimethoxychalcone	-	2'-OH-4'-OCH3	3'',5''-Di-OCH3	298.33

MOL #108696

Table 2: The structural arrangement of 5 ER α / β heterodimer selective compounds.

Compound No.	Structure	BRET
4		α/β at 10 μ M
6		α/β at 1 μ M
9		α/β at 1 μ M
10		α/β at 1 μ M

MOL #108696

Figure 1

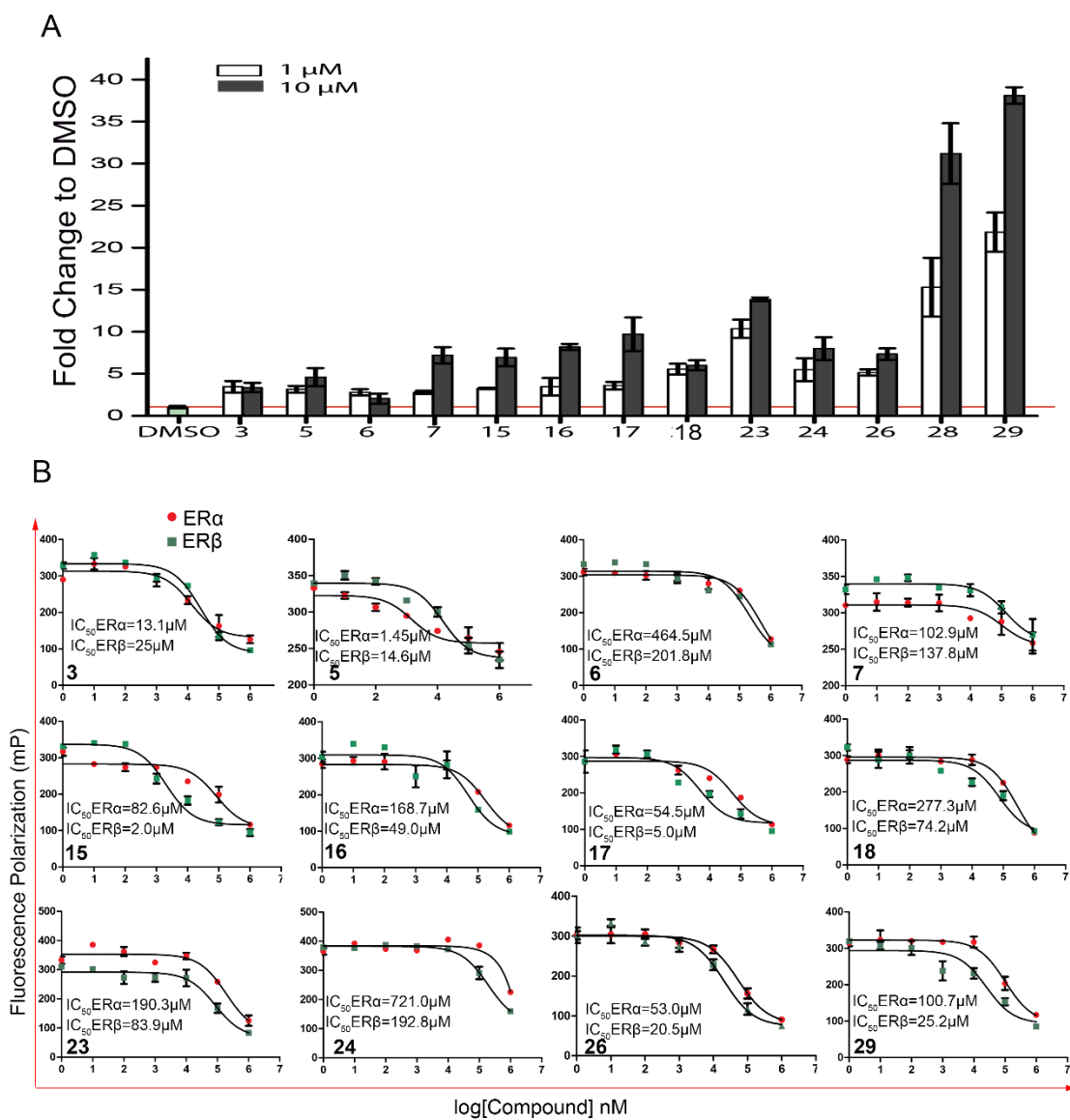
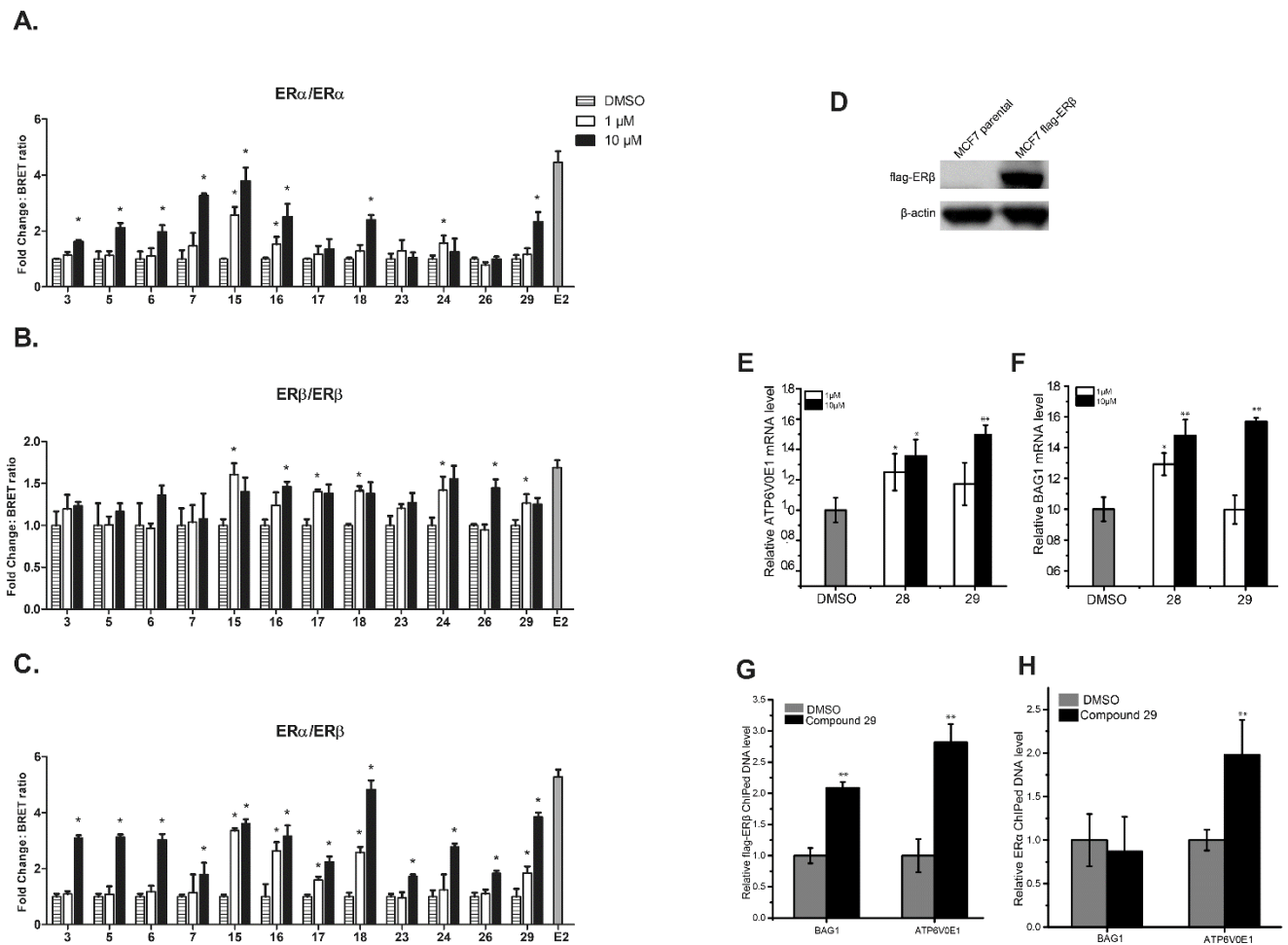
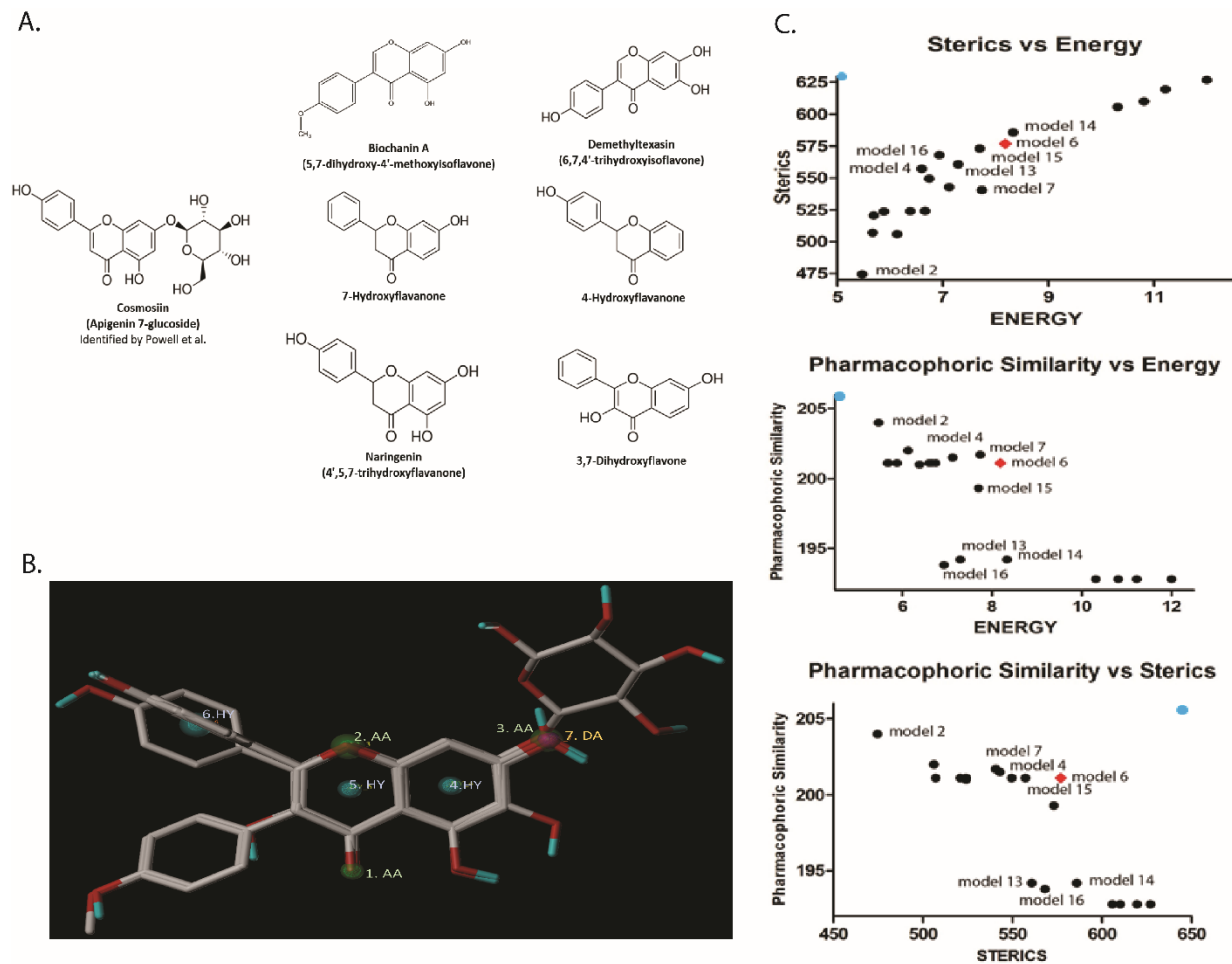


Figure 2

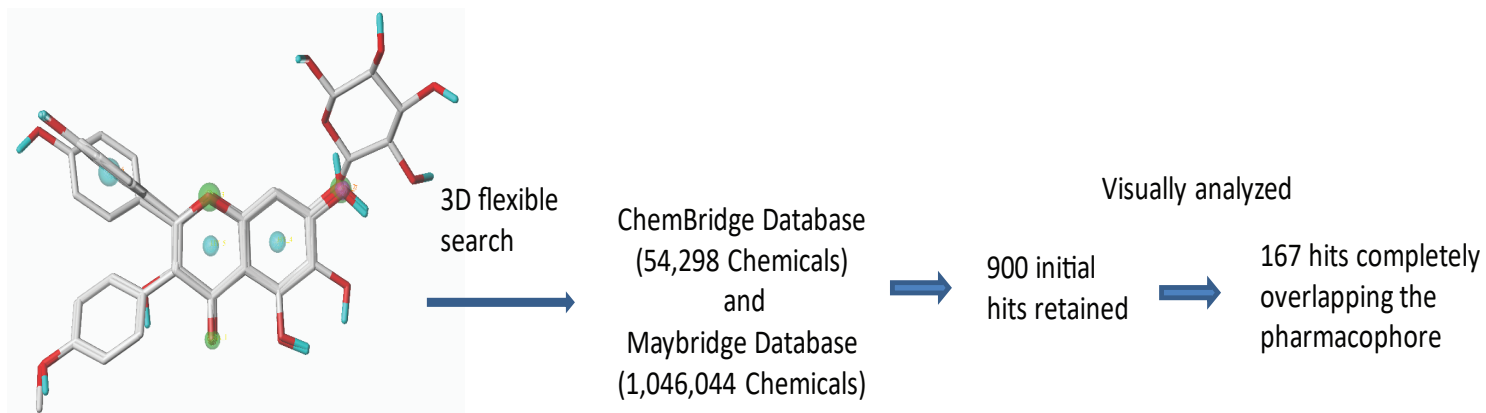


MOL #108696

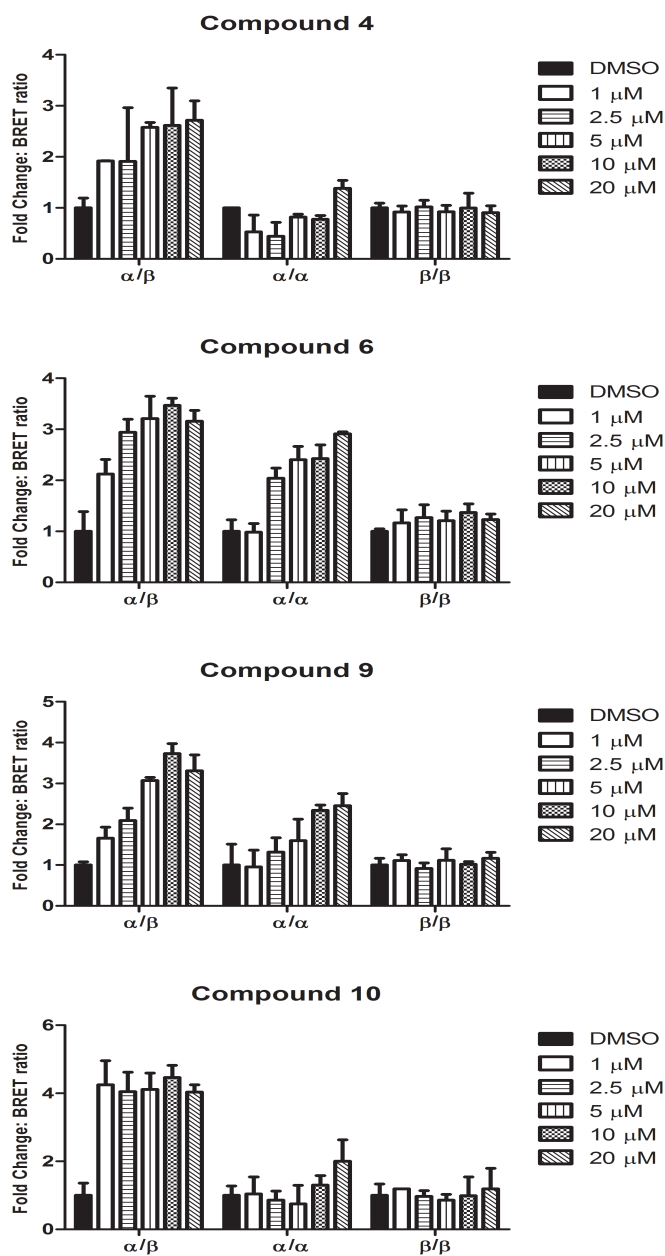
Figure 3



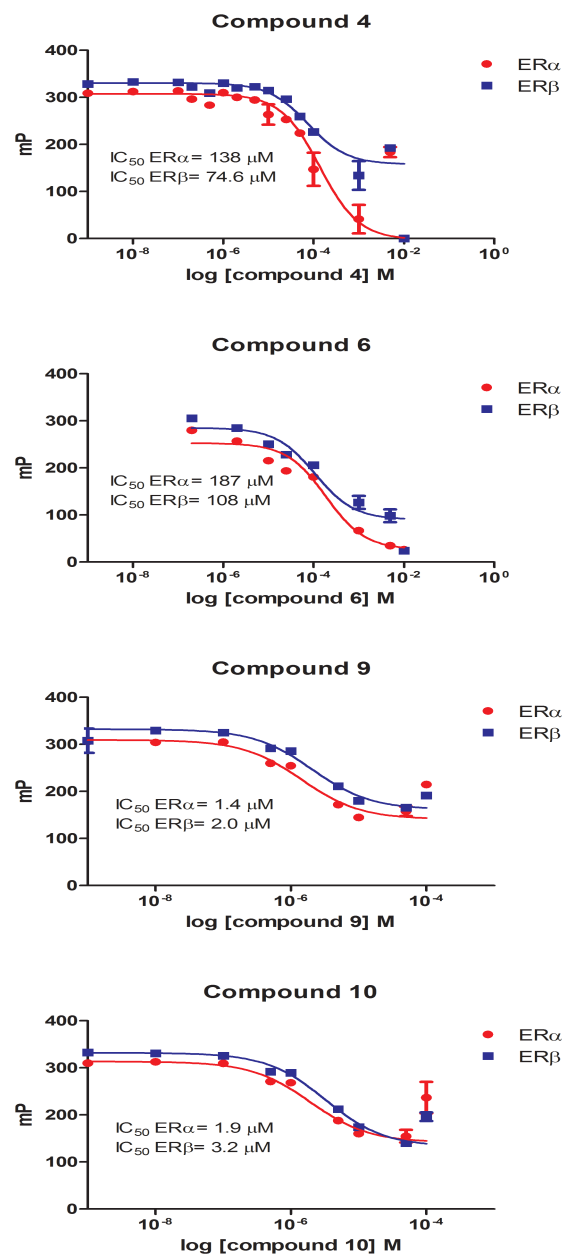
A.



B.



C.



MOL #108696

Figure 5

