PROTEINS:
Structure, Function, and Bioinformatics

Exploring the Molecular Basis of Action of ring D Aromatic Steroidal Antiestrogens

Journal:	*PROTEINS: Structure, Function, and Bioinformatics*
Manuscript ID:	Prot-00021-2015.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Alvarez, Lautaro; University of Buenos Aires, Organic Chemistry Veleiro, Adriana; University of Buenos Aires, Organic Chemistry Burton, Gerardo; University of Buenos Aires, Organic Chemistry
Key Words:	estrogen receptor, salpichrolide, withanolide, antiestrogenic, molecular dynamics

Exploring the Molecular Basis of Action of ring D Aromatic Steroidal Antiestrogens

Lautaro D. Alvarez, Adriana S. Veleiro and Gerardo Burton*

Departamento de Química Orgánica and UMYMFOR (CONICET-UBA), Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires

Keywords: Estrogen Receptor, Salpichrolide, Withanolide, antiestrogenic, Molecular Dynamics.

Running title: D-aromatic Antiestrogens

Correspondence to: Gerardo Burton, Departamento de Química Orgánica, Facultad de Ciencias Exactas y Naturales, Pabellón 2, Ciudad Universitaria, C1428EGA Ciudad de Buenos Aires, Argentina. Phone/Fax: 54 11 4576-3385. Email: burton@qo.fcen.uba.ar

ABSTRACT. Salpichrolides are natural plant steroids that contain an unusual six-membered aromatic ring D. We recently reported that some of these compounds, and certain analogues with a simplified side chain, exhibited antagonist effects toward the human estrogen receptor (ER), a nuclear receptor whose endogenous ligand has an aromatic A ring (estradiol). Drugs acting through the inhibition or modulation of ERs, are frequently used as a hormonal therapy for ER(+) breast cancer. Previous results suggested that the aromatic D ring was a key structural motif for the observed activity, thus this modified steroid nucleus may provide a new scaffold for the design of novel antiestrogens. Using Molecular Dynamics simulation we have modelled the binding mode of the natural salpichrolide A and a synthetic analogue with an aromatic D ring within the ERα. These results taken together with the calculated energetic contributions associated to the different ligand binding modes are consistent with a preferred inverted orientation of the steroids in the ligand binding pocket with the aromatic ring D occupying a position similar to that observed for the A ring of estradiol. Major changes in both dynamical behavior and global positioning of H11, caused by the loss of the ligand-His524 interaction might explain, at least in part, the molecular basis of the antagonism exhibited by these compounds. Using steered molecular dynamics we also found a putative unbinding pathway for the steroidal ligands through a cavity formed by residues in H3, H7 and H11, that requires only minor changes in the overall receptor conformation.

INTRODUCTION

The estrogen receptors (ERs) are members of the nuclear receptor superfamily (NR), which are soluble intracellular proteins that act as ligand-regulated transcription factors controlling specific gene expression in most mammalian cells.[1,2] The human NR superfamily includes 48 proteins that are essential in embryonic development, maintenance of differentiated cellular phenotypes, metabolism and apoptosis. The ERs together with the progesterone (PR), mineralocorticoid (MR), androgen (AR) and glucocorticoid (GR) receptors, form the steroid receptors (SR) family and represent one of the most important drug targets for the pharmaceutical industry.[3] In particular, drugs acting through the inhibition of ERs, antiestrogens, are frequently used as a hormonal therapy for patients who exhibit ER positive breast cancer. However, long-term treatment with these compounds is prone to severe adverse consequences due to their carcinogenic and genotoxic effects.[4]

In the Ligand Binding Domain (LBD) of SRs, the arrangement of α-helices creates a residue free cavity in its bottom half, where the steroidal ligand is bound.[5] The ligand-receptor interaction is mediated by non-specific/non-polar and specific/polar interactions. It is accepted that while the former interactions contribute with most of the binding energy, the latter are involved in the recognition of ligands. Due the aromatic character of the A ring of estradiol (**1**), the natural ER ligand (see Fig. 1 for structures), a major difference exists in the ligand binding mode for the ERs compared to other SRs. Thus, while in most SRs a ketone group in ring A of the steroid is oriented towards a Gln/Arg pair of the receptor, the aromatic A ring of estradiol exposes a phenolic hydroxyl to the Glu/Arg pair of the ERs. According to several ERα/estradiol crystal structures (e.g. pdb:1qku, Fig. 2a and 2b), Glu353 and Arg394 form part of a polar pocket that accommodates the planar A-

ring moiety.[6,7,8] These structures also show that at the other end of the mostly hydrophobic ligand molecule, the C17-hydroxyl contacts His524, establishing a hydrogen bond which is considered essential to maintain an active receptor conformation.[9,10] The presence of estradiol, or other agonists within the ER ligand binding pocket (LBP), induces the docking of helix 12 (H12) against helices 3, 4 and 11 in a highly conserved conformation, thus allowing coactivator recruitment and consequent transcriptional activation of target genes.

Currently, a large amount of structurally diverse synthetic estrogens have been reported that can be classified as pure agonists, pure antagonists or Selective Estrogen Receptor Modulators (SERMs).[11] In several pure antagonists and SERMs, such as tamoxifen or fulvestrant (ICI 182.780), a bulky side chain blocks H12 from assuming an agonist position, thereby preventing the formation of the coactivator binding pocket. Due to this mode of action, these type of compounds are considered as active antagonists. However, there are also ER ligands that lack these bulky groups and still have the ability to inhibit agonist action (passive antagonists).[12,13] This suggests that the interaction between ligands and ERs does not just switch on or off the receptor through helix H12, but that a more complex mode of action needs to be considered to fully understand the biochemistry of ER action.

The finding of an ideal antiestrogen without secondary effects is a pending issue in the treatment of ER related diseases and intense efforts are being made to identify new drugs with improved activity profiles. In recent years, there has been a growing interest in ER ligands of plant origin or phytoestrogens.[14,15,16] Mainly isolated from Leguminosae family, phytoestrogens are non-steroidal compounds[17] and until recently, no phytosteroids acting through the ERs had been described. Withanolides are a group of C-28 steroidal lactones and lactols isolated from several genera of the Solanaceae family.[18,19] A small group of

these compounds, termed salpichrolides,[20,21] were isolated from *Salpichroa origanifolia* and exhibited antiestrogenic effects.[22] Salpichrolides such as Salpichrolide A (Fig. 1), have a six-membered aromatic D ring, an unusual structural modification of the steroid nucleus. To further investigate this new type of ER ligands, we have synthesized salpichrolide analogues with a simplified side chain, and tested the antiestrogenic activity in the ER+ human breast cancer cells MCF-7.[23] It is noteworthy that a very simple salpichrolide analogue such as compound **3** (Fig. 1a), that has neither the side chain nor any of the functionalities present in rings A and B of the salpichrolides, retains the antiestrogenic activity. This result suggested that the aromatic D ring is a key structural motif for the observed activity. Combining docking and Molecular Dynamics simulation (MD) methods we also performed a preliminary analysis of the ligand binding mode of compound **3** in the ER LBP, finding that an inverse orientation might be stable.[23] Since this new scaffold could represent an opportunity to design novel antiestrogens with an improved therapeutic profile, we have now investigated in detail the binding mode of this type of compounds and explored the molecular basis of action of compounds **2** and **3** through MD simulations. Our results show that a favored inverted ligand binding mode produces important conformational changes on the receptor structure that would explain the passive antagonism of the D-aromatic analogues.

METHODS

Initial structures of ERα complexes.

The starting coordinates of the ER ligand binding domain were taken from the crystal structure of the ERα/estradiol complex (pdb:1qku, chain A). To build the ERα/estradiol system the co-crystallized estradiol molecule was conserved. All ERα/**3** systems were

constructed from our previous docking study,[23] using ligand coordinates from the best solutions found in clusters A, B and C to obtain the corresponding ERα/**3**-A, ERα/**3**-B and ERα/**3**-C systems. The ERα/**2** system was constructed from the 20 ns snapshot of the ERα/**3**-A system and the HF/631G** optimized structure of **2**, which was introduced by overlapping skeleton carbon atoms with the corresponding atoms in **3**. In all systems hydrogen atoms were added with the Tleap module considering the ε-tautomer of His524. To build the corresponding force field parameters of the ligand, RESP (restraint electrostatic potential) atomic partial charges were computed using the HF method with the 6-31G** basis set in the quantum chemistry program Gaussian 03[24] for the corresponding HF-optimized structures.

Molecular Dynamics

Molecular dynamics (MD) were performed with the AMBER 12 software package.[25] The ligand parameters were assigned according to the general AMBER force field (GAFF) and the corresponding RESP charges using the Antechamber. The Amber99 force field parameters were used for all receptor residues.[26] The complexes were immersed in an octahedral box of TIP3P water molecules using the Tleap module, giving final systems of around 32000 atoms. Two chloride ions were added in order to balance the charges in the complexes. The systems were initially optimized and then gradually heated to a final temperature of 300 K. Starting from these equilibrated structures, MD production runs of 100 ns were performed. All simulations were performed at 1 atm and 300 K, maintained with the Berendsen barostat and thermostat respectively, using periodic boundary conditions and the particle mesh Ewald method (grid spacing of 1 Å) for treating long-range electrostatic interactions with a uniform neutralizing plasma. The SHAKE algorithm

was used to keep bonds involving H atoms at their equilibrium length, allowing the use of a 2 fs time step for the integration of Newton's equations.

Steered Molecular Dynamics (SMD) Simulation was carried out to investigate the ligand unbinding pathway from the 100 ns snapshot of the ERα/**3**-C system. Taking these coordinates as the initial structure, a pulling force was applied to increase to 30 Å the distance between the C8 atom of the steroid (Fig. 1) and the CA atom of Glu353. A force constant of 4 Kcal/mol $Å^2$ was used and 5 ns of Amber-SMD simulation was performed at a constant temperature of 300 K.

Analysis of results

The root mean squared deviation (RMSD) of ligand atoms, the root mean square fluctuations (RMSF) of CA receptor residues, the time evolution of the distances among selected atoms and the time evolution of torsion angles between selected atoms were monitored with the Ptraj module. Trajectories were visualized and representative snapshots were obtained using VMD.[27]

The MM/PBSA.py tool implemented in AMBER was used to compute the electrostatic (ele) and Van der Waals (vdw) contributions to the total energy of the molecular mechanics (MM) force field in the gas phase. The desolvation term of the ligand was not taken into account as it is located in the interior of the LBP, hence that term should largely cancel in the comparison of the different poses of compound **3**. Moreover, due to the high similarity among the simulated systems, the entropic term has not been calculated.

The MM/QM-COSMO calculations were performed following the method of Anisimov and Cavasotto[28] with the MOPAC program.[29] Local energy minimizations limited to 100 cycles were carried out on each ligand-receptor structure using the PM6 Hamiltonian[30] In

this case, solvation terms were evaluated by the COSMO continuum solvent model using the standard atomic radii and a dielectric constant of 78.4. The surface accessible solvent area was built using a 1.3 Å solvent probe radius. For the non-polar contribution a standard surface tension of 0.0037 Kcal/(mol Å^2) was used. Entropic contributions have not been considered. Both MM and MM/QM-COSMO calculations were performed over 8000 snapshots of the last 80 ns of the trajectory, previous deletion of water molecules.

The CAVER 3.0 program[31] was used to detect tunnels in the ERα/**3**-C system; 500 snapshots were extracted from the MD trajectory and then ligand and water molecules were removed. The coordinates of the C8 of the steroid in the first snapshot, were used as the initial starting point to the tunnel search. A probe radius of 0.9 Å and a clustering threshold of 3.5 were used. Tunnels were visualized with VMD.[27]

RESULTS AND DISCUSSION

Ligand binding mode of compound 3

A central aspect of the molecular basis of antiestrogenic activity of D-aromatic steroid analogues resides in the orientation that these molecules may assume within the ERs. To obtain a preliminary insight on the binding mode of compound **3**, we have previously docked compound **3** into the crystal structure of ERα/estradiol complex (pdb:1qku)[23] and found that it can acquire three globally different orientations inside the ERα LBP. Two of them (poses A and B) correspond to an inverse binding mode in which ring D points towards the polar pocket flanked by Glu353 and Arg394 (Fig. 3a and 3b, respectively), while one (pose C) corresponds to the normal mode (Fig. 3c). The difference between both inverse modes resides in that pose B exhibits a second 180° inversion along the long axis of the molecule. The statistical analysis of the docking solutions revealed that pose A was

both the more frequent and the more energetically favorable,[23] suggesting that this could be a relevant binding mode of D-aromatic analogues. Based on these findings and to further investigate the molecular basis of the interaction between **3** and the ERα LBP, we constructed three ERα/**3** complexes using the best solution of each docking pose and the ERα/estradiol crystal structure (ERα/**3**-A, ERα/**3**-B and ERα/**3**-C systems). From these initial ligand-receptor coordinates, 100 ns of Amber-MD simulations were performed for each complex. The ERα/estradiol MD simulation was also performed as a control trajectory. All the simulated complexes were constructed by removing the crystallized water molecules and immersing the protein and ligand atoms in a box of explicit TIP3P water molecules.

ERα LBD/estradiol complex. As expected, the time-dependent residue fluctuation (root-mean square deviations, RMSD) measured over the heavy atoms of the steroid from the initial structure (Fig. 4a) reveals that the position of the ligand molecule was conserved during all the simulation, maintaining very stable interactions with the LBP residues. Fig. 4b shows a representative snapshot of the estradiol binding mode observed during the MD simulation. Consistently, the polar interactions between the C3-hydroxyl and both Arg394 and Glu353 and between the C17-hydroxyl and His524 remained stable during the time-scale of the simulation. Moreover, a water molecule occupied a position similar to that found in several ER/estradiol crystal structures, taking part in a polar network together with Arg394 and Glu353. In this way, our 100 ns MD of the ERα/estradiol complex is consistent with the crystal structure data and with previous simulations performed with this system.[32,33]

*ERα LBD/**3** complexes.* We were first interested in analyzing the dynamical behavior of the ERα/**3** complex with the steroid oriented as the estradiol molecule (ERα/**3**-C system, "normal" binding mode). Notably, the visual inspection of this trajectory reveals that the final position of the steroid molecule differs considerably from the initial one (Fig. 4e). The RMSD curve shows that the position of the ligand atoms changes abruptly during the first 20 ns of the simulation reaching a stable value of 6.5 Å (Fig. 4a). In this new position the steroid does not interact with any of the polar LBP residues. The average distance between the center of mass of the ligand and the CZ atom of Arg394 and the CD atom of Glu353 are 15.24 and 13.92 Å, respectively; these values are much larger than those observed in the estradiol system (Table I). Interestingly, while the steroid is moving towards the His524, there is a major change in the side chain conformation of this residue. Thus the interchange of the aromaticity between A and D rings causes a rearrangement of LBP residues that moves the molecule of compound **3** away from the LBP center, indicating that this orientation would not be adequate for binding **3**.

In contrast with the above results, the MD simulations on the steroid in both the inverted binding orientation (ERα/**3**-A system) and the doubly inverted binding orientation (ERα/**3**-B system) exhibited small and constant RMSD values (Fig. 4a), with averages similar to those observed in the ERα/estradiol system (Table I). In both systems the steroid remained practically in its original position (Table I), with the C20-carbonyl interacting with the Arg394/Glu353 pair and a water molecule, thus establishing a polar network that resembled that observed in the ERα/estradiol complex (Fig. 4c and 4d). At the other end of the ligand molecule, no interactions were observed between the C1-carbonyl and the His524. Thus,

the MD simulations showed that both the inverted and the doubly inverted orientations have the ability to achieve a stable binding mode.

Energy of ligand-receptor interactions. To determine which of the above inverted binding modes is more favorable, we calculated the energetic contribution associated to the ligand binding mode in each trajectory. We used two approaches to estimate the energy of the ligand-receptor interaction. First, the MM/PBSA method was used to compute the energetic contributions from the electrostatic energy (ele) and Van der Waals interactions (vdw) arising from bond, angle and dihedral terms in the force field, the sum of which gave the total gas phase binding energy (MM). Interestingly, the vdw contributions in ERα/**3** systems were larger than those calculated for the ERα/estradiol system, probably due to the higher number of carbons atoms in the **3** molecule. However, as could be inferred from the observed polar interactions, the ele contribution in the case of estradiol was significantly more favorable than in the case of **3**. The results also revealed that the interaction is 3.2 Kcal/mol more stable in the inverted binding mode (ERα/**3**-A system) compared to the doubly inverted binding mode (ERα/**3**-B system), mainly as consequence of a more efficient hydrophobic contact between the steroid and the LBP residues (Table I). In the case of the "normal" binding orientation of **3** (ERα/**3**-C system), the ele contribution in the final position is smaller than in the other systems, resulting in a less favorable MM energy.

In order to obtain a more precise estimation of the ligand-receptor interaction energy, we also used the MM/QM-COSMO method to estimate the binding free energy of **3** in both inverted binding orientations (A and B poses). In this calculation developed by Anisimov and Cavasotto,[28] MD trajectories are re-evaluated using a semiempirical PM6 Hamiltonian and a continuum solvent model, while translational and rotational entropy contributions are calculated through their corresponding configurational integrals and the internal

(vibrational) entropy is evaluated through normal mode analysis. As was observed with the MM/PBSA method, the MM/QM-COSMO results show that the binding free energy is significantly more favorable when **3** is in the inverted binding orientation (pose A) compared to the doubly inverted orientation (pose B) (Table I).

Taken the previous docking study and the above MD results altogether, we conclude that the more favorable binding mode for compound **3** corresponds to the inverted orientation (pose A) compared to the natural ligand estradiol.

Ligand binding mode of compound 2

Next, we were interested in investigating the ligand binding mode of the more complex salpichrolide A molecule (**2**, Fig. 1). In contrast with the simplified analogue, this natural antiestrogen has a very bulky side chain attached to the aromatic D ring and therefore the main questions reside in how the receptor can accommodate this moiety and if there are any specific polar interactions involved in its recognition. Since we had already found the preferred binding orientation of compound **3**, we used this model as starting point to investigate the binding mode of **2**. The HF/6-31G** optimized structure of **2** was introduced in the 20 ns snapshot of the ERα/**3**-A system superimposing the carbon atoms of the steroid nucleus of both compounds. In this initial structure, the side chain of **2** was located in a cavity formed by residues of H3, H4, the H1-H3 loop and the β-sheet, with the Glu353/Arg394 pair at one side. In this free cavity of the binding pocket the steroid side chain may be tightly accommodated, with only minor sterical superposition between receptor and ligand atoms that disappeared during the minimization step of the equilibration protocol. Visual inspection of the MD production, revealed a pronounced conformational change of Leu349 located at H3, thus allowing for the ligand side chain. The salpichrolide

molecule achieved a globally stable binding mode with minor changes from its initial position (Fig. 5a), as indicated by the small RMSD values (Fig. 5b). Furthermore, an interaction was observed between the salpichrolide side chain and the receptor, the 26-hydroxyl formed a stable and strong hydrogen bond with the backbone oxygen atom of Leu346 also located at H3. This interaction was maintained during all the time scale of the simulation (Fig. 5c). The RMSD values measured for backbone atoms of the protein indicated that the structure of the receptor did not experience any major changes either (Fig. 5b). Thus, starting from the inverse orientation, we obtained an ERα/**2** complex in which not only the receptor was able to accommodate the voluminous side chain of **2** by modifying the conformation of only one residue, but a specific polar interaction participated in its recognition. The previous report that other salpichrolides that differ from **2** in the functional groups on rings A or D exhibited similar antiestrogenic activity,[23] suggests that the tight fit of the withanolide side chain in the secondary cavity described above and the observed specific interaction would play a key role in their affinity for the receptor. The changes in the functionalities in rings A or D would have a minor effect.

Molecular basis of the antagonism of compound 3

To investigate the molecular basis involved in the passive antagonism of **3**, we compared the dynamical behavior and overall structures of ERα/estradiol (agonist system) and ERα/**3**-A (antagonist system) complexes. First, we calculated the overall backbone fluctuation of the protein over the last 60 ns of MD simulation (RMSF) that provides a time-average representation of per-residue fluctuation. Fig. 6 shows the difference between RMSF values of ERα/estradiol and ERα/**3**-A complexes, revealing two main regions where the fluctuation results are considerably different between both systems. On one side, the

fluctuation of the H8-H9 loop is smaller in the presence of **3** than in presence of estradiol. Interestingly, this is a nine residues long loop that, according to crystal structures,[6,7,8] forms part of the dimer interface. The change in the dynamical behavior of the H8-H9 loop may thus affect the ability of the complex to dimerize.

On the other hand, the H11-H12 loop is significantly more mobile when the D-aromatic analogue is bound. As mentioned in the introduction, the agonist action of estradiol has been associated to its ability to interact with the His524 (H11), since this interaction fixes a His524 conformation that can participate in an extensive hydrogen bonding network involving Glu419, Glu339 and Lys531. The ionic interaction between the side chains of Glu339 (H3) and Lys531 (H11) works as a "zipper", coupling the N-terminal end of H3 with the C-terminal end of H11 and stabilizing a H11 conformation that positions H12 in the agonist conformation.[32] Fig. 7a shows a representative snapshot of the ERα/estradiol complex in which the formation of this "zipper" is depicted. The analysis of key distances between the residue pairs involved, reveals that the "zipper" is formed during most of the time scale of the ERα/estradiol simulation (Fig. 7b). Instead, in the ERα/**3**-A system we observed that the "zipper" stays until *ca* 25 ns and then disappears (Fig. 7c). This results in a global conformational change of the receptor that leads to the separation of the H3 and H11 helices (compare Fig. 8a and 8b), a phenomenon that can be observed analyzing the time evolution of the distance between the CA atoms of Glu339 and Lys531 (Fig. 8c). Thus, we conclude that the inability of **3** to interact with the His524 is translated into major changes in both dynamical behavior and global position of the H11, destabilizing the agonist conformation of the ER and giving rise to the passive antagonism of this D-aromatic analogue.

Unbinding pathway of compound 3 in pose C

As was described above, when we evaluated the binding mode of **3** in the normal orientation (ERα/**3**-C system), a large displacement of the steroid towards the H11 was observed achieving a final localization under His524. To the best of our knowledge, this is the first time that such pronounced ligand mobility is reported for an SR/ligand complex using classical MD. This strongly suggests that the presence of a non-aromatic ring near the Arg/Glu pair and an aromatic ring near the His524 is highly unfavorable. To evaluate the presence of tunnels connecting the LBP with bulk solvent, the ERα/**3**-C trajectory was reanalyzed with the CAVER program.[31,34] Only two relevant tunnel clusters were detected during the entire MD simulation (Fig. 9a). Tunnel I was shaped by residues at H3, H7 and H11 helices involving half of the LBP and a cavity formed by the terminal residues of these helices. Tunnel II had the opposite direction, contacting the bulk solvent through a cavity located between H3, H5, H8 and the H1-H3 loop. Analysis of the characteristics of these tunnels showed that Tunnel I was far more favored than Tunnel II. Thus, over 500 snapshots analyzed, Tunnel I was identified practically in all of them (98 %) while Tunnel II was less detected (67 %). Furthermore, the average value of the bottleneck radii was considerably larger for Tunnel I compared to Tunnel II (2.1 Å and 1.0 Å respectively).

On the basis of these findings, we envisaged that the pathway followed by the ligand in the first 20 ns of the ERα/**3**-C system might represent a putative unbinding pathway. The fact that the final localization was conserved during the next 80 ns of the MD simulation suggested that this position might be considered as a stable intermediate state. Since very large times could be required to visualize a spontaneous and complete unbinding using classical MD, several modified methods have been applied to understanding the binding/unbinding pathways in NRs.[35-39] Here, we used steered MD (SMD), which

provides a means of accelerating the unbinding process by applying external forces that lower the energy barriers, to drive the ligand along its unbinding path on a nanosecond time scale.

In order to induce the unbinding of ligand from the 100 ns snapshot of the ERα/**3**-C system, 5 ns of SMD were carried out with a pulling force applied so as to increase the distance between the C8 atom of the steroid and the CA atom of Glu353, driving the steroid along Tunnel I. Fig. 9b shows the temporal evolution of distance between pulled atoms and Fig. 9c reveals that continuous ligand RMSD evolution is obtained during the SMD. Remarkably, receptor RMSD values are always smaller than 2 Å (Fig 9c), indicating that the complete ligand unbinding (Fig 9d) can be achieved without major changes of the receptor backbone conformation. In this way, the SMD simulation shows that the ligand can be actually expelled through Tunnel I. Using random acceleration molecular dynamics (RAMD), Burendahl et al have found that estradiol exhibits a strong preference for a similar unbinding pathway.[36]

CONCLUSIONS

Several withanolide containing plants have been used in traditional medicine for centuries. A subgroup of withanolides having an aromatic ring D were found to inhibit the action of estradiol, a steroid hormone characterized by an aromatic ring A. This led to the synthesis of several simplified analogues that suggested that the aromatic D ring was key to the observed activity. Our MD simulation results for the natural salpichrolide A (**2**) and the simplified synthetic analogue **3**, show that those orientations in which the aromatic ring D of these compounds plays the role of the A ring of estradiol in the ER LBP, are more favorable than the "normal" steroid orientation. Thus, the localization of a planar ring in

this region of the LBP appears to be essential for adequate steroid recognition by the ER. When a non-aromatic ring is introduced there, the system becomes highly unstable and the receptor tends to expel the steroid molecule, following a pathway that is coincident with that found for the unbinding of other NR ligands. Unable to form a polar interaction with the His524, the binding of the D-aromatic analogue in the inverted mode disrupts the contact between H3 and H11 that is required to stabilize the H12 agonist conformation. Thus, the molecular basis of the passive antagonism of **3** would not originate in a direct disarrangement of H12, but through an indirect perturbation of H11 dynamics.

Finally, molecular modelling results showed how complex molecules such as **2** can bind to the ERα exploiting a secondary cavity generated by the rotation of Leu349. This observation may be used for the design of novel analogues with improved ER activity.

ACKNOWLEDGMENTS

We thank Agencia Nacional de Promoción Científica y Tecnológica (PICT2010-0623), CONICET (Argentina) Grant 11220110100702 and Universidad de Buenos Aires (Grant 20020130100367BA) for financial support.

REFERENCES

1. Gronemeyer H, Gustafsson JA, Laudet V. Principles for modulation of the nuclear receptor superfamily. Nat Rev Drug Discov 2004;3:950-964.
2. D'Auria MV, Sepe V, Zampella A. Natural ligands for nuclear receptors: biology and potential therapeutic applications. Curr Top Med Chem 2012;12:637-669.
3. Moore JT, Collins JL, Pearce KH. The nuclear receptor superfamily and drug discovery. ChemMedChem 2006;1:504-523.

4. Yang G, Nowsheen S, Aziz K, Georgakilas AG. Toxicity and adverse effects of Tamoxifen and other anti-estrogen drugs. Pharmacol Ther 2013;139:392-404.

5. Brelivet Y, Rochel N, Moras D. Structural analysis of nuclear receptors: from isolated domains to integral proteins. Mol Cell Endocrinol 2012;348:466-473.

6. Tanenbaum DM, Wang Y, Williams SP, Sigler PB. Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains. Proc Natl Acad Sci U S A 1998;95:5998-6003.

7. Gangloff M, Ruff M, Eiler S, Duclaud S, Wurtz JM, Moras D. Crystal structure of a mutant hERalpha ligand-binding domain reveals key structural features for the mechanism of partial agonism. J Biol Chem 2001;276:15059-15065.

8. Brzozowski AM, Pike AC, Dauter Z, Hubbard RE, Bonn T, Engström O, Ohman L, Greene GL, Gustafsson JA, Carlquist M. Molecular basis of agonism and antagonism in the oestrogen receptor. Nature 1997;389:753-758.

9. Ekena K, Weis KE, Katzenellenbogen JA, Katzenellenbogen BS. Different residues of the human estrogen receptor are involved in the recognition of structurally diverse estrogens and antiestrogens. J Biol Chem 1997;272:5069-5075.

10. Ekena K, Weis KE, Katzenellenbogen JA, Katzenellenbogen BS. Identification of amino acids in the hormone binding domain of the human estrogen receptor important in estrogen binding. J Biol Chem 1996;271:20053-20059.

11. Bolli A, Marino M. Current and future development of estrogen receptor ligands: applications in estrogen-related cancers. Recent Pat Endocr Metab Immune Drug Discov 2011;5:210-229.

12. Shiau AK, Barstad D, Loria PM, Cheng L, Kushner PJ, Agard DA, Greene GL. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. Cell 1998;95:927-937.

13. Shiau AK, Barstad D, Radek JT, Meyers MJ, Nettles KW, Katzenellenbogen BS, Katzenellenbogen JA, Agard DA, Greene GL. Structural characterization of a subtype-selective ligand reveals a novel mode of estrogen receptor antagonism. Nat Struct Biol 2002;9:359-364.

14. Michel T, Halabalaki M, Skaltsounis AL. New concepts, experimental approaches, and dereplication strategies for the discovery of novel phytoestrogens from natural sources. Planta Med 2013;79:514-532.

15. Castelo-Branco C, Soveral I. Phytoestrogens and bone health at different reproductive stages. Gynecol Endocrinol 2013;29:735-743.

16. Cassidy A. Dietary phytoestrogens and bone health. J Br Menopause Soc 2003; 9:17-21.

17. Patisaul HB, Jefferson W. The pros and cons of phytoestrogens. Front Neuroendocrinol 2010;31:400-419.

18. Misico RI, Nicotra VE, Oberti JC, Barboza G, Gil RR, Burton G. Withanolides and related steroids. Prog Chem Org Nat Prod 2011;94:127-229.

19. Veleiro AS, Oberti JC, Burton G. Chemistry and bioactivity of withanolides from southamerican Solanaceae. Studies in Natural Products Chemistry. Bioactive Natural Products 2005; 32:1019-1052.

20. Veleiro AS, Oberti JC, Burton G. A Ring-D Aromatic Withanolide from Salpichroa origanifolia. Phytochemistry 1992;31:935-937.

21. Tettamanzi MC, Veleiro AS, de la Fuente JR, Burton G. Withanolides from Salpichroa origanifolia. J Nat Prod 2001;64:783-786.

22. Machin RP, Veleiro AS, Nicotra VE, Oberti JC, M Padrón. Antiproliferative activity of withanolides against human breast cancer cell lines. J Nat Prod 2010; 73:966-968.

23. Sonego JM, Rivero EM, Gargiulo L, Lüthy I, Alvarez LD, Veleiro AS, Burton G. Synthesis and biological evaluation of salpichrolide analogs as antiestrogenic agents. Eur J Med Chem 2014;82:233-241.

24. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA, Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA. Gaussian, Version 03, Revision C. Wallingford: Gaussian Inc.; 2004.

25. Case DA, Cheatham III TA, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts B, Hayik S, Roitberg A, Seabra A, Swails J, Goetz AW, Kolossváry I, Wong, Paesani F, Vanicek J, Wolf RM, Liu J, Wu X, Brozell

SR, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Cui G, Roe DR, Mathews DH, Seetin MG, Salomon-Ferrer, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A, Kollman PA. AMBER 12, University of California, San Francisco. 2012.

26. Cheatham III TE, Cieplak P, Kollman PA. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. J Biomol Struct Dyn 1999;16:845-862.

27. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14:33-38.

28. Anisimov VM, Cavasotto CN. Quantum mechanical binding free energy calculation for phosphopeptide inhibitors of the Lck SH2 domain. J Comput Chem 2011; 32:2254-2263

29. MOPAC2009, JJPS, Stewart Computational Chemistry, Colorado Springs, CO, USA. 2008.

30. Stewart JJ. Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. J Mol Model 2007;13:1173-1213.

31. Chovancová E, Pavelka A, Beneš P, Strnad O, Brezovský J, Kozlíková B, Gora A, Šustr V, Klvaňa M, Medek P, Biedermannová L, Sochor J, Damborský J. CAVER 3.0: A Tool for the analysis of transport pathways in dynamic protein structures. PLoS Comp Biol 2012;8:e1002708

32. Celik L, Lund JD, Schiott B. Conformational dynamics of the estrogen receptor alpha: molecular dynamics simulations of the influence of binding site structure on protein dynamics. Biochemistry 2007;46:1743-1758.

33. Gao L, Tu Y, Ågren H, Eriksson LA. Characterization of agonist binding to His524 in the estrogen receptor alpha ligand binding domain. J Phys Chem B 2012;116:4823-4830.

34. Petrek M, Otyepka M, Banáš P, Kosinová P, Koča J, Damborský J. CAVER: A new tool to explore routes from protein clefts, pockets and cavities. BMC Bioinformatics 2006;7:316

35. Kosztin D, Izrailev S, Schulten K. Unbinding of retinoic acid from its receptor studied by steered molecular dynamics. Biophys J 1999;76:188-197.

36. Burendahl S, Danciulescu C, Nilsson L. Ligand unbinding from the estrogen receptor: a computational study of pathways and ligand specificity. Proteins 2009;77:842-856.

37. Perakyla M. Ligand unbinding pathways from the vitamin D receptor studied by molecular dynamics simulations. Eur Biophys J 2009;38:185-198.

38. Martinez L, Polikarpov I, Skaf MS. Only subtle protein conformational adaptations are required for ligand binding to thyroid hormone receptors: simulations using a novel multipoint steered molecular dynamics approach. J Phys Chem B 2008;112:10741-10751.

39. Capelli AM, Bruno A, Entrena Guadix A, Costantino G. Unbinding pathways from the glucocorticoid receptor shed light on the reduced sensitivity of glucocorticoid ligands to a naturally occurring, clinically relevant mutant receptor. J Med Chem 2013;56:7003-7014.

Figure legends

FIGURE 1. Structures of ER ligands.

FIGURE 2. a) General view of the crystal structure of the ERα LBD/estradiol complex (pdb:1qku). b) Ligand binding mode of estradiol (**1**) in the ERα LBD/estradiol crystal structure (pdb:1qku) showing the estradiol molecule and the polar amino acid side chains involved in ligand binding.

FIGURE 3. Docking solutions of compound **3** (light gray) into the ERα ligand binding pocket.[23] a) Pose A; b) Pose B; c) Pose C. The estradiol molecule is shown as reference in dark gray.

FIGURE 4. a) RMSD from the initial structures measured over ligand atoms of the simulated systems. b) Representative snapshot of the estradiol binding mode during the MD simulation. c-e) Representative snapshots of compound **3** binding mode during the MD simulation in pose A (c), pose B (d) and pose C (e).

FIGURE 5. a) Representative snapshot of the binding mode of **2** during the MD simulation of the ERα/**2** complex. b) RMSD of the initial structures measured over ligand atoms and over receptor backbone atoms. c) Time evolution of the distances between the hydrogen atom of the 26-hydroxyl group of **2** and the backbone oxygen atom of Leu346.

FIGURE 6. Difference between the RMSF values of the ERα/estradiol and the ERα/**3**-A systems ($\Delta RMSF = RMSF_{ER\alpha/estradiol} - RMSF_{ER\alpha/\mathbf{3}\text{-}A}$). The secondary structure of ER LBD is schematized along the *x*-axis.

FIGURE 7. Representative snapshot of ERα/estradiol complex showing the extensive hydrogen bonding network involving residues of H3, H7 and H11. b-c) Time evolution of the distances between the hydrogen atom of the C17-hydroxyl of estradiol and the ND1

nitrogen atom of His524 (A distance), the HE2 hydrogen atom of His524 and the oxygen of the backbone carbonyl of Glu419 (B distance), the hydrogen atom of Glu419 amino group and the CD oxygen atom of Glu339 (C distance); the CD oxygen atom of Glu339 and the NZ nitrogen atom of Lys531 (D distance) and the CD oxygen atom of Glu419 and the NZ nitrogen atom of Lys531 (E distance) in the ERα/estradiol (b) and ERα/**3**-A (c) systems.

FIGURE 8. a-b) 100 ns snapshots of the ERα/estradiol (a) and the ERα/**3**-A (b) complexes with the CA atoms of Lys531 and Glu339 shown as dark gray balls. c) Time evolution of the distances between CA atoms of Lys531 and Glu339 in the ERα/estradiol and the ERα/**3**-A complexes.

FIGURE 9. a) Initial snapshot of the ERα/**3**-C system showing tunnels detected by CAVER 3.0. Tunnel I and Tunnel II are represented by a sequence of dark gray and light gray balls, respectively. b) Time evolution of the distance between pulled atoms in the SMD simulation. c) RMSD of the initial structures measured over ligand atoms (gray) and over receptor backbone atoms (black); d) 1 ns, 3 ns and 5 ns snapshots of the SMD simulation. Pulled atoms are shown as dark gray balls.

Table I. Structural and thermodynamical information of ERα/ligand complexes

	estradiol	3-A	3-B	3-C
Average RMSD (Å)[a]				
Ligand	0.76	1.64	0.94	5.98
Protein	1.68	1.85	1.90	1.99
Average Distance (Å)				
CZ(Arg394) - CD(Glu353)	4.18	4.33	4.11	4.28
Ligand - CZ(Arg394)[b]	10.54	10.30	10.59	15.24
Ligand - CD(Glu353)[b]	9.22	9.99	9.54	13.92
Ligand - CG(His524)[b]	8.31	9.31	9.11	5.75
MM (kcal/mol)[c]				
vdw	-40.6	-48.2	-45.5	-44.2
ele	-16.2	-7.0	-6.5	-3.2
MM	-56.8	-55.2	-52.0	-47.4
MM/QM-COSMO (kcal/mol)[d]				
Total	nd	-23.9	-18.0	nd

[a]Average RMSD values measured over ligand or receptor backbone heavy atoms.
[b]Average distances between the center of mass of the ligand and the residue atom.
[c]Interaction energy contributions to the total energy of the molecular mechanics (MM) force field in the gas phase computed using the MM/PBSA method (vdw: Van der Waals; ele: electrostatic; MM: total gas phase binding energy).
[d]Binding free energies computed using the MM/QM-COSMO method (nd: not determined).

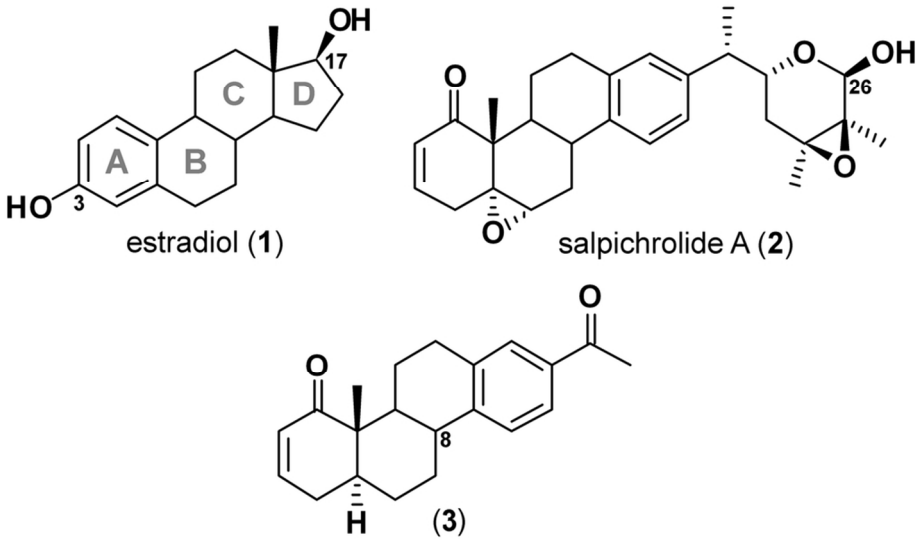

FIGURE 1. Structures of ER ligands
47x26mm (600 x 600 DPI)

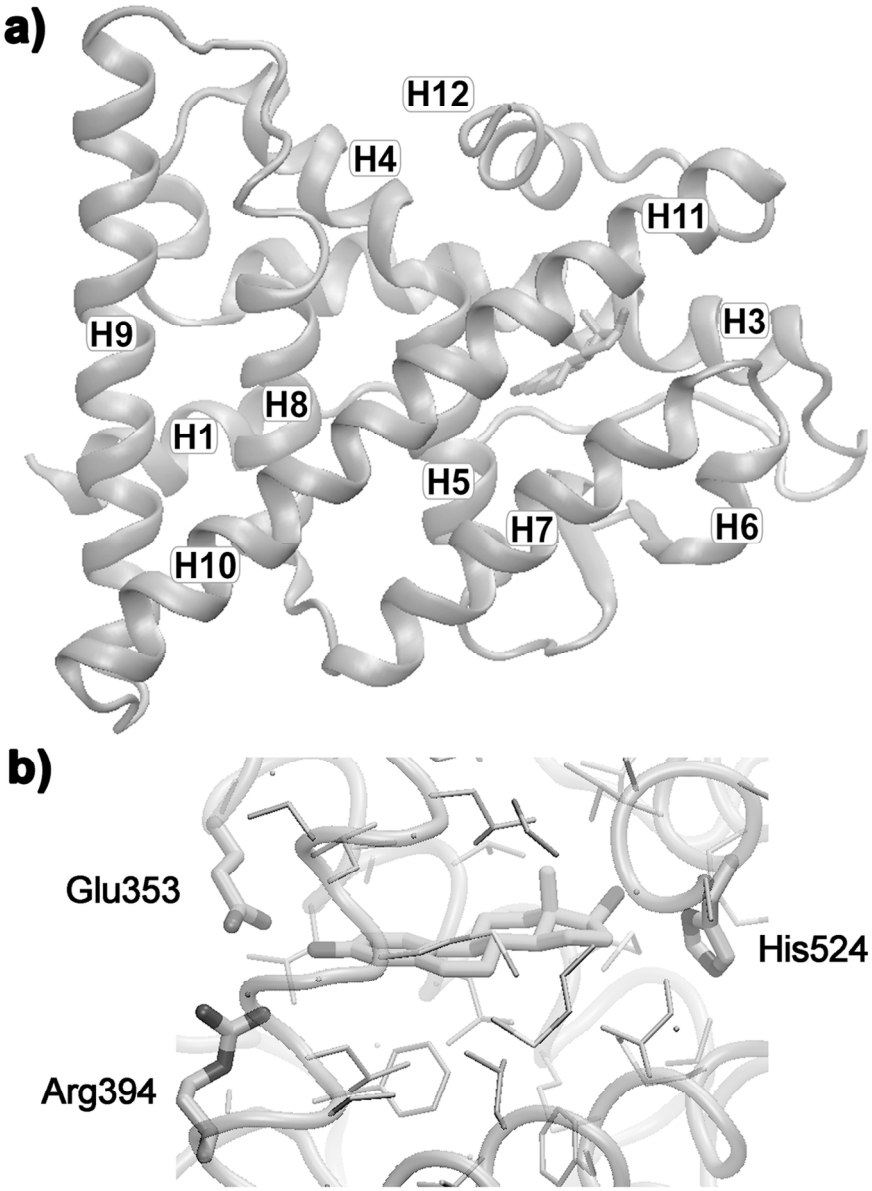

FIGURE 2. a) General view of the crystal structure of the ERα LBD/estradiol complex (pdb:1qku). b) Ligand binding mode of estradiol (**1**) in the ERα LBD/estradiol crystal structure (pdb:1qku) showing the estradiol molecule and the polar amino acid side chains involved in ligand binding.
110x145mm (300 x 300 DPI)

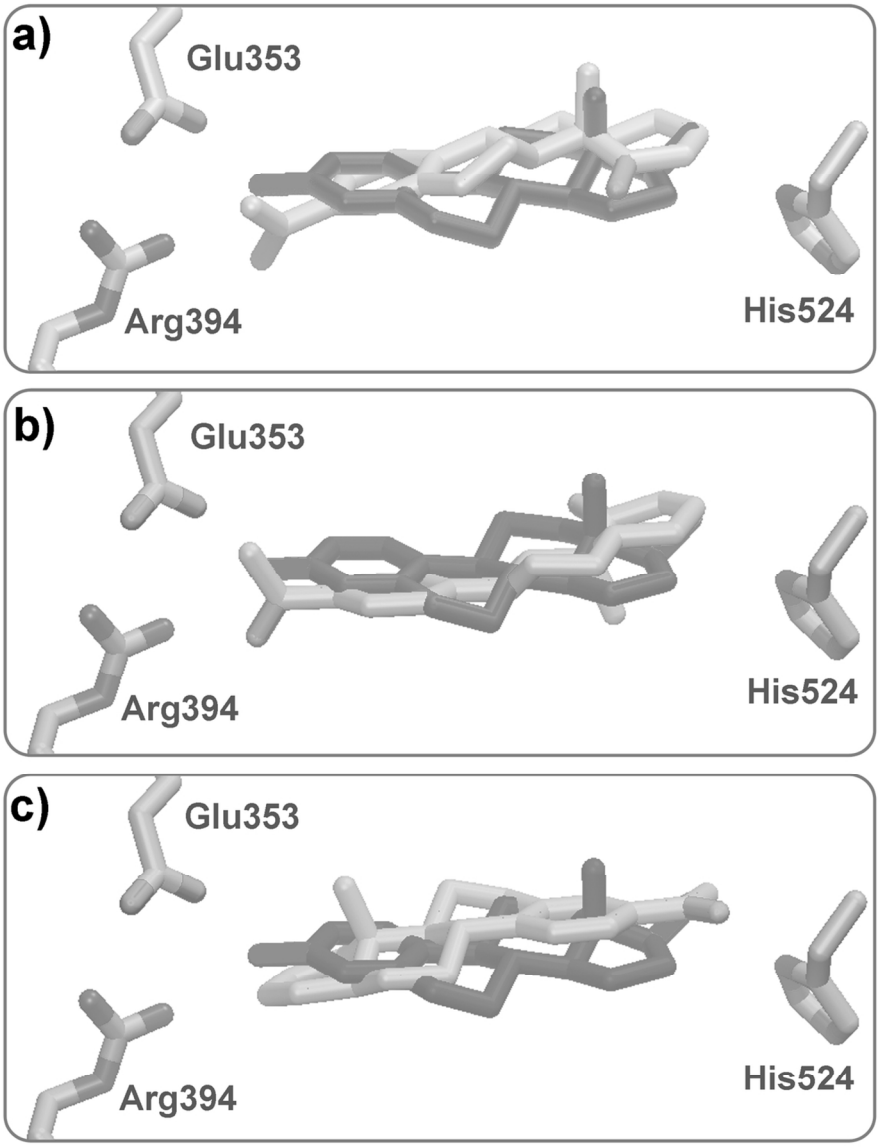

FIGURE 3. Docking solutions of compound **3** (light gray) into the ERα ligand binding pocket.[23] a) Pose A; b) Pose B; c) Pose C. The estradiol molecule is shown as reference in dark gray
109x133mm (300 x 300 DPI)

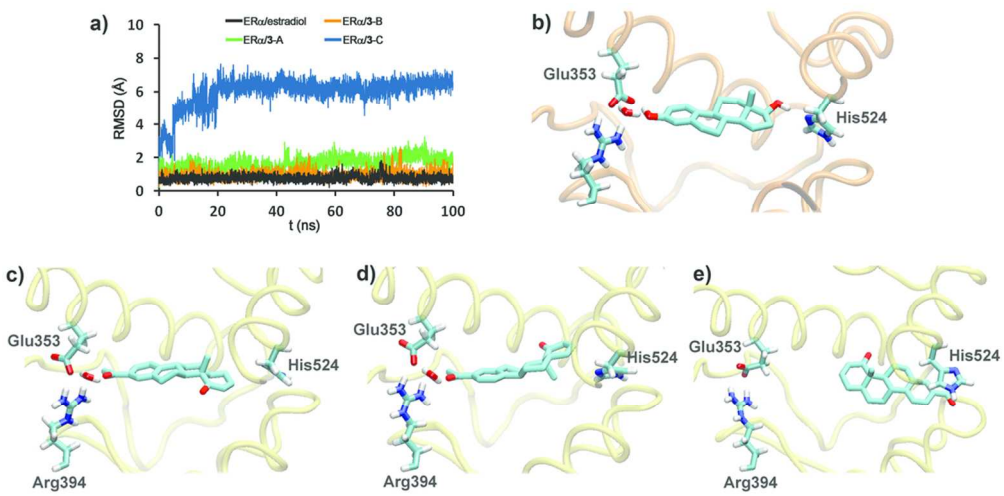

FIGURE 4. a) RMSD from the initial structures measured over ligand atoms of the simulated systems. b) Representative snapshot of the estradiol binding mode during the MD simulation. c-e) Representative snapshots of compound **3** binding mode during the MD simulation in pose A (c), pose B (d) and pose C (e) 100x50mm (300 x 300 DPI)

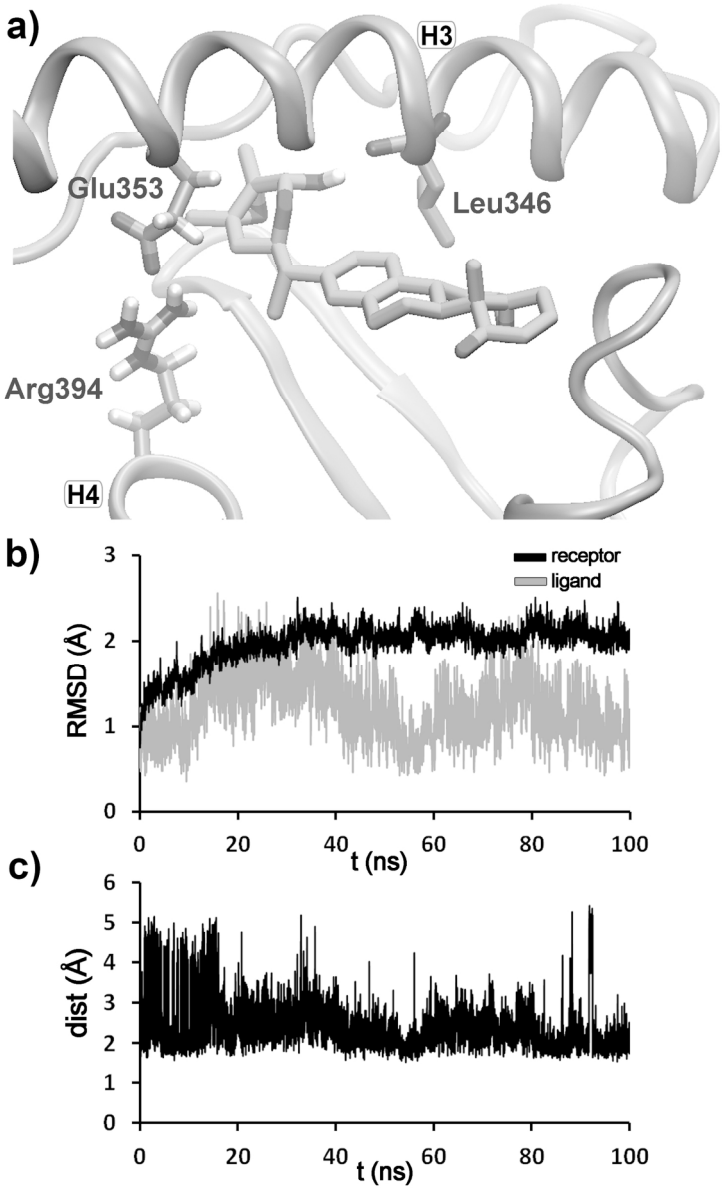

FIGURE 5. a) Representative snapshot of the binding mode of **2** during the MD simulation of the ERα/**2** complex. b) RMSD of the initial structures measured over ligand atoms and over receptor backbone atoms. c) Time evolution of the distances between the hydrogen atom of the 26-hydroxyl group of **2** and the backbone oxygen atom of Leu346
131x205mm (300 x 300 DPI)

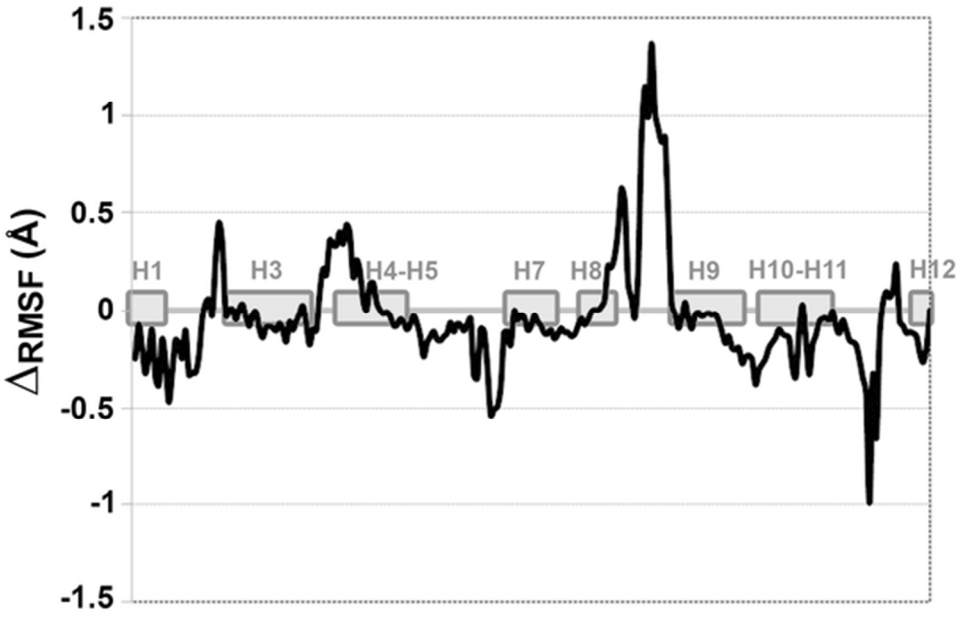

FIGURE 6. Difference between the RMSF values of the ERα/estradiol and the ERα/**3**-A systems (ΔRMSF = $RMSF_{ERα/estradiol}$ - $RMSF_{ERα/\mathbf{3}\text{-A}}$). The secondary structure of ER LBD is schematized along the *x*-axis
53x34mm (300 x 300 DPI)

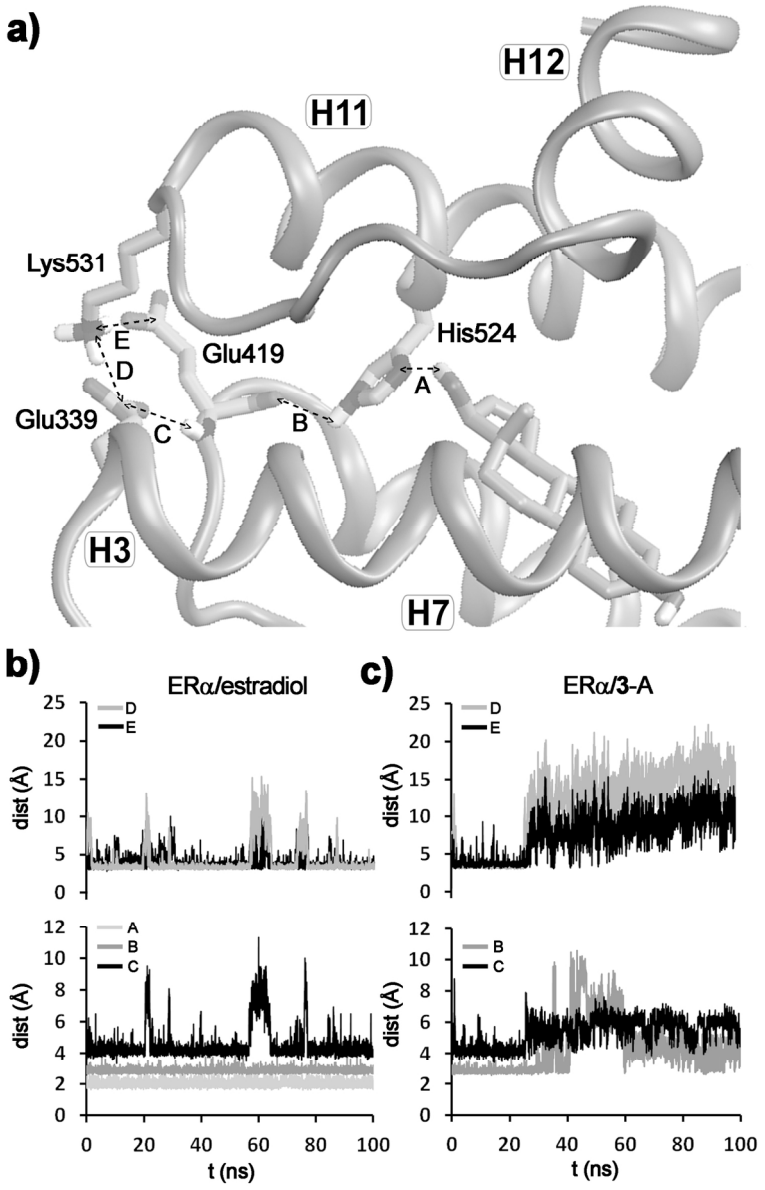

FIGURE 7. Representative snapshot of ERα/estradiol complex showing the extensive hydrogen bonding network involving residues of H3, H7 and H11. b-c) Time evolution of the distances between the hydrogen atom of the C17-hydroxyl of estradiol and the ND1 nitrogen atom of His524 (A distance), the HE2 hydrogen atom of His524 and the oxygen of the backbone carbonyl of Glu419 (B distance), the hydrogen atom of Glu419 amino group and the CD oxygen atom of Glu339 (C distance); the CD oxygen atom of Glu339 and the NZ nitrogen atom of Lys531 (D distance) and the CD oxygen atom of Glu419 and the NZ nitrogen atom of Lys531 (E distance) in the ERα/estradiol (b) and ERα/**3**-A (c) systems
129x203mm (300 x 300 DPI)

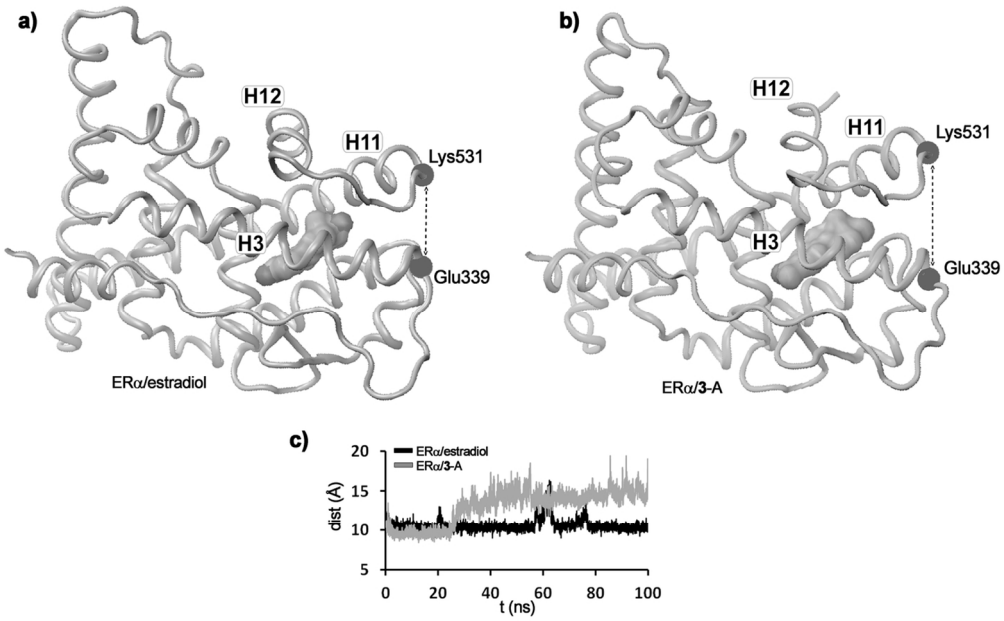

FIGURE 8. a-b) 100 ns snapshots of the ERα/estradiol (a) and the ERα/**3**-A (b) complexes with the CA atoms of Lys531 and Glu339 shown as dark gray balls. c) Time evolution of the distances between CA atoms of Lys531 and Glu339 in the ERα/estradiol and the ERα/**3**-A complexes.
119x75mm (300 x 300 DPI)

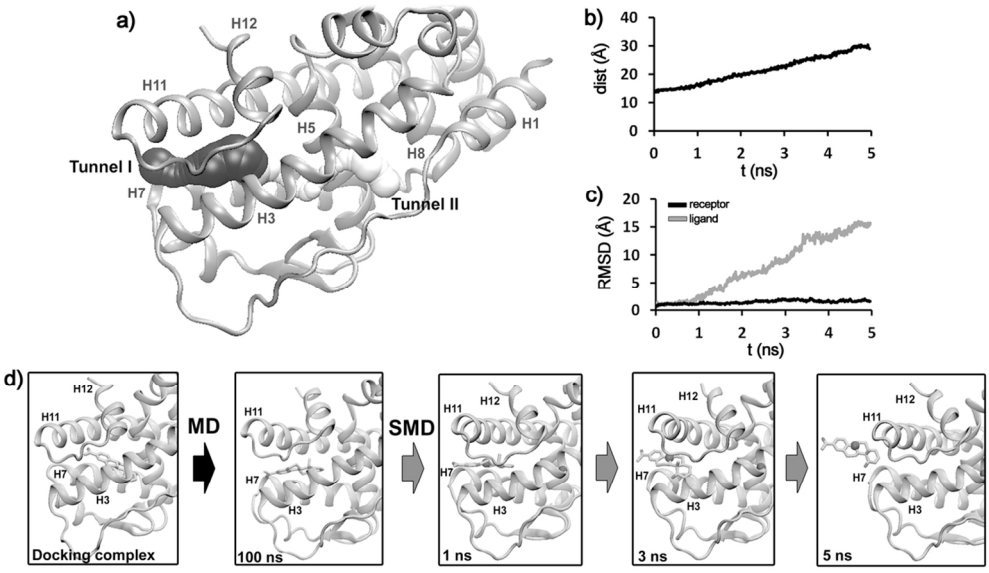

FIGURE 9. a) Initial snapshot of the ERα/**3**-C system showing tunnels detected by CAVER 3.0. Tunnel I and Tunnel II are represented by a sequence of dark gray and light gray balls, respectively. b) Time evolution of the distance between pulled atoms in the SMD simulation. c) RMSD of the initial structures measured over ligand atoms (gray) and over receptor backbone atoms (black); d) 1 ns, 3 ns and 5 ns snapshots of the SMD simulation. Pulled atoms are shown as dark gray balls
110x64mm (300 x 300 DPI)