Accepted Manuscript

Potent aromatase inhibitors and molecular mechanism of inhibitory action

Hongjun Kang, Xingqing Xiao, Chao Huang, Yan Yuan, Dongyan Tang, Xiaochang Dai, Xianghui Zeng

PII: S0223-5234(17)30960-1

DOI: 10.1016/j.ejmech.2017.11.057

Reference: EJMECH 9932

To appear in: *European Journal of Medicinal Chemistry*

Received Date: 9 August 2017

Revised Date: 28 October 2017

Accepted Date: 20 November 2017

Please cite this article as: H. Kang, X. Xiao, C. Huang, Y. Yuan, D. Tang, X. Dai, X. Zeng, Potent aromatase inhibitors and molecular mechanism of inhibitory action, *European Journal of Medicinal Chemistry* (2017), doi: 10.1016/j.ejmech.2017.11.057.

Potent aromatase inhibitors and molecular mechanism of inhibitory action

Hongjun Kang[a,1], Xingqing Xiao[b,1], Chao Huang[c], Yan Yuan[d], Dongyan Tang[a], Xiaochang Dai[a,*], Xianghui Zeng[a, e,*]

[a] *Key Laboratory of Medicinal Chemistry for Natural Resource, Yunnan University, Ministry of Education, School of Chemical Science and Technology, Yunnan University, Kunming 650091, P.R. China*

[b] *Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, North Carolina 27695-7905, USA*

[c] *Engineering Research Center of Biopolymer Functional Materials of Yunnan, Yunnan Minzu University, Kunming, 650500, P.R. China*

[d] *School of Ethnic Medicine, Yunnan Minzu University, Kunming, 650504, P.R. China*

[e] *Current address: Department of Pharmacy, Faculty of Health and Medical Sciences, University of Copenhagen. DK-2100 Copenhagen, Denmark*

[1] *Co-first author*

Correspondence: xchdai@ynu.edu.cn (X.D.), zeng_89@hotmail.com (X.Z.)

Highlights

► Discovery of novel non-steroidal inhibitors against aromatase

► Novel compounds demonstrate higher potencies in aromatase inhibitory than letrozole

► Computational models of aromatase with different substrates are built

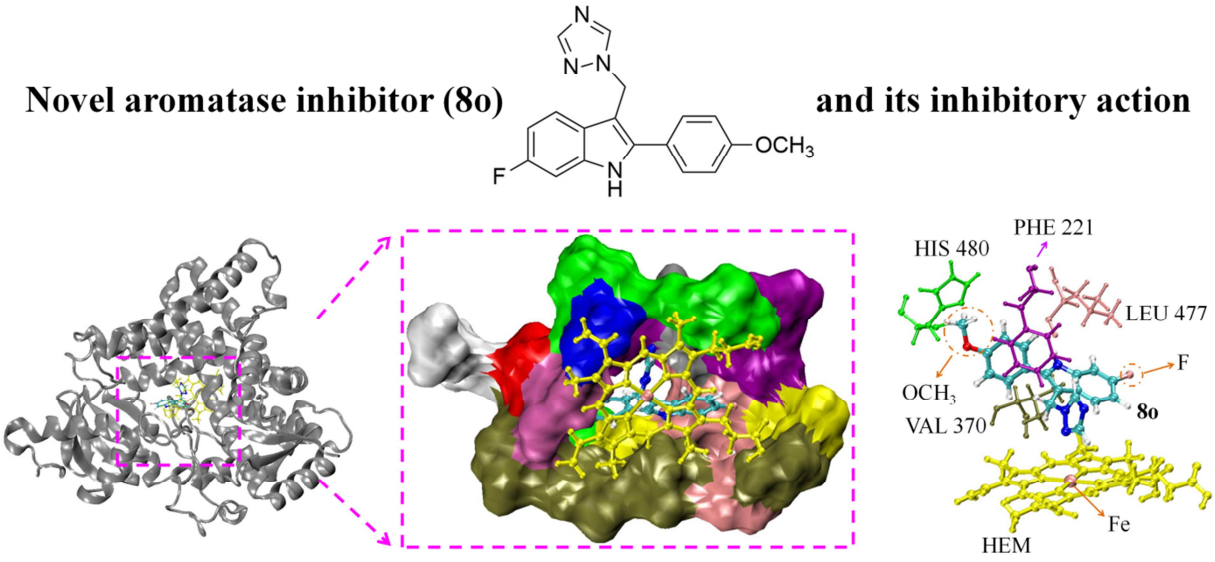
► Binding mode of substrates to aromatase has significant impact on pharmacological effect

Abstract

Estrogen is a significant factor in the maintenance and progression of hormone-dependent breast cancer. As well known, aromatase mediates the production of estrogen. Thus, reduction of aromatase with chemical molecules has been considered to be an effective treatment for estrogen receptor-positive (ER+) breast cancer. In this work, we designed and synthesized a series of novel non-steroidal molecules containing 2-phenylindole scaffold and moiety of either imidazole or 1,2,4-triazole to enhance their binding capacity with the aromatase. Among these molecules, a compound named as **8o** was confirmed experimentally to have the highest inhibitory activity to aromatase. Further cell activity assay proved that compound **8o** has low cytotoxicity and is a promising lead for

developing novel aromatase inhibitors. Molecular modelling and simulation techniques were performed to identify the binding modes of letrozole and **8o** with the aromatase. Analysis of energy of the two compound-aromatase complexes revealed that the **8o** has a low binding energy (strong binding affinity) to the aromatase as compared to letrozole, which was in accordance with the experimental results. As concluded, a combination of experimental and computational approaches facilitates us to understand the molecular mechanism of inhibitory action and discover more potent non-steroidal AIs against aromatase, thereby opening up a novel therapeutic strategy for hormone-dependent breast cancer.

Graphical Abstract



Keywords

Estrogen; breast cancer; aromatase inhibitors (AIs); letrozole; molecular mechanism of inhibitory action

1. Introduction

Breast cancer, which is well known as a common disease in women, remains a high rate of death for female patients worldwide [1, 2]. It is widely accepted that the level of estrogens and the profile of estrogen receptors are two measurable indicators towards a risk assessment for breast tumors [3]. Since nearly 70% of the patients with breast cancer are estrogen receptor-positive (ER+) and may be estrogen dependent, estrogen deprivation has been considered an attractive therapeutic strategy for ER-positive breast cancer [4]. Aromatase, otherwise called CYP19，is a rate-limiting

enzyme in the biosynthesis of estrogens that is responsible for the conversion of androgens including androstenedione and testosterone into estrone and estradiol [5]. For the postmenopausal women with ER-positive breast cancer, aromatase is usually overexpressed in the breast tissues and is the main source of local estrogen production in neoplastic tissues [5]. Interfering with aromatase activity and reducing the level of estrogens in tumor tissues may slow down the growth of breast cancer cells, eventually extending the lifespan of patients [6].

According to the diversity in molecular skeletons, aromatase inhibitors can be categorized into two types: steroidal and non-steroidal blockers [6]. Steroidal AI (e.g., exemestane in Fig.1) derived from the substrate androstenedione interacts with aromatase through chemical actions, resulting in an irreversible binding process of the species; while non-steroidal AI (e.g., anatrozole and letrozole in Fig.1) binds to enzyme through non-covalent interactions, resulting in a reversible binding process.

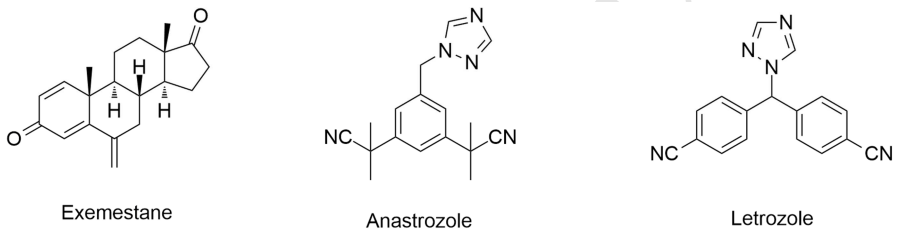


Fig.1. Chemical structures for steroidal (exemestane) and nonsteroidal (anastrozole, letrozole) aromatase inhibitors (AIs).

Nowadays, aromatase inhibitors (AIs) constitute the first-line drugs for ER-positive breast cancer in postmenopausal women. The third generation of AIs (exemestane, anatrozole, and letrozole) was approved by Food and Drug Administration as first-line therapy for hormonally-responsive breast cancer in postmenopausal women, since they had been proven to be superior to tamoxifene known as a representative of selective estrogen receptor modulators (SERMs) [7]. However, with the broad applications of AIs in clinical practices, some unexpected problems are gradually shown up, such as non-responses to some of patients, resistance to AI treatment and inhibition of some CYP450 enzymes [8]. Hence, there are urgent needs to discover and develop new generation of AIs to overcome the defects.

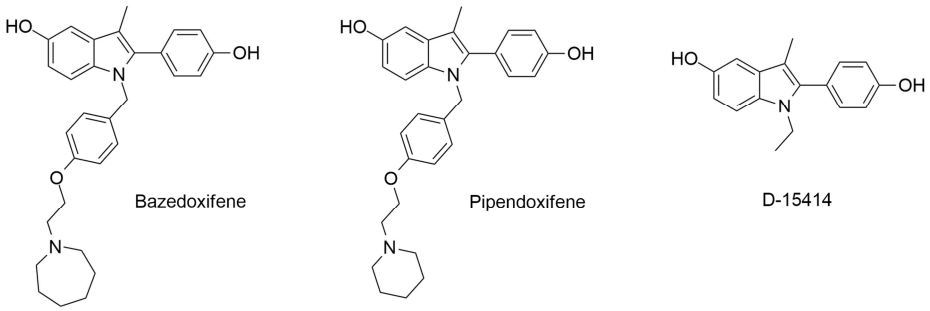


Fig.2. Chemical structures for selective estrogen receptor modulators (SERMs) based on 2-phenyl indole scaffold.

2-Phenyl indole, an estrogen-mimicking chemical structure, has been used widely in drug designs for treatment of estrogen-related diseases [9, 10, 11]. Some SERMs (e.g., bazedoxifene and pipendoxifene, Fig.2) are based on 2-phenyl indole skeletons. D-15414 (Fig.2), a 2-phenyl indole analogue, is a non-steroidal estrogen with fairly high binding affinity for the estrogen receptor [12]. Thus, we are reasonable to postulate that 2-phenyl indole derivatives may enter into the substrate-binding pocket of aromatase, because they are mimics of estrone and estradiol, the products of aromatization reactions, with high binding affinity to aromatase enzyme. On the other hand, imidazole and 1,2,4-triazole are nitrogen-containing heterocyclic rings that can coordinate with the heme of aromatase [13]. Based on the knowledge above, a series of novel non-steroidal aromatase inhibitors were designed by introducing the azole group at the 3 position in the 2-phenyl indole framework (Fig.3A). The indole moiety accounts for fitting the binding site of aromatase, whereas the azole moiety chelates the iron atom in the heme existing in the active site of aromatase. Since our designed molecules are structurally similar to letrozole (Fig.3B), it is possible for the azole substituted 2-phenyl indole to have certain bioactivity in inhibiting aromatase.

To study the impact of the modified 2-phenylindoles by either imidazole or 1,2,4-triazole on aromatase activity, we synthesized twenty imidazole or 1,2,4-triazole substituted 2-phenyl indoles (namely, compounds **8a**-**t** (Table 1)), and tested their inhibitory activities against aromatase in our lab. To explore the molecular mechanism of inhibitory action, an integrated computational strategy was then applied to investigate the binding of letrozole and compound **8o** to aromatase.

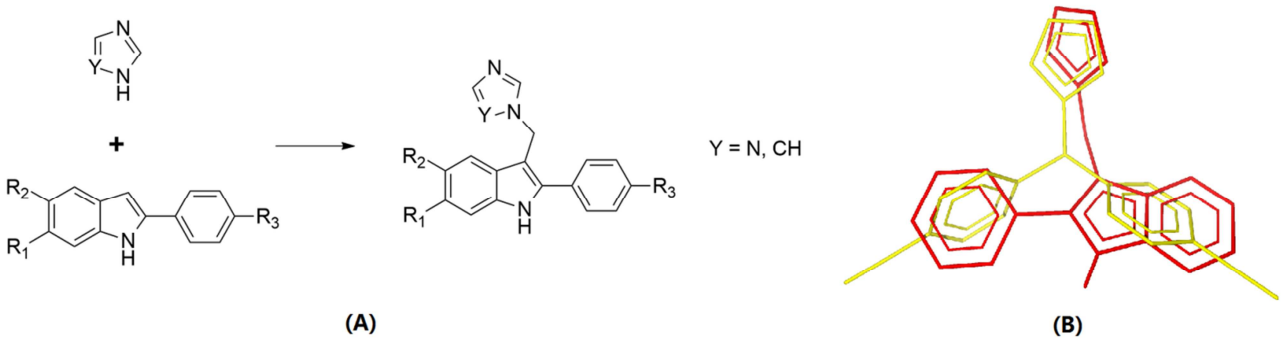
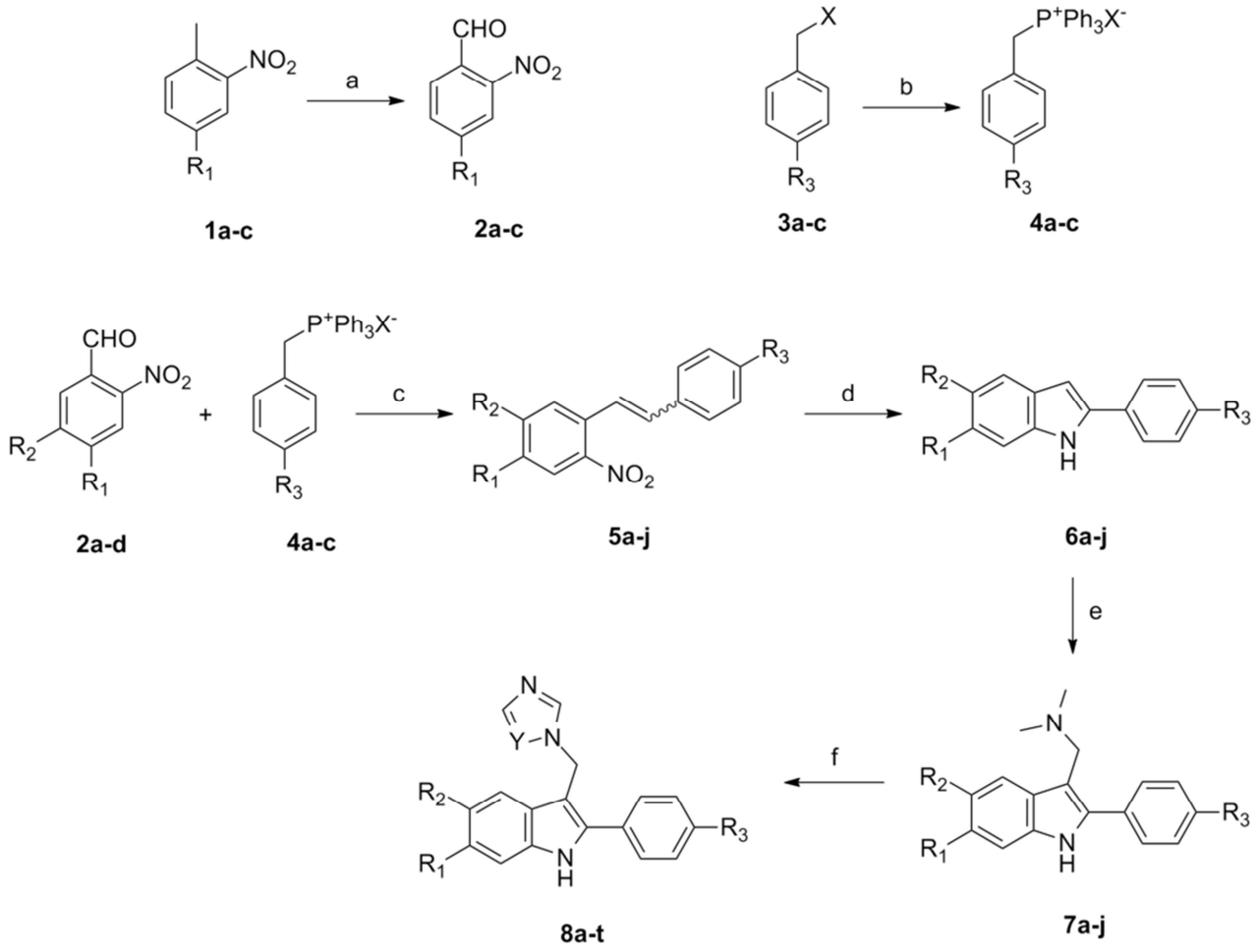


Fig.3. (A) Rational design of new AIs based on 2-phenyl indole scaffold; (B) The overlapping of one designed molecule (red) and letrozole (yellow).

2. Results and Discussion

2.1. Chemistry

The synthesis route of the twenty imidazole or 1,2,4-triazole substituted 2-phenyl indole aromatase inhibitors **8a-t** is shown in Scheme 1. The precursors **5a-j** were synthesized by Wittig reaction of substituted 2-nitrobenzaldehydes **2a-d** and quaternary phosphonium salts **4a-c** in the presence of organic base DBU. Reductive cyclization of nitrostyrene derivatives **5a-j** to the 2-phenyl indole intermediates **6a-j** was readily accomplished by triphenylphosphine using a dichlorodioxomolybdenum (VI) complex ($MoO_2Cl_2(dmf)_2$) as catalyst. Compounds **7a-j** were prepared from the precursors **6a–j** by Mannich reaction. Finally, the target compounds **8a-t** were prepared by treatment of 3-dimethylaminomethyl-2-phenyl indoles **7a–j** with imidazole or 1,2,4-triazole in hot xylene. The structures of all target compounds **8a-t** were confirmed by IR, LC-MS, ^{1}H NMR and ^{13}C NMR spectroscopy.



8a, R_1=F, R_2=H, R_3=F, Y=CH
8b, R_1=F, R_2=H, R_3=Cl, Y=CH
8c, R_1=Cl, R_2=H, R_3=F, Y=CH
8d, R_1=Cl, R_2=H, R_3=Cl, Y=CH
8e, R_1=F, R_2=H, R_3=OCH_3, Y=CH
8f, R_1=Cl, R_2=H, R_3=OCH_3, Y=CH
8g, R_1=OCH_3, R_2=H, R_3=F, Y=CH
8h, R_1=OCH_3, R_2=H, R_3=Cl, Y=CH
8i, R_1,R_2=5, 6-methylenedioxy, R_3=F, Y=CH
8j, R_1,R_2=5, 6-methylenedioxy, R_3=Cl, Y=CH
8k, R_1=F, R_2=H, R_3=F, Y=N
8l, R_1=F, R_2=H, R_3=Cl, Y=N
8m, R_1=Cl, R_2=H, R_3=F, Y=N
8n, R_1=Cl, R_2=H, R_3=Cl, Y=N
8o, R_1=F, R_2=H, R_3=OCH_3, Y=N
8p, R_1=Cl, R_2=H, R_3=OCH_3, Y=N
8q, R_1=OCH_3, R_2=H, R_3=F, Y=N
8r, R_1=OCH_3, R_2=H, R_3=Cl, Y=N
8s, R_1,R_2=5, 6-methylenedioxy, R_3=F, Y=N
8t, R_1,R_2=5, 6-methylenedioxy, R_3=Cl, Y=N

Scheme 1. Synthesis route of twenty imidazole or 1,2,4-triazole substituted 2-phenyl indole AIs **8a-t**. Reagents and conditions: (a) i. DMF-DMA, pyrrolidine, DMF, 125℃, 6h, 95-100%; ii. $NaIO_4$, THF/H_2O, rt, overnight, 63-67%; (b) PPh_3, toluene, reflux, 6h, 72-98%; (c) DBU, THF, reflux, overnight, 91-96%; (d) $MoO_2Cl_2(dmf)_2$, PPh_3, toluene, reflux, 16h, 50-67%; (e) $CH_2{=}N^+(CH_3)_2Cl^-$, CH_2Cl_2, rt, 8h, 88-95%; (f) azole, xylene, 130℃, 1h, 25-56%.

2.2. Biological activity

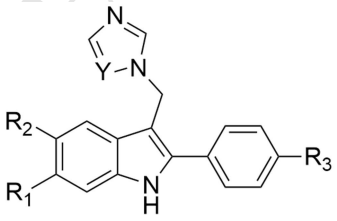
2.2.1. Evaluate the inhibitory potencies of AIs against aromatase

Aromatase inhibitory activities of the twenty 2-phenyl indole compounds **8a-t** were determined using enzyme-linked immunosorbent assay (ELISA) as described in the experimental section, and the values of IC_{50}/aromatase were given in Table 1. The most potent compounds in this series are **8o**, **8c** and **8e** (IC_{50}/aromatase: 14.1 nM, 32.3 nM, and 36.1 nM, respectively), which are more powerful in the inhibition of aromatase, as compared to letrozole (IC_{50}/aromatase: 49.5 nM). Additionally, compounds **8a**, **8k** and **8s** are also found to have good inhibitory activity to aromatase (IC_{50}/aromatase: 51.2 nM, 54.4 nM and 58.3 nM, respectively).

As summarized in Table 1, we learnt that the R_1-substituted compounds with a large methoxy group in benzene ring of indole exhibit a weak inhibitory activity to aromatase, indicating that the R_1 position is sensitive to steric hindrance. Thus, small group is required at the R_1 position. The atom F with larger electronegativity is a better choice rather than the atom Cl, attributing to the steric flexibility for the polar amino acids on aromatase in vicinity of the R_1 position. However, when the methoxy group is replaced in the 2-phenyl group for the R_3-substituted compounds, they become potent AIs. Comparing the imidazole- and triazole-substituted compounds, it seems that triazole has a stronger AI activity. This can be explained by the different polarities in the active pocket of the enzyme as well as the spatial orientations.

Table 1

Aromatase inhibition activity and cytotoxicity of the twenty 2-phenyl indole AIs **8a-t** and letrozole.



Compound No.	R_1	R_2	R_3	Y	IC_{50}/aromatase (nM)	IC_{50}/MCF-7 (μM)
8a	F	H	F	CH	51.2	83.2
8b	F	H	Cl	CH	82.9	45.0
8c	Cl	H	F	CH	32.3	41.5
8d	Cl	H	Cl	CH	260.8	>1000
8e	F	H	OCH_3	CH	36.1	>1000
8f	Cl	H	OCH_3	CH	276.8	>1000
8g	OCH_3	H	F	CH	108.7	>1000
8h	OCH_3	H	Cl	CH	1058	>1000

8i	5, 6-methylenedioxy		F	CH	210.6	>1000
8j	5, 6-methylenedioxy		Cl	CH	2635	>1000
8k	F	H	F	N	54.4	>1000
8l	F	H	Cl	N	217.4	>1000
8m	Cl	H	F	N	96.9	625
8n	Cl	H	Cl	N	207.5	382
8o	F	H	OCH_3	N	14.1	325
8p	Cl	H	OCH_3	N	196.5	14.4
8q	OCH_3	H	F	N	237.6	12.2
8r	OCH_3	H	Cl	N	7772	612
8s	5, 6-methylenedioxy		F	N	58.3	334
8t	5, 6-methylenedioxy		Cl	N	677.4	354
Letrozole					49.5	4.73

2.2.2. Cytotoxicity assay

The cytotoxicity of the twenty 2-phenyl indole AIs **8a-t** and letrozole was further tested in MCF-7 cells, as listed in Table 1. Results of MCF-7 cell activity assay revealed that all the azole substituted 2-phenyl indole AIs had lower cytotoxicity than letrozole. The most toxic compounds in this series are compound **8p**, **8q** (IC_{50}/MCF-7: 14.4 μM and 12.2 μM, respectively), which are less toxic than letrozole (IC_{50}/MCF-7: 4.73 μM). Compound **8o**, the most potent AI in this series, exhibits low cytotoxicity (IC_{50}/MCF-7: 325 μM). This signified that compound **8o** has a good safety profile.

Anticancer effect of AIs is resultant of their impacts on human body by reduction of estrogen levels in the circulation and tumor tissues, making it hard for ER-positive breast cancer cells to grow and spread [14]. Hence, the AIs are not required to have high cytotoxic activity to directly kill cancer cells. Novel potent, more selective, less toxic AIs are needed because of many side effects resulted from the long time use of them. Compound **8o** is an excellent lead for the development of the new generation of AIs, as it possesses highly potent aromatase inhibitory activity and low toxicity.

2.3. Molecular modeling studies

2.3.1. Computationally Probe binding mechanism of drug molecules with aromatase at a molecular level

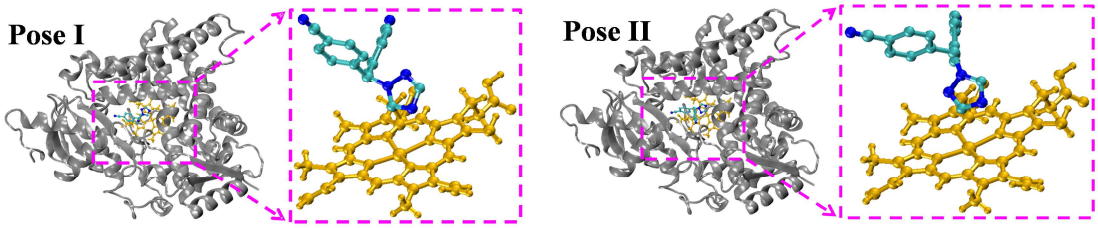
To explore molecular mechanism of inhibitory action, we applied an integrated computational strategy combining quantum mechanics (QM) calculation, molecular docking and atomistic molecular dynamics (MD) simulation in this work to investigate the binding modes of aromatase with letrozole and **8o** (the most powerful compound, IC_{50}/aromatase: 14.1 nM) [15]. QM calculations were done for letrozole, **8o** and iron porphyrin to optimize geometry and derive partial atomic charges. Molecular docking technique was then used to determine binding sites of aromatase and to predict possible binding poses of letrozole and **8o**. Finally, atomistic MD simulations were employed to examine the stability of the complexes letrozole-aromatase and **8o**-aromatase. Binding free energy was calculated on the last 5 ns simulation trajectory using molecular mechanics/generalized born surface area (MM/GBSA) approach with variable dielectric constant model [16, 17]. Details of our computational protocol are given in the experimental section and the supporting information file.

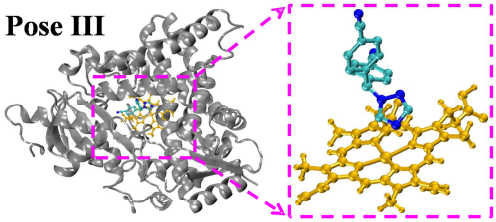
2.3.2. Determine binding structures of aromatase with AIs

Structure of aromatase containing an iron porphyrin is available in Protein Data Bank (PDB entry: 3EQM) [18]. Binding modes of letrozole and compound **8o** on aromatase were studied using AutoDocK Vina program based on the minimum-energy conformers of letrozole (Fig.S1-C) and **8o** (Fig.S2-C) obtained from geometrical optimizations [19]. Nine putative poses were yielded from each individual docking attempt for letrozole and **8o**. Six out of the nine poses showed that both the letrozole and **8o** reside in the central pocket of aromatase. (It is mentioned that the iron porphyrin is centered in the pocket of aromatase.) The other three poses of letrozole and **8o** are located in varied surficial regions of aromatase. Thus, it is reasonable to believe that the binding site of aromatase is in its central pocket. After removing some similar poses from the six poses of letrozole and **8o** in the aromatase pocket, three potential poses of letrozole and **8o** were reserved (Fig.4). For convenience, we named the three docking poses of letrozole as Poses I, II and III (Fig.4A), and called the three poses of **8o** as Poses 1, 2 and 3 (Fig.4B). Both the letrozole and **8o** are buried in interior of aromatase and interact with iron porphyrin.

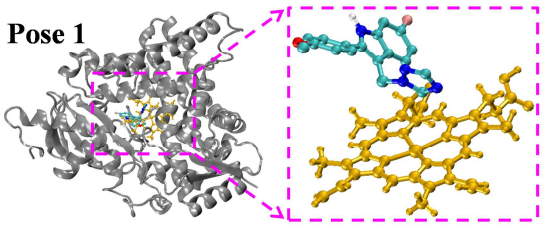
The molecular docking results show only the static orientation and interactions of AIs with aromatase. Therefore, it is necessary to employ explicit-solvent atomistic MD simulations to study their dynamics properties of AIs and aromatase. We examined the binding stability of all the six poses of letrozole and compound **8o** in six independent 50-ns simulations. Binding free energy of the AIs and aromatase over the last 5 ns of the simulation trajectory was calculated using the

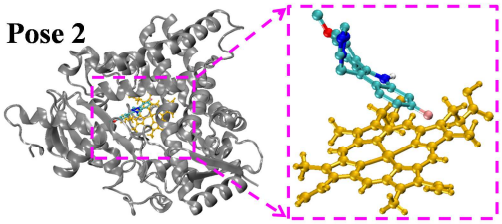
MM/GBSA approach with the variable dielectric constant model. Fig.4C lists our computed binding free energies of the six docking poses associated with the letrozole-aromatase and **8o**-aromatase. For the letrozole-aromatase complex, Pose I has a lower binding free energy, -19.96±0.16 kcal/mol, than Poses II and III. Thus, we can reasonably infer that Pose I is the most likely binding mode of letrozole and aromatase. For the **8o**-aromatase complex, Pose 1 is considered to be the most likely binding mode due to the low binding free energy -23.43±0.21 kcal/mol. Additionally, by comparing Pose I of the letrozole-aromatase and Pose 1 of the **8o**-aromatase, we found that Pose 1 has a lower binding free energy (-23.43±0.21 kcal/mol) than Pose I (-19.96±0.16 kcal/mol). It signifies that aromatase has a stronger preference to compound **8o** rather than letrozole, which is in agreement with our experimental measurements in Table 1: the IC_{50}/aromatase value (14.1 nM) of **8o** is lower than that (49.5 nM) of letrozole.

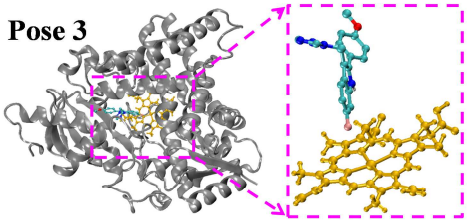




(a) letrozole & aromatase







(b) 8o & aromatase

(c)

letrozole & aromatase	Pose I	Pose II	Pose III
$\Delta G_{binding}^{avg}$	**-19.96±0.16**[i]	-14.33±0.18	-16.89±0.16
8o & aromatase	Pose 1	Pose 2	Pose 3
$\Delta G_{binding}^{avg}$	**-23.43±0.21**	-20.97±0.14	-20.93±0.18

(i) Calculation errors are estimated by the standard errors of the mean.

Fig.4. (A) Three poses of letrozole in aromatase, viz. Poses I, II and III, are obtained via the AutoDock Vina program, where the letrozole is buried in interior of aromatase and interacts with iron porphyrin. (B) Three poses of compound **8o** in aromatase, viz. Poses 1, 2 and 3, are obtained. Aromatase is represented by a silver ribbon, iron porphyrin is specified

in yellow, and letrozole and **8o** are exhibited in multi-colored bead-and-bond model. (C) Comparison of computed binding free energies between letrozole and aromatase and between compound **8o** and aromatase. (unit: kcal/mol)

2.3.3. Analysis of energy decomposition of AIs bound to aromatase

Clustering analysis was performed on the last 5 ns simulation trajectories of Pose I (letrozole-aromatase) and Pose 1 (**8o**-aromatase) to obtain representative structures of the complexes at equilibrium [20]. Figs.5A and 6A show the representative structures of the complexes letrozole-aromatase and **8o**-aromatase, respectively. Both letrozole and **8o** with their triazole pointing at iron porphyrin were buried in a limited pocket of aromatase, leading to a restricted motion of molecules. To explore which parts of the AIs contribute to the binding affinity and which parts contribute to binding specificity, we calculated the inter-chain van der Waals (VDW) and charge-charge (ELE+EGB) interactions between the AIs and aromatase. The first two terms in Equation (S-3) were calculated and plotted in Figs.5B and 6B. Detailed calculations of energy decomposition were given in the Supporting information. To study partial energy contributions of AIs to the whole binding with aromatases, we purposely divided letrozole into two components, viz. triazole and other (Fig.S1-A), and purposely divided **8o** into four components, viz. F, OCH_3, triazole and other (Fig.S2-A). Meanwhile, iron porphyrin was also separated into two components: HEM and Fe (Fig.S3). The component Fe contains only one Fe atom centered in iron porphyrin, and the HEM part involves all the non-Fe atoms in iron porphyrin.

Firstly, let us discuss the situation of letrozole bound to aromatase (Fig.5B). Triazole of letrozole has strong inter-chain VDW interactions with the HEM of iron porphyrin, and slight inter-chain ELE+EGB interactions with the HEM and Fe. The other part of letrozole interacts with the HEM relying on not only mutual VDW attraction, but also mutual charge-charge (ELE+EGB) repulsion. Besides, strong VDW interactions were observed between the other part of letrozole and the residue Trp at site 224 (TRP 224) on aromatase. The residues VAL at site 373 (VAL 373) and Met at site 374 (MET 373) prefer to interact with letrozole via both inter-chain VDW and ELE+EGB energies. The interaction domain of letrozole and aromatase was enlarged in Fig.5C to highlight the relative locations of letrozole, iron porphyrin and the key residues on aromatase. As evidenced by our previous work [21, 22], the inter-chain VDW energy contributes to the binding specificity for AIs to iron porphyrin-contained aromatase, and the inter-chain ELE+EGB energy between AIs and iron porphyrin-contained aromatase is responsible for the binding affinity.

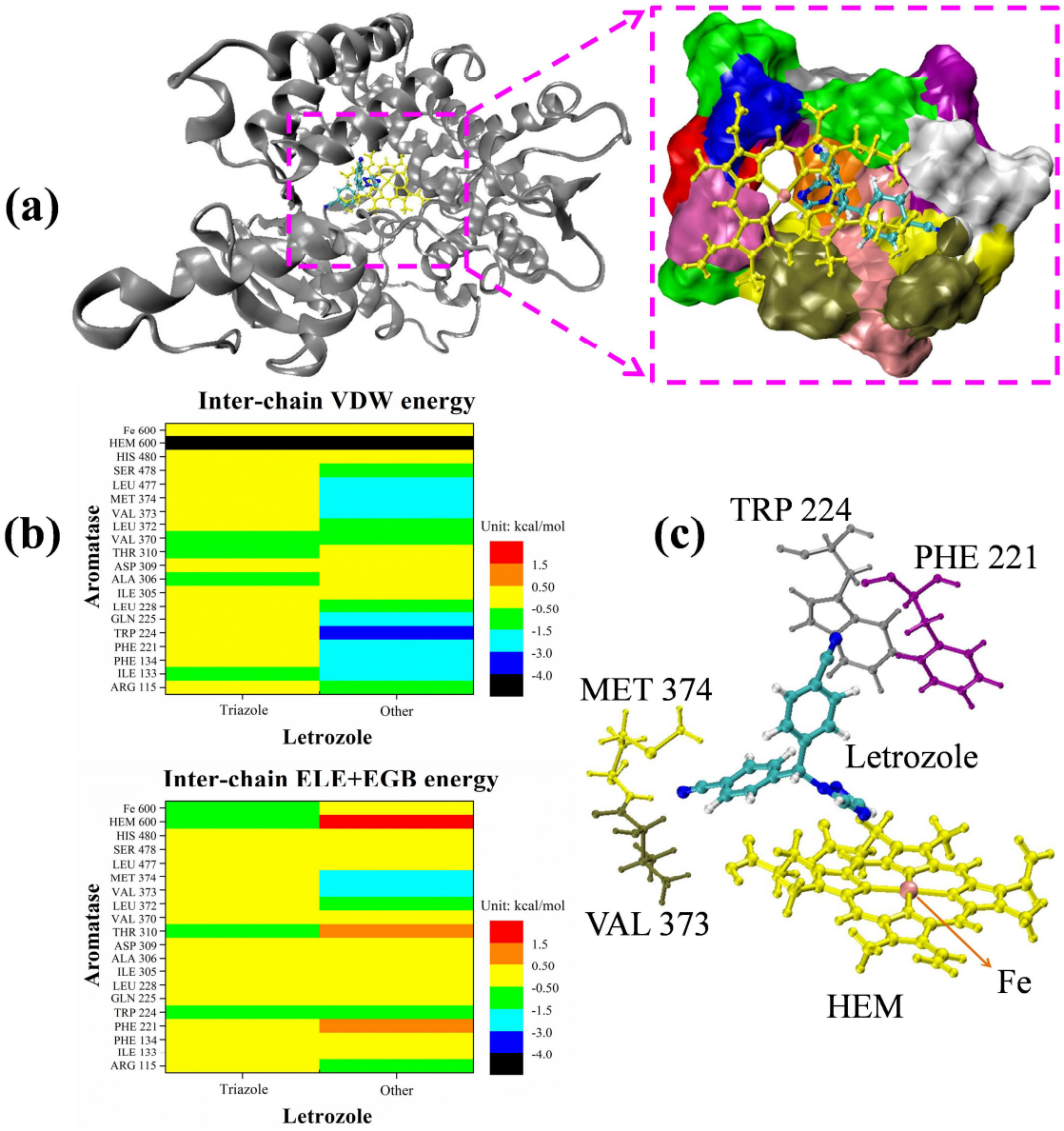


Fig.5. (A) The most likely binding structure of letrozole and aromatase. The picture enlarged on the right side demonstrates a detailed binding position and orientation of letrozole in the aromatase, where the amino acids that are found in the aromatase to be near letrozole are shown as multi-colored isosurfaces. (B) Energy panels show the inter-chain van der Waals (VDW) and charge-charge (ELE+EGB) interactions between letrozole and aromatase. The *x*-axis represents the two primary components of letrozole: triazole and other, the *y*-axis represents the iron porphyrin (HEM and Fe) and the amino acids that are near letrozole, and the color bar on the right scales the value of the energies. (C) The snapshot shows that letrozole has strong inter-chain interactions with Phe at site 221, Trp at site 224, Val at site 373, Met at site 374, and HEM and Fe.

Next, our focus is transferred on the situation of compound **8o** bound to aromatase (Fig.6B). Triazole of **8o** attracts the HEM part of iron porphyrin through the inter-chain VDW energy, but repels the HEM part through the inter-chain ELE+EGB energy. A medium inter-chain ELE+EGB

interaction was also found in between the triazole of **8o** and the component Fe of iron porphyrin. Although the atom F of **8o** does not provide strong energy contributions in the binding of aromatase, a substitution of the atom F by a larger atom Cl easily cause an atomic overlap with aromatase due to a limited pocket space (Fig.6C), thereby weakening the affinity of AIs. This explains why the compound **8p** exhibits a weak inhibitory activity against aromatase (Table 1). The group OCH_3 of **8o**, which sticks out of the central pocket of aromatase, has certain conformational flexibility so as to contact well with the residues Phe at site 221 (PHE 221), Asp at site 309 (ASP 309) and His at site 480 (HIS 480) via the inter-chain VDW energies, strengthening the thermal stability of the **8o**-aromatase. If the group OCH_3 was substituted by stiff spherical atoms, such as F and Cl, a loss of conformational flexibility might trigger either an inadequate contact for the small atom F or an atomic overlap for the large atom Cl with aromatase. For this reason, the capabilities of compounds **8k** and **8l** in the inhibition of aromatase are relatively weaker than compound **8o** (Table 1).

Based on the above discussions about letrozole and **8o**, we reasonably believed that introducing the triazole group favors molecular interactions of AIs and iron porphyrin, eventually arriving at our initial goal of suppressing bioactivity of aromatase. Meanwhile, steric effect is also an important factor in the design process of potent AIs. Proper substitutions chosen in AIs facilitate the binding of AIs with aromatase by either increasing molecular contacts or avoiding inappropriate overlaps.

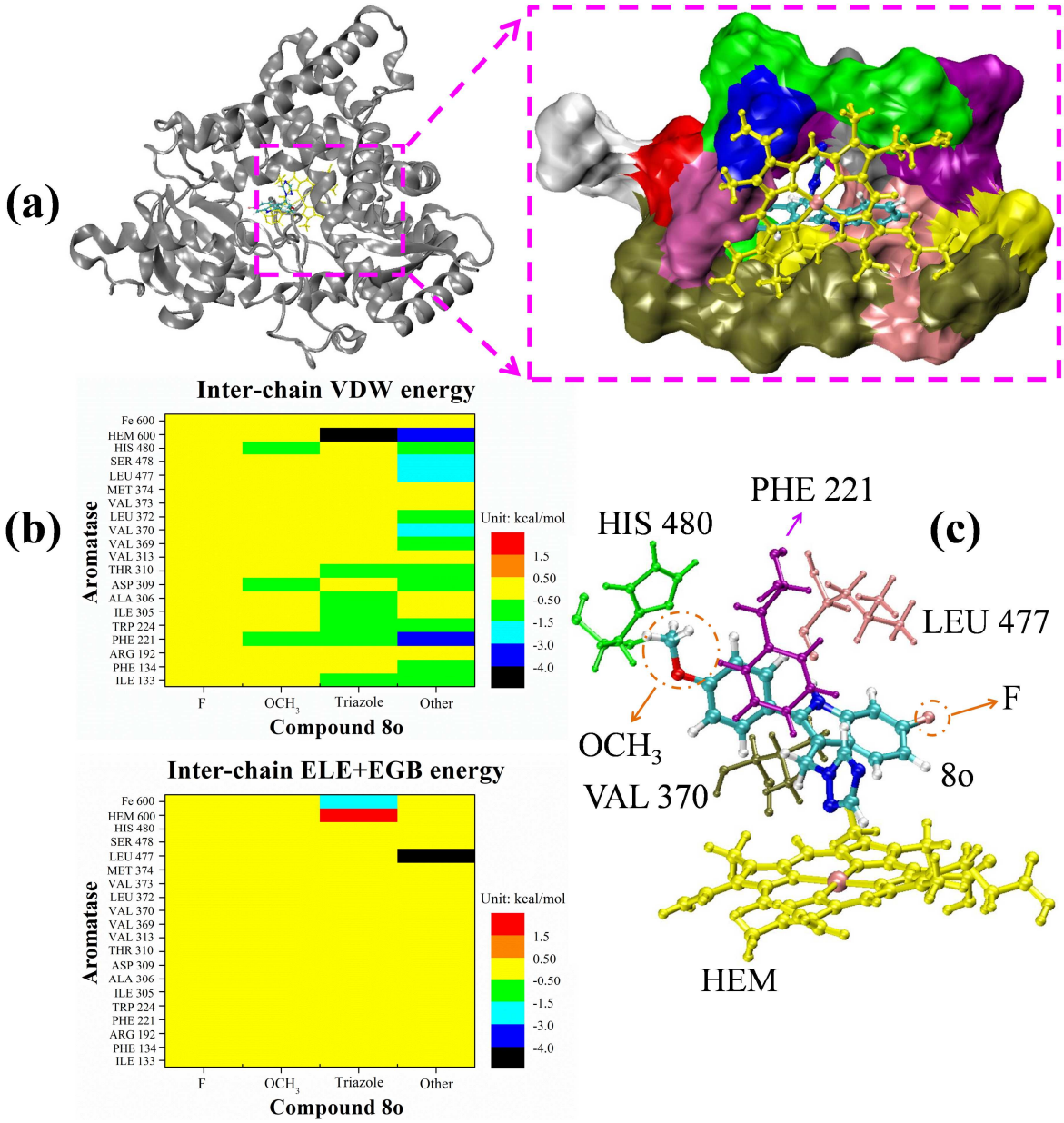


Fig.6. (A) The most likely binding structure of compound **8o** and aromatase. The picture enlarged on the right side demonstrates a detailed binding position and orientation of **8o** in the aromatase, where the amino acids that are found in the aromatase to be near **8o** are shown as multi-colored isosurfaces. (B) Energy panels show the inter-chain VDW and ELE+EGB interactions between **8o** and aromatase. The *x*-axis represents the four primary components of **8o**: fluorin (F), methy (OCH_3), triazole and other, the *y*-axis represents the iron porphyrin (HEM and Fe) and the amino acids that are near **8o**. (C) The snapshot shows that **8o** has strong inter-chain interactions with Phe at site 221, Val at site 370, Leu at site 477, His at site 480, and HEM and Fe.

3. Conclusion

The use of aromatase inhibitor or ER modulator as adjuvant therapy has been the mainstay of treatment for postmenopausal women with ER-positive early-stage breast cancer. Persistent risk of tumor recurrence remains a clinical and scientific challenge. Therefore, it is worthwhile to prevent tumor recurrence by developing an alternative strategy with better efficacy. In this work, we strived to develop novel non-steroidal AIs based on a 2-phenylindole scaffold to suppress the bioactivity of aromatase, and three compounds **8o**, **8c** and **8e** were proved experimentally to be more powerful than letrozole does. Compound **8o** is considered to be the most potent AI due to its lowest value (14.1 nM) of IC_{50}/aromatase. Additionally, results of MCF-7 cell activity assay revealed that compound **8o** exhibits much lower cytotoxicity (IC_{50}/MCF-7: 325 μM) as compared to letrozole (IC_{50}/MCF-7: 4.73 μM), indicating a better safety profile for compound **8o**. Furthermore, we computationally studied the binding structures of letrozole and compound **8o** in aromatase, providing significant impact on the AIs' efficacy. The discovery of new highly potent AIs and understanding their molecular mechanism of inhibitory action is central to further improving therapeutic options for ER-positive breast cancers.

4. Experimental and simulation section

4.1. Chemistry

All compounds were fully characterized by spectroscopic techniques. The NMR spectra were recorded on a Bruker-Avance 300 MHz spectrometer (^{1}H: 300 MHz, ^{13}C: 75 MHz) with tetramethylsilane (TMS) as the internal standard (δ 0.0 ppm), chemical shifts (δ) are expressed in ppm, and *J* values are given in Hz. Deuterated DMSO was used as a solvent. IR spectra were recorded on a FT-IR Thermo Nicolet Avatar 360 using a KBr pellet. The reactions were monitored by thin layer chromatography (TLC) using silica gel GF_{254}. The melting points were determined on an XT-4A melting point apparatus and are uncorrected. HRMS was performed on an Agilent LC-MSD TOF instrument.

All chemicals and solvents were used as received without further purification unless otherwise stated. Column chromatography was performed on silica gel (200–300 mesh).

*4.1.1. Method for the synthesis of compound **8a***

Substituted 2-nitrobenzaldehydes **2a** reacted with quaternary phosphonium salts **4a** to produce the precursors **5a** in the presence of organic base DBU. Reductive cyclization of the nitrostyrene **5a**

to the 2-phenyl indole derivative **6a** was readily accomplished by triphenylphosphine using a catalyst $MoO_2Cl_2(dmf)_2$. Compound **7a** was prepared from the intermediate **6a** by Mannich reaction. Finally, the target compounds **8a** was prepared by treatment of **7a** with imidazole in hot xylene.

*4.1.1.1. 4-Fluoro-2-nitrobenzaldehyde (**2a**)*

To a stirred solution containing 4-fluro-2-nitrotoluene (**1a**) (1.0 g, 6.45 mmol) in 15 mL of anhydrous N,N-dimethylformamide, was added pyrrolidine (0.69 g, 9.70 mmol) and N,N-dimethylformamide dimethyl acetal (1.54 g, 12.92 mmol). The resulting reaction mixture was allowed to stir at 125 □ for 6 hours and then poured into 20 mL of water. The product was extracted with ethyl acetate three times (20 mL×3). The combined organic layers were washed successively with 30 mL of water and 30 mL of brine, then dried by anhydrous sodium sulfate and concentrated under reduced pressure to provide dark red oil. The oil was dissolved in 20 mL of THF/water (1:1, v/v) and the resultant solution was treated with sodium periodate (4.13 g, 19.3 mmol) at 0 □. The resulting reaction mixture was allowed to stir at room temperature overnight and was filtered through Celite. The filtrate was diluted with 10 mL of water and extracted by ethyl acetate three times (15 mL×3). The combined organic layers were washed successively with 25 mL of water and 25 mL of brine, then dried by anhydrous sodium sulfate and concentrated under reduced pressure to provide dark oil that was further purified by column chromatography on silica gel (6:1 petroleum ether/ethyl acetate) to yield 4-fluoro-2-nitrobenzaldehyde (**2a**) (0.69 g, 4.08 mmol, 63%) as a yellow solid [23].

*4.1.1.2. (4-Flurobenzyl)triphenylphosphonium bromide (**4a**)*

To a stirred solution containing triphenylphosphine (1.66 g, 6.33 mmol) in 15 mL of toluene, 4-fluorobenzyl bromide (**3a**) (1.0 g, 5.29 mmol) was added and the resulting mixture was allowed to stir at 120 □ for 6 hours, then cooled to room temperature. The precipitate was filtered off, washed with 10 ml of organic solvent (4:1 petroleum ether/ethyl acetate), and dried by vaccum to yield (4-flurobenzyl) triphenylphosphonium bromide (**4a**) (2.34 g, 5.18 mmol, 98 %) as a white solid [24].

*4.1.1.3. 4-Fluoro-2-nitro-4'-fluoro stilbene (**5a**)*

4-Fluoro-2-nitrobenzaldehyde (**2a**) (1.0 g, 5.91 mmol) and (4-Flurobenzyl) triphenylphosphonium bromide (**4a**) (2.94 g, 6.51 mmol) were dissolved in 20 mL of THF. DBU (1.08 g, 7.10 mmol) was added to the solution. The resulting reaction mixture was refluxed for 4 hours. Then the solvent was evaporated and the residue was dissolved in 30 mL of ethyl acetate.

The organic layer was washed by brine, then dried by anhydrous sodium sulfate and concentrated under reduced pressure to provide crude 4-fluoro-2-nitro-4'-fluoro stilbene (**5a**) (1.45 g, 5.55 mmol, 94%) as a brown oil [24].

*4.1.1.4. 6-Fluoro-2-(4-fluorophenyl)-1H-indole (**6a**)*

To a stirring solution containing 4-fluoro-2-nitro-4'-fluoro stilbene (**5a**) (1.45 g, 5.55 mmol) in 25 mL of toluene, triphenylphosphine (3.88 g, 14.8 mmol) and $MoO_2Cl_2(dmf)_2$ (0.20 g, 0.58 mmol) were added successively. The resulted solution was refluxed for 16 hours under nitrogen atmosphere. The solvent was evaporated and the residue was purified by column chromatography on silica gel (10:1 petroleum ether/ethyl acetate) to yield 6-fluoro-2-(4-fluorophenyl)-1*H*-indole (**6a**) (0.84 g, 3.66 mmol, 66%) as a white solid [25].

*4.1.1.5. 3-Dimethylaminomethyl-6-Fluoro-2-(4-fluorophenyl)-1H-indole (**7a**)*

6-Fluoro-2-(4-fluorophenyl)-1*H*-indole (**6a**) (1.0 g, 4.36 mmol) and (methylene)- dimethyl ammonium chloride (0.82 g, 8.76 mmol) were dissolved in 15 mL of dry dichloromethane. The resulting reaction mixture was stirred for 8 hours at room temperature, and then treated with 10 ml of saturated sodium carbonate solution. The organic layer was separated and washed with water two times (10×2), then dried by anhydrous sodium sulfate and concentrated under reduced pressure to provide 3-dimethylaminomethyl-6-Fluoro-2-(4-fluorophenyl)-1*H*-indole (**7a**) (1.15 g, 4.01 mmol, 92 %) as a white solid [26].

*4.1.1.6. 3-((1H-imidazol-1-yl)methyl)-6-fluoro-2-(4-fluorophenyl)-1H-indole (**8a**)*

To a solution containing 3-dimethylaminomethyl-6-Fluoro-2-(4-fluorophenyl)-1*H*-indole (**7a**) (1.0 g, 3.49 mmol) in 20 ml of xylene was added imidazole (1.19 g, 17.45 mmol). The reaction mixture was allowed to stir at 130℃ for 1 hour. Then xylene was removed under reduced pressure and the residue was dissolved by 30 ml of ethyl acetate. The organic layer was washed with water three times (10 ×3), then dried by anhydrous sodium sulfate and concentrated under vaccum. The residue was recrystalized from ethyl acetate to afford 3-((1*H*-imidazol-1-yl)methyl)-6-fluoro-2-(4-methoxyphenyl)-1*H*-indole (**8a**) (0.50 g, 1.60 mmol) [27]: Yield 46 %; White solid; Mp 211-220 ^{o}C; IR (KBr) (ν_{max}, cm^{-1}): 3144, 1502, 1226, 806, 744; ^{1}H NMR (300 MHz, DMSO-*d6*): δ 11.75 (s, 1H), 7.66-7.62 (d, J = 10.8 Hz, 3H), 7.48-7.35 (m, 3H), 7.19-7.16 (m, 1H), 6.95-6.84 (m, 3H), 5.40 (s, 2H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 163.6-160.8 (d, J = 210.0 Hz), 160.4-157.7 (d, J = 202.5

Hz), 136.8, 136.5, 135.9, 130.3, 128.5, 128.1, 124.7, 119.6, 119.0, 116.1, 115.8, 108.3, 106.6, 97.6, 97.3, 40.7; HRMS (ESI) *m/z* calcd for $C_{18}H_{14}N_3F_2^+$ $[M+H]^+$, 310.1150; found, 310.1150.

The following compounds (**8b**-**t**) were prepared using the general method described for the synthesis of compound **8a**.

4.1.2. 3-((1*H*-imidazol-1-yl)methyl)-2-(4-chlorophenyl)-6-fluoro-1*H*-indole (**8b**): Yield 42 %; White solid; Mp 215-223 ^oC; IR (KBr) (ν_{max}, cm^{-1}): 3327, 1625, 1221, 1080, 804, 731; ^1H NMR (300 MHz, DMSO-*d6*): δ 11.79 (s, 1H), 7.63-7.61 (m, 5H), 7.49-7.44 (m, 1H), 7.20-7.17 (d, J = 9.0 Hz, 1H), 6.96-6.85 (m, 3H), 5.42 (s, 2H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 160.9, 157.8, 136.8, 136.0, 133.0, 130.4, 129.8, 129.0, 128.5, 124.7, 119.7, 119.0, 108.4, 108.1, 107.1, 97.7, 97.3, 40.7; HRMS (ESI) *m/z* calcd for $C_{18}H_{14}N_3FCl^+$ $[M+H]^+$, 326.0854; found, 326.0858.

4.1.3. 3-((1*H*-imidazol-1-yl)methyl)-6-chloro-2-(4-fluorophenyl)-1*H*-indole (**8c**): Yield 42 %; White solid; Mp 233-238 ^oC; IR (KBr) (ν_{max}, cm^{-1}): 3005, 1672, 1499, 1227, 836, 749; ^1H NMR (300 MHz, DMSO-*d6*): δ 11.81 (s, 1H), 7.67-7.61 (m, 3H), 7.48-7.36 (m, 4H), 7.07-7.04 (m, 1H), 6.94 (s, 1H), 6.84 (s, 1H), 5.40 (s, 2H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 163.7, 160.5, 136.8, 136.2, 130.5, 130.5,130.4, 128.5, 127.9, 126.7, 120.0, 119.7, 119.0, 116.2, 115.9, 111.0, 106.7, 40.6; HRMS (ESI) *m/z* calcd for $C_{18}H_{14}N_3FCl^+$ $[M+H]^+$, 326.0854; found, 326.0853.

4.1.4. 3-((1*H*-imidazol-1-yl)methyl)-6-chloro-2-(4-chlorophenyl)-1*H*-indole (**8d**): Yield 44 %; White solid; Mp 232-244 ^oC; IR (KBr) (ν_{max}, cm^{-1}): 3024, 1509, 1229, 830; ^1H NMR (300 MHz, DMSO-*d6*): δ 11.82 (s, 1H), 7.62 (m, 5H), 7.48-7.43 (m, 2H), 7.07-7.04 (m, 1H), 6.94 (s, 1H), 6.83 (s, 1H), 5.42 (s, 2H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 136.8, 136.4, 136.3, 133.2, 130.2, 130.2, 129.9, 129.0, 129.0, 128.5, 126.9, 126.7, 120.0, 119.9, 119.0, 111.0, 107.2, 40.5; HRMS (ESI) *m/z* calcd for $C_{18}H_{12}N_2O_2FCl^+$ $[M+H]^+$, 342.0565; found, 342.0563.

4.1.5. 3-((1*H*-imidazol-1-yl)methyl)-6-fluoro-2-(4-methoxyphenyl)-1*H*-indole (**8e**): Yield 53 %; yellow solid; Mp 210-213 ^oC; IR (KBr) (ν_{max}, cm^{-1}): 3123, 1617, 1410, 1031, 829, 590; ^1H NMR (300 MHz, DMSO-*d6*): δ 11.65 (s, 1H), 7.62 (s, 1H), 7.55-7.52 (d, 2H), 7.46-7.42 (dd, 1H), 7.18-7.15 (m, 1H), 7.11-7.09 (d, 2H), 6.92 (s, 1H), 6.91-6.89 (d, 1H), 6.85 (s, 1H), 5.39 (s, 2H), 3.81 (s, 3H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 160.7-157.6 (d, *J* = 232.5 Hz), 159.3, 137.6, 136.8, 135.8-

135.6 (d, *J* = 15 Hz), 129.4, 128.5, 124.9, 124.0, 119.2-119.1 (d, *J* = 7.5 Hz), 119.0, 114.5, 108.1-107.8 (d, *J* = 22.5 Hz), 105.6, 97.6-97.2 (d, *J* = 30 Hz), 55.3, 40.9; HRMS (ESI) *m/z* calcd for $C_{19}H_{17}FN_3O^+$ $[M+H]^+$, 322.1350; found, 322.1352.

4.1.6. 3-((1*H*-imidazol-1-yl)methyl)-6-chloro-2-(4-methoxyphenyl)-1*H*-indole (**8f**): Yield 48 %; yellow solid; Mp 213-221 ^oC; IR (KBr) (ν_{max}, cm^{-1}): 3115, 1616, 1504, 1459, 1252, 837, 632; ^1H NMR (300 MHz, DMSO-*d6*): δ 11.70 (s, 1H), 7.61 (s, 1H), 7.55-7.47 (d, 2H), 7.44-7.41 (dd, 1H), 7.12-7.09 (m, 1H), 7.05-7.05 (d, 2H), 7.02-6.94 (d, 1H), 6.84 (s, 1H), 5.39 (s, 1H), 3.82 (s, 3H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 159.4, 137.9, 136.7, 136.1-136.0 (d, *J* = 7.5 Hz), 129.5, 128.5, 126.9, 126.3, 123.7, 119,8, 119,4, 118,9, 114,5, 110.9, 105.7, 55.3, 40.7; HRMS (ESI) *m/z* calcd for $C_{19}H_{17}N_3OCl^+$ $[M+H]^+$, 338.1054; found, 338.1052.

4.1.7. 3-((1*H*-imidazol-1-yl)methyl)-2-(4-fluorophenyl)-6-methoxy-1*H*-indole (**8g**): Yield 32 %; yellow solid; Mp 208-220 ^oC; IR (KBr) (ν_{max}, cm^{-1}): 3119, 1632, 1499, 1155, 1068, 857, 663; ^1H NMR (300 MHz, DMSO-*d6*): δ 11.45 (s, 1H), 7.63-7.59 (m, 3H), 7.40-7.32 (m, 3H), 6.96 (s, 1H), 6.91-6.91 (dd, 1H), 6.85 (s, 1H), 6.72-6.69 (m, 1H), 5.37 (s, 2H), 3.78 (s, 3H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 163.5-162.3 (d, *J* = 90 Hz), 156.3, 136.8, 134.5, 130.1, 129.9, 128.4, 122.3, 119.0, 116.1, 115.8, 109.9, 106.4, 94.5, 55.2, 40.9; HRMS (ESI) *m/z* calcd for $C_{19}H_{17}N_3OF^+$ $[M+H]^+$, 322.1350; found, 322.1354.

4.1.8. 3-((1*H*-imidazol-1-yl)methyl)-2-(4-chlorophenyl)-6-methoxy-1*H*-indole (**8h**): Yield 55 %; White solid; Mp 224-230 ^oC; IR (KBr) (ν_{max}, cm^{-1}): 3453, 3055, 1631, 1459, 1158, 1068, 837, 570; ^1H NMR (300 MHz, DMSO-*d6*): δ 11.49 (s, 1H), 7.62 (s, 1H), 7.58 (s, 1H), 7.35-7.33 (d, 3H), 6.96, 6.91-6.90, 6.85, 6.72-6.69 (m, 1H), 5.39 (s, 1H), 3.78 (s, 3H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 156.4, 136.8, 136.8, 134.0, 132.4, 130.9, 129.5, 129.0, 128.4, 122.3, 119.1, 119.0, 110.0, 106.9, 94.4, 55.2, 40.9; HRMS (ESI) *m/z* calcd for $C_{19}H_{17}N_3OCl^+$ $[M+H]^+$, 338.1054; found, 338.1058.

4.1.9. 7-((1*H*-imidazol-1-yl)methyl)-6-(4-fluorophenyl)-5*H*-[1,3]dioxolo[4,5-f]indole (**8i**): Yield 56 %; White solid; Mp 232-238 ^oC; IR (KBr) (ν_{max}, cm^{-1}): 3152, 1551, 1474, 1349, 1210, 840, 660; ^1H NMR (300 MHz, DMSO-*d6*): δ 11.45 (s, 1H), 7.60-7.56 (t, 3H), 7.37-7.31 (m, 2H), 6.96-6.92 (d, 3H), 6.84 (s, 1H),5.95 (1H), 5.54 (s, 2H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 163.3, 160.0, 144.6, 142.6, 136.7, 134.3, 130.8, 129.8, 129.7, 128.6, 128.4, 121.9, 118.9, 116.0, 115.7, 106.8, 100.3, 96.9, 92.1, 40.9; HRMS (ESI) *m/z* calcd for $C_{19}H_{15}N_3O_2F^+$ $[M+H]^+$, 336.1142; found, 336.1144.

4.1.10. 7-((1*H*-imidazol-1-yl)methyl)-6-(4-chlorophenyl)-5*H*-[1,3]dioxolo[4,5-f]indole (**8j**): Yield 51 %; White solid; Mp 244-254 ^{o}C; IR (KBr) (ν_{max}, cm^{-1}): 3207, 1466, 1306, 1194, 1074, 825; ^{1}H NMR (300 MHz, DMSO-*d6*): δ 11.49 (s, 1H), 7.61-7.56 (d, 5H), 6.96-6.92 (t, 3H), 6.84 (s, 1H), 5.95 (s, 2H), 5.36 (s, 2H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 144.8, 142.7, 136.7, 133.8, 132.2, 131.0, 130.9, 129.3, 128.9, 128.4, 122.0, 118.9, 107.3, 100.4, 97.0, 92.1, 40.8; HRMS (ESI) *m/z* calcd for $C_{19}H_{15}N_3O_2Cl^+$ $[M+H]^+$, 352.0847; found, 352.0847.

4.1.11. 3-((1*H*-1,2,4-triazol-1-yl)methyl)-6-fluoro-2-(4-fluorophenyl)-1*H*-indole (**8k**): Yield 40 %; White solid; Mp 218-224 ^{o}C; IR (KBr) (ν_{max}, cm^{-1}): 3251, 1504, 1225, 1016, 838, 699; ^{1}H NMR (300 MHz, DMSO-*d6*): δ 11.72 (s, 1H), 8.73 (s, 1H), 7.92-7.89 (m, 1H), 7.64-7.59 (t, 2H), 7.44-7.38 (m, 1H), 7.19-7.15 (m, 1H), 6.95-6.89 (m, 2H), 5.58 (s, 2H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 163.8, 160.9, 160.5, 157.8, 151.5, 144.0, 136.9, 135.9, 135.7, 135.6, 130.6, 130.5, 128.9, 124.6, 119.7, 119.6, 116.1, 115.8, 108.3, 108.0, 105.7, 97.7, 97.3, 43.8; HRMS (ESI) *m/z* calcd for $C_{17}H_{12}N_4F_2Na^+$ $[M+Na]^+$, 333.0922; found, 333.0919.

4.1.12. 3-((1*H*-1,2,4-triazol-1-yl)methyl)-2-(4-chlorophenyl)-6-fluoro-1*H*-indole (**8l**): Yield 40 %; White solid; Mp 247-254 ^{o}C; IR (KBr) (ν_{max}, cm^{-1}): 3215, 1627, 1131, 830, 679; ^{1}H NMR (300 MHz, DMSO-*d6*): δ 11.76 (s, 1H), 8.74 (s, 1H), 7.97-7.88 (m, 3H), 7.64-7.61 (d, 2H), 6.96-6.90 (t, 1H), 5.60 (s, 2H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 160.9, 157.8, 151.5, 144.0, 136.5, 136.0, 135.8, 133.1, 130.3, 130.0, 128.9, 124.6, 119.8, 119.7, 108.4, 108.0, 106.1, 97.7, 97.3, 43.7; HRMS (ESI) *m/z* calcd for $C_{17}H_{12}N_4FNaCl^+$ $[M+Na]^+$, 349.0626; found, 349.0627.

4.1.13. 3-((1*H*-1,2,4-triazol-1-yl)methyl)-6-chloro-2-(4-fluorophenyl)-1*H*-indole (**8m**): Yield 49 %; White solid; Mp 198-205 ^{o}C; IR (KBr) (ν_{max}, cm^{-1}): 3230, 1504, 1135, 840, 512; ^{1}H NMR (300 MHz, DMSO-*d6*): δ 11.77 (s, 1H), 8.72 (s, 1H), 7.97 (s, 1H), 7.94-7.90 (m, 3H), 7.65-7.62 (d, 1H), 7.45-7.39 (m, 3H), 7.09-7.06 (d, 1H), 5.58 (s, 2H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 163.8, 160.6, 151.5, 144.0, 137.3, 136.2, 130.7, 130.6, 127.8, 126.7, 126.6, 119.9, 116.1, 115.8, 111.0, 105.7, 43.7; HRMS (ESI) *m/z* calcd for $C_{17}H_{12}N_4FNaCl^+$ $[M+Na]^+$, 349.0626; found, 349.0626.

4.1.14. 3-((1*H*-1,2,4-triazol-1-yl)methyl)-6-chloro-2-(4-chlorophenyl)-1*H*-indole (**8n**): Yield 40 %; White solid; Mp 229-234 ^{o}C; IR (KBr) (ν_{max}, cm^{-1}): 3230, 1431, 1156, 1011, 856; ^{1}H NMR (300

MHz, DMSO-*d6*): δ 11.81 (s, 1H), 8.73 (s, 1H), 7.97 (s, 1H), 7.91-7.88 (d, 2H), 7.65-7.62 (d, 3H), 7.44-7.43 (d, 1H), 7.09-7.06 (m, 1H), 5.60 (s, 2H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 151.5, 144.0, 136.9, 136.3, 133.3, 130.1, 129.0, 126.9, 126.6, 120.0, 111.0, 106.2, 43.6; HRMS (ESI) *m/z* calcd for $C_{17}H_{12}N_4NaCl_2^+$ $[M+Na]^+$, 365.0331; found, 365.0333.

4.1.15. 3-((1*H*-1,2,4-triazol-1-yl)methyl)-6-fluoro-2-(4-methoxyphenyl)-1*H*-indole (**8o**): Yield 45 %; yellow solid; Mp 206-210 ^{o}C; IR (KBr) (ν_{max}, cm^{-1}): 3150, 1615, 1460, 1139, 842, 676; ^{1}H NMR (300 MHz, DMSO-*d6*): δ 11.60 (s, 1H), 8.69 (s, 1H), 7.98 (s, 1H), 7.81-7.78 (d, 2H), 7.61-7.57 (dd, 1H), 7.17-7.12 (m, 1H), 6.94-6.88 (d, 2H), 5.57 (s, 1H), 3.82 (s, 3H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 160.7-157.5 (d, *J* = 240 Hz), 159.4, 151.4, 143.9, 137.9, 135.8-135.6 (d, *J* = 15 Hz), 129.7, 124.8, 123.9, 119.4-119.2 (d, *J* = 15 Hz), 114.4, 108.0-107.7 (d, *J* = 22.5 Hz), 104.6, 97.5-97.2 (d, *J* = 22.5 Hz), 55.2, 44.0; HRMS (ESI) *m/z* calcd for $C_{18}H_{15}N_4OFNa^+$ $[M+Na]^+$, 345.1127; found, 345.1126.

4.1.16. 3-((1*H*-1,2,4-triazol-1-yl)methyl)-6-chloro-2-(4-methoxyphenyl)-1*H*-indole (**8p**): Yield 45 %; yellow solid; Mp 211-216 ^{o}C; IR (KBr) (ν_{max}, cm^{-1}): 3126, 1615, 1505, 1136, 1030, 841, 677; ^{1}H NMR (300 MHz, DMSO-*d6*): δ 11.66 (s, 1H), 8.68 (s, 1H), 7.97 (s, 1H), 7.81-7.78 (d, *J* = 9 Hz, 2H), 7.61-7.58 (d, *J* = 9 Hz, 1H), 7.41-7.41 (d, 1H), 7.15-7.12 (d, *J* = 9 Hz, 2H), 7.07-7.03 (d, *J* = 12 Hz, 1H), 5.57 (s, 2H), 3.83 (s, 3H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 159.5, 151.5, 143.9, 138.3, 136.1, 129.8, 126.8, 126.4, 123.6, 119.7-119.6 (d, *J* = 7.5 Hz), 114.5, 110.8, 104.8, 55.3, 43.9; HRMS (ESI) *m/z* calcd for $C_{18}H_{15}N_4ONaCl^+$ $[M+Na]^+$, 361.0826; found, 361.0824.

4.1.17. 3-((1*H*-1,2,4-triazol-1-yl)methyl)-2-(4-fluorophenyl)-6-methoxy-1*H*-indole (**8q**): Yield 25 %; yellow solid; Mp 178-180 ^{o}C; IR (KBr) (ν_{max}, cm^{-1}): 3126, 1632, 1504, 1160, 1028, 843, 677; ^{1}H NMR (300 MHz, DMSO-*d6*): δ 11.42 (s, 1H), 8.69 (s, 1H), 7.97 (s, 1H), 7.91-7.86 (d, 5H), 7.50-7.47 (d, 1H), 7.43 (s, 1H), 7.40-7.38 (t, 3H), 6.91-6.90 (d, H), 6.73-6.70 (m, 1H), 5.55 (s, 2H), 3.78 (s, 2H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 163.5, 160.3, 156.3, 151.4, 143.9, 136.7, 135.0, 130.3, 130.2, 128.5, 122.1, 119.2, 116.0, 115.7, 109.8, 105.4, 94.5, 55.3, 44.0; HRMS (ESI) *m/z* calcd for $C_{18}H_{15}N_4OFNa^+$ $[M+Na]^+$, 345.1127; found, 345.1122.

4.1.18. 3-((1*H*-1,2,4-triazol-1-yl)methyl)-2-(4-chlorophenyl)-6-methoxy-1*H*-indole (**8r**): Yield 38 %; White solid; Mp 203-207 ^{o}C; IR (KBr) (ν_{max}, cm^{-1}): 3425, 3128, 1633, 1457, 1161, 1029, 837,

731; ^{1}H NMR (300 MHz, DMSO-*d6*): δ 11.46 (s, 1H), 8.70 (s, 1H), 7.97 (s, 1H), 7.88-7.85 (d, 2H), 7.62-7.48 (q, 1H), 6.90-6.89 (d, 1H), 6.73-6.70 (m, 2H), 5.56 (s, 2H), 3.78 (s, 3H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 156.4, 151.4, 143.9, 136.8, 134.5, 132.6, 130.8, 129.8, 128.9, 122.1, 119.3, 109.9, 105.9, 94.4, 55.2, 43.9; HRMS (ESI) *m/z* calcd for $C_{18}H_{15}N_4ONaCl^+$ $[M+Na]^+$, 361.0832; found, 361.0829.

4.1.19. 7-((1*H*-1,2,4-triazol-1-yl)methyl)-6-(4-fluorophenyl)-5*H*-[1,3]dioxolo[4,5-f]indole (**8s**): Yield 46 %; White solid; Mp 245-248 ^{o}C; IR (KBr) (ν_{max}, cm^{-1}): 3174, 1502, 1471, 1293, 1135, 843, 672; ^{1}H NMR (300 MHz, DMSO-*d6*): δ 11.41 (s, 1H), 8.71 (s, 1H), 7.97 (s, 1H), 7.89-7.84 (t, 2H), 7.40-7.34 (d, 2H), 7.13 (s, 1H), 6.91 (s, 1H), 5.95 (s, 1H), 5.51 (s, 2H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 163.4, 160.2, 151.4, 144.6, 143.9, 142.6, 134.8, 134.6, 130.8, 130.7, 130.1, 130.0, 128.5, 121.8, 115.9, 115.6, 105.9, 100.4, 97.2, 92.2, 44.0; HRMS (ESI) *m/z* calcd for $C_{18}H_{13}N_4O_2FNa^+$ $[M+Na]^+$, 359.0914; found, 359.0913.

4.1.20. 7-((1*H*-1,2,4-triazol-1-yl)methyl)-6-(4-chlorophenyl)-5*H*-[1,3]dioxolo[4,5-f]indole (**8t**): Yield 49 %; White solid; Mp 257-261 ^{o}C; IR (KBr) (ν_{max}, cm^{-1}): 3168, 1634, 1472, 1041, 676; ^{1}H NMR (300 MHz, DMSO-*d6*): δ 11.46 (s, 1H), 8.73 (s, 1H), 7.97 (s, 1H), 7.86-7.83 (t, 3H), 7.60-7.57 (d, 1H), 7.14 (s, 1H), 6.92 (s, 1H) 5.96 (s, 1H), 5.52 (s, 3H); ^{13}C NMR (75 MHz, DMSO-*d6*): δ 151.4, 144.8, 144.0, 142.7, 134.3, 132.4, 131.0, 130.8, 129.6, 128.9, 121.9, 106,4, 100.4, 97.2, 92.2, 43.9; HRMS (ESI) *m/z* calcd for $C_{18}H_{13}N_4O_2NaCl^+$ $[M+Na]^+$, 375.0619; found, 375.0623.

4.2. Biological evaluation

4.2.1 Placental sample collection and microsome preparation

Studies were performed in accordance with the code of Ethics of the World Medical Association and the ethical standards of the relevant national and institutional committees on human investigation. Subjects were free of diseases and other substance abuse. Placenta samples were collected and dissected immediately following delivery. Specimens were quickly washed with cold PBS to eliminate contaminating blood. Aliquots of 103 g of placental samples were pulverized and homogenized in TES solution (20mM TES, 1mM EDTA, 0.15M KCl, 0.1mM PMSF, pH 7.4) on ice, and samples were centrifuged at 10,000 g for 25 min. The supernatant was centrifuged at 100,000 g for 75 min at 4 °C, and the aliquot was re-suspended in TES buffer containing 20 mM

TES, 1 mM EDTA, 0.15 M KC and 0.25 M sucrose (pH 7.4). A Bradford assay (Bio-Rad) was used for sample protein quantification, and samples was stored at −80 °C.

4.2.2 Aromatase inhibition assay

Aromatase inhibition assays were performed using the Estrone ELISA kit (BioVendor, Brno, Czech Republic) androstenedione, NADPH and human aromatase with letrozole as positive control [28]. The 96-well black plate was used to determine the activity of the compounds. Solutions of the substrate and inhibitors were prepared and the ELISA procedures were performed as described in the kits manuals. Quantification of samples was performed by applying the linear regression equation of the standard curve to the absorbance response. Each compound was tested in triplicate measurements and the average IC_{50} value can be seen in Table 1.

4.2.3. Cytotoxicity assay

MCF-7 cells (obtained from Kunming Institute of Zoology. CAS) were cultured at 37°C in a 5 % CO_2 humidified atmosphere in DMEM medium supplemented with 10 % FBS, 1 % penicillin-streptomycin. Cells were collected using 0.25% trypsin and seeded in 96 well plates at a density of 4,000 cells/well, then incubated at 37 °C overnight. After 48 h incubation with various concentrations of tested compounds, the medium containing compounds was removed and fresh medium containing 10 μL of MTT (5 mg/ml) was added to each well, then the resultant solution was incubated for another 4 h. The MTT medium was discarded before adding 150 μL of DMSO to each well. The optical densities at 490 nm were measured by an ELISA microplate reader (Multiskan Go, Thermo Fisher Scientific, Gemany). The IC_{50}/MCF-7 values were calculated using GraphPad Prism 6 software.

4.3. Molecular modeling studies

Quantum mechanics (QM) calculations were performed on Gaussion09 package to scan the potential energy profile of letrozole (Fig.S1-B) and the potential energy surface of **8o** (Fig.S2-B) to determine their minimum-energy structures in solution. The electrostatic surface potential (ESP) of letrozole and **8o** was calculated by using the DFT B3LYP method with the 6-31G(d) basis set, and the partial atomic charges of letrozole and **8o** was fit by performing two-stage restrained

electrostatic potential (RESP) charge-fitting protocol [29]. The general amber force field (GAFF) was employed to describe the atoms on letrozole and **8o** [30].

The crystal structure of aromatase in complex with androstenedione is freely available in Protein Data Bank (PDB entry: 3EQM). We removed the androstenedione and phosphate ions from the original PDB file, and reserved the remaining 452 amino acids and an iron porphyrin for subsequent computational study. AMBER 14SB force field is used to describe the standard amino acids [31]. Iron porphyrin, an organometallic compound, is consisting of an Fe ion contained in the centre of porphyrin (Fig.S3). To build the force field for the simulations of iron porphyrin with aromatase inhibitors, we followed the Metal Center Parameter Builder (MCPB) parametrization scheme to derive force constants and charge parameters for iron porphyrin [32].

Docking studies of letrozole and compound **8o** on aromatase were carried out via AutoDock Vina program. The energy-minimum structures of letrozole and **8o** obtained from the QM calculations (Figs.S1-C and S2-C) were set as a ligand, and the aromatase was treated as a rigid receptor. The Gasteiger charges were used in our molecular docking studies. The dimensionality of docking grid box was fixed at *size_x*=40 Å, *size_y*=40 Å and *size_z*=40 Å. By changing the positions of box center, we conducted multiple independent docking attempts to search for possible binding sites and determine all binding modes for the two letrozole-aromatase and **8o**-aromatase systems.

All predicted docking poses of letrozole and compound **8o** on the aromatase obtained from the docking studies need to undergo further evaluations in atomistic molecular dynamics (MD) simulations. We solvated the aromatase complexed with letrozole and **8o** into a periodic truncated octahedral box filled with TIP3P water molecules. Chloride counterions were added to neutralize the system. Details of our simulation procedures are given below. (1) A 10000-step energy minimization of solvent was performed using steepest descent method with the letrozole-aromatase and **8o**-aromatase constrained by a force of 200 kcal/mol. (2) The system was heated over 10 ps at 298 K to relax the water molecules. A short 40-ps isothermal-isobaric (NPT) ensemble MD simulation was carried out with the letrozole-aromatase and **8o**-aromatase constrained by a force of 200 kcal/mol. (3) Another 10000-step energy minimization was performed with a restraint of 25 kcal/mol to the letrozole-aromatase and **8o**-aromatase, after which a 20-ps NPT MD simulation was run with a restraint of 25 kcal/mol to the letrozole-aromatase and **8o**-aromatase. (4) A 10000-step unconstrained energy minimization was conducted to all the atoms in the system, after which the system was reheated over 40 ps at a constant volume to 298 K. (5) Production simulations were

performed in 50 ns in a canonical (NVT) ensemble on the AMBER 12 package. Berendsen thermostat was used to maintain the simulation temperature at 298 K. We used partial mesh Ewarld summation to calculate long-ranged electrostatic interactions with a 12 Å cut-off distance and 1E-5 tolerance. The SHAKE algorithm was used to constrain the motions of all bonds involving hydrogen atom. The calculation of the root mean square deviation (RMSD) was restrained to fit the heavy atoms on backbone chains of the aromatase and the non-hydrogen atoms on iron porphyrin, letrozole and **8o**. The RMSD *vs.* time was examined over the entire simulation to ascertain whether the binding process reached a stable state or not.

Acknowledgments

Financial support for this work was provided by the National Natural Science Foundation of China (No. 81101688). We would like to thank NC State University High Performance Computing Center for providing us computing resources, and Dr. Lihui Peng for assistance on placental sample collection.

Appendix A. Supporting information

Detailed descriptions of computational protocol, and ^{1}H and ^{13}C NMR Spectra of compounds **8a-t** are given in the Supporting information.

References

[1] E.O. Fourkala, O. Blyuss, H. Field, R. Gunu, A. Ryan, J. Barth, I. Jacobs, A. Zaikin, A. Dawnay, U. Menon, Sex hormone measurements using mass spectrometry and sensitive extraction radioimmunoassay and risk of estrogen receptor negative and positive breast cancer: Case control study in UK Collaborative Cancer Trial of Ovarian Cancer Screening (UKCTOCS), Steroids 110 (2016) 62-69.

[2] Y. Omoto, H. Iwase, Clinical significance of estrogen receptor β in breast and prostate cancer from biological aspects, Cancer Sci. 106 (2015) 337-343.

[3] E.O. Fourkala, A. Zaikin, M. Burnell, A. Gentry-Maharaj, J. Ford, R. Gunu, Association of serum sex steroid receptor bioactivity and sex steroid hormones, Endocr. Relat. Cancer 19 (2012) 137-147.

[4] J. Zhang, G. Li, Z. Li, X. Yu, Y. Zheng, K. Jin, H. Wang, Y. Gong, X. Sun, X. Teng, J. Cao, L. Teng, Estrogen-independent effects of ER-α36 in ER-negative breast cancer, Steroids 77 (2012) 666-673.

[5] R.W. Brueggemeier, J.C. Hackett, E.S. Diaz-Cruz, Aromatase inhibitors in the treatment of breast cancer, Endocr. Rev. 26 (2005) 331-345.

[6] N. Adhikari, S.A. Amin, A. Saha, T. Jha, Combating breast cancer with non-steroidal aromatase inhibitors (NSAIs): Understanding the chemico-biological interactions through comparative SAR/QSAR study, Eur. J. Med. Chem. 137 (2017) 365-438.

[7] R. Riemsma, C.A. Forbes, A. Kessels, K. Lykopoulos, M.M. Amonkar, D.W. Rea, J. Kleijnen, Systematic review of aromatase inhibitors in the first-line treatment for hormone sensitive advanced or metastatic breast cancer, Breast Cancer Res. Treat. 123 (2010) 9-24.

[8] G.W. Sledge, E.P. Mamounas, G.N. Hortobagyi, H.J. Burstein, P.J. Goodwin, A.C. Wolff, Past, present, and future challenges in breast cancer treatment, J. Clin. Oncol. 32 (2014) 1979-1986.

[9] C.P. Miller, M.D. Collini, B.D. Tran, H.A. Harris, Y.P. Kharode, J.T. Marzolf, R.A. Moran, R.A. Henderson, R.H.W. Bender, R.J. Unwalla, L.M. Greenberger, J.P. Yardley, M.A. Abou-Gharbia, C.R. Lyttle, B.S. Komm, Design, synthesis, and preclinical characterization of novel, highly selective indole estrogens, J. Med. Chem. 44 (2001) 1654-1657.

[10] E.V. Angerer, J. Prekajac, J. Strohmeier, 2-Phenylindoles. Relationship between structure, estrogen receptor affinity, and mammary tumor inhibiting activity in the rat, J. Med. Chem. 27 (1984) 1439-1447.

[11] W. Lv, J. Liu, D. Lu, D.A. Flockhart, M. Cushman, Synthesis of mixed (E,Z)-, (E)-,and (Z)-Norendoxifen with dual aromatase inhibitory and estrogen receptor modulatory activities, J. Med. Chem. 56 (2013) 4611-4618.

[12] S.P. Robinson, R. Koch, V.C. Jordan, In vitro estrogenic actions in rat and human cells of hydroxylated derivatives of D16726 (zindoxifene), an agent with known antimammary cancer activity in vivo, Cancer Res. 48 (1988) 784-787.

[13] M.A. Neves, T.C. Dinis, G. Colombo, M.L. Sá e Melo, Fast Three Dimensional Pharmacophore Virtual Screening of New Potent Non-Steroid Aromatase Inhibitors, J. Med. Chem. 52 (2009) 143-150.

[14] M.M. Ghorab, M.S. Alsaid, N. Samir, G.A. Abdel-Latif, A.M. Soliman, F.A. Ragab, D.A. Abou El Ella DA. Aromatase inhibitors and apoptotic inducers: Design, synthesis, anticancer activity and molecular modeling studies of novel phenothiazine derivatives carrying sulfonamide moiety as hybrid molecules, Eur. J. Med. Chem. 134 (2017) 304-315.

[15] X. Xiao, B. Zhao, L. Yang, X. Liang, Y. Ren, Probe the binding mode of aristololactam-β-D-glucoside to phenylalanine transfer RNA in silico, ChemistrySelect 1 (2016) 5430-5439.

[16] H. Gohlke, C. Kiel, D.A. Case, Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes, J. Mol. Biol. 330 (2003) 891-913.

[17] K. Zhu, M.R. Shirts, R.A. Friesner, Improved methods for side chain and loop predictions via the protein local optimization program: variable dielectric model for implicitly improving the treatment of polarization effects, J. Chem. Theory Comput. 3 (2007) 2108-2119.

[18] D. Ghosh, J. Griswold, M. Erman, W. Pangborn, Structural basis for androgen specificity and estrogen synthesis in human aromatase, Nature 457 (2009) 219-223.

[19] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2010) 455-461.

[20] J. Shao, S.W. Tanner, N. Thompson, T.E. Cheatham, Clustering molecular dynamics trajectories: 1. characterizing the performance of different clustering algorithms, J. Chem. Theory Comput. 3 (2007) 2312-2334.

[21] X. Xiao, P.F. Agris, C.K. Hall, Molecular recognition mechanism of peptide chain bound to the $tRNA^{Lys3}$ anticodon loop *in silico*, J. Biomol. Struct. Dyn. 33 (2015) 14-27.

[22] X. Xiao, B. Zhao, P.F. Agris, C.K. Hall, Simulation study of the ability of a computationally-designed peptide to recognize target $tRNA^{Lys3}$ and other decoy tRNAs, Protein Sci. 25 (2016) 2243-2255.

[23] D.A. Patrick, S.A. Bakunov, S.M. Bakunova, E.V Kumar, R.J. Lombardy, S.K. Jones, A.S. Bridges, O. Zhirnov, J.E. Hall, T. Wenzler, R. Brun, R.R. Tidwell, Synthesis and in Vitro Antiprotozoal Activities of Dicationic 3,5-Diphenylisoxazoles, J. Med. Chem., 50 (2007) 2468-2485.

[24] W. Porcal, P. Hernández, G. Aguirre, L. Boiani, M. Boiani, A. Merlino, A. Ferreira, R.D. Maio, A. Castro, M. González, H. Cerecetto, Second generation of 5-ethenylbenzofuroxan derivatives as inhibitors of Trypanosoma cruzi growth: Synthesis, biological evaluation, and structure–activity relationships, Bioorg. & Med. Chem., 15 (2007) 2768-2781.

[25] R. Sanz, J. Escribano, M.R. Pedrosa, R. Aguado, F.J. Arnáiz, Dioxomolybdenum(VI)-Catalyzed Reductive Cyclization of Nitroaromatics. Synthesis of Carbazoles and Indoles, Adv. Synth. & Catal., 349 (2007) 713-718.

[26] A.P. Kozikowski, D. Ma, J. Brewer, S. Sun, E. Costa, E. Romeo, A. Guidotti, Chemistry, Binding Affinities, and Behavioral Properties of a New Class of "Antineophobic" Mitochondrial DBI Receptor Complex (mDRC) Ligands, J. Med. Chem, 36 (1993) 2908-2920.

[27] S. Kamiya, H. Matsui, H. Shirahase, S. Nakamura, K. Wada, M. Kanda, H. Shimaji, N. Kakeya, Thromboxane A2 Synthetase Inhibitors with Histamine H1-Blocking Activity: Synthesis and Evaluation of a New Series of Indole Derivatives, Chem. Pharm. Bull. (Tokyo), 43 (1995) 1692-1695.

[28] K. Ohno, N. Araki, T. Yanase, H. Nawata, M. Iida, A novel nonradioactive method for measuring aromatase activity using a human ovarian granulosa-like tumor cell line and an estrone ELISA, Toxicol Sci. 82 (2004) 443-450.

[29] P. Cieplak, W.D. Cornell, C. Bayly, P.A. Kollman, Application of the multimolecule and multiconformational RESP methodology to biopolymers: charge derivation for DNA, RNA, and proteins, J. Comput. Chem. 16 (1995) 1357-1377.

[30] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, J. Comput. Chem. 25 (2004) 1157-1174.

[31] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB, J. Chem. Theory Comput. 11 (2015) 3696-3713.

[32] P. Li, K.M. Merz Jr., MCPB.py: A python based metal center parameter builder, J. Chem. Inf. Model 56 (2016) 599-604.

Highlights

► Discovery of novel non-steroidal inhibitors against aromatase

► Novel compounds demonstrate higher potencies in aromatase inhibitory than letrozole

► Computational models of aromatase with different substrates are built

► Binding mode of substrates to aromatase has significant impact on pharmacological effect