

# Rational Design of Allosteric Modulators of the Aromatase Enzyme: An Unprecedented Therapeutic Strategy to Fight Breast Cancer

Angelo Spinello,<sup>1</sup> Silvia Martini,<sup>2</sup> Federico Berti,<sup>3</sup> Marzia Pennati,<sup>2</sup> Matic Pavlin,<sup>1</sup> Jacopo Sgrignani,<sup>4</sup> Giovanni Grazioso,<sup>5</sup> Giorgio Colombo,<sup>6,7</sup> Nadia Zaffaroni,<sup>2</sup> Alessandra Magistrato<sup>1,\*</sup>

1) CNR-IOM-Democritos c/o International School for Advanced Studies (SISSA), Trieste, Italy.

2) Fondazione IRCCS Istituto Nazionale dei Tumori, Milano, Italy.

3) Dipartimento di Scienze Chimiche e Farmaceutiche, University of Trieste, Italy.

4) Institute for Research in Biomedicine, Bellinzona, Switzerland

5) Dipartimento di Scienze Farmaceutiche, University of Milano, Italy.

6) Dipartimento di Chimica, University of Pavia, Italy.

7) Istituto di Chimica del Riconoscimento Molecolare, CNR, Milano, Italy

\* Corresponding author. [alessandra.magistrato@sissa.it](mailto:alessandra.magistrato@sissa.it)

## Abstract

Estrogens play a key role in cellular proliferation of estrogen-receptor-positive (ER+) breast cancers (BCs). Suppression of estrogen production by competitive inhibitors of the enzyme aromatase (AIs) is currently one of the most effective therapies against ER+ BC. Yet, the development of acquired resistance, after prolonged treatments with AIs, represents a clinical major concern. Serendipitous findings indicate that aromatase may be non-competitively inhibited by clinically employed drugs and/or industrial chemicals. Here, by performing *in silico* screening on two putative allosteric sites, molecular dynamics and free energy simulations, supported by enzymatic and cell-based assays, we identified five leads inhibiting the enzyme via a non-active site-directed mechanism. This study provides new compelling evidences for the existence of an allosteric regulation of aromatase and for the possibility of exploiting it to modulate estrogens biosynthesis. Such modulation can aptly reduce side effects caused by the complete estrogen deprivation therapy, and, possibly, delay/avoid the onset of resistance.

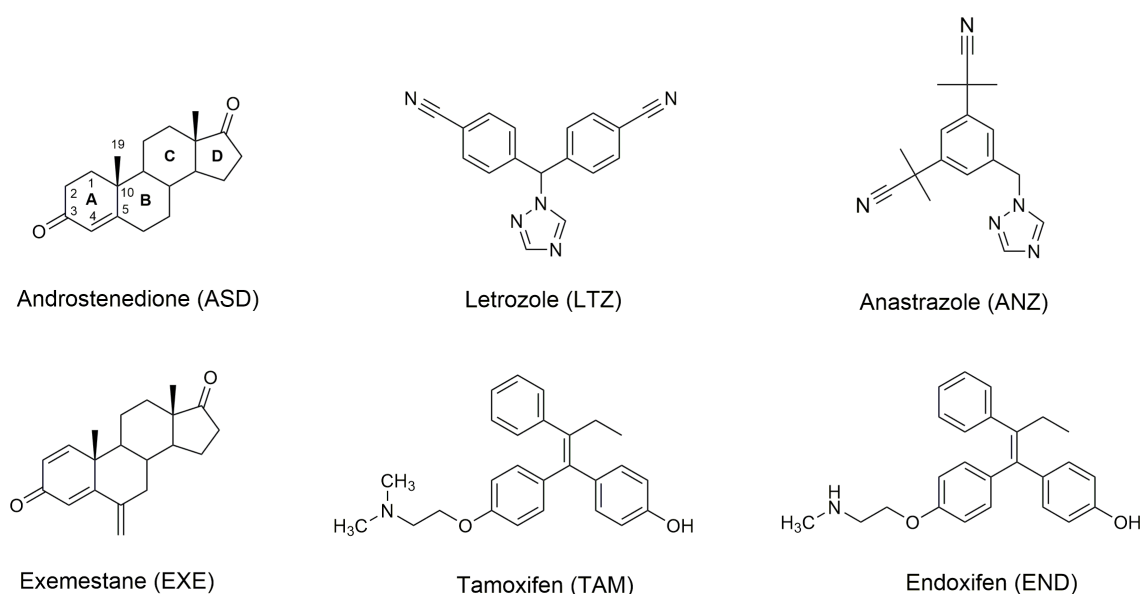
**Keywords:** cytochromes P450; aromatase; molecular dynamics; aromatase inhibitors; docking; breast cancer; resistance onset; mixed inhibition mechanism.

## 1. Introduction

The steroidogenic cytochromes P450 (CYP450s), CYP19A1 and CYP17A1, are relevant therapeutic targets for hormone-related tumors, such as breast and prostate cancers. In particular, CYP19A1, commonly referred as human aromatase (HA) enzyme, promotes the conversion of androgens (androstenedione (ASD), testosterone and 16 $\alpha$ -hydroxytestosterone), into estrogens, (estrone, 17 $\beta$ -estradiol and 17 $\beta$ ,16 $\alpha$ -estriol, respectively) via a three-step catalysis, requiring molecular oxygen, protons and electrons.<sup>1,2</sup> While the protons necessary for catalysis are supplied via a substrate/water channel opening in the vicinity of the active site,<sup>3,4</sup> electrons are provided by a NADPH cytochrome P450 reductase (CPR) for all microsomal CYP450s. Once synthesized, estrogens bind to estrogen receptor (ER), whose major form in breast cancer (BC) is ER $\alpha$ , encoded by the gene *ESR1*. ER $\alpha$  is a transcription factor, controlling genes associated with cell survival and proliferation.<sup>5</sup> This receptor stimulates cellular growth and proliferation, thus playing a critical role in estrogen receptor positive (ER+) BC, where abnormal concentrations of estrogens are produced by HA.

BC is the most frequent female cancer, accounting for nearly 30% of all cases in women,<sup>6</sup> among which about 60% and 75% of BCs in pre- and post-menopausal women, respectively, are ER+. Three types of endocrine therapies currently employed to treat ER+ BC patients are based on the use of: (i) selective estrogen receptor modulators (SERMs), such as tamoxifen, which block estrogen-induced ER $\alpha$  activation;<sup>7</sup> (ii) selective estrogen receptor degraders (SERDs), such as fulvestrant, which antagonize and degrade ER $\alpha$ ,<sup>8</sup> and (iii) HA inhibitors (Aromatase Inhibitors, AIs), which counteract estrogen biosynthesis.<sup>9</sup> The clinical use of AIs such as letrozole (LTZ), anastrozole (ANZ) and exemestane (EXE, Figure 1) represented a major breakthrough in the treatment of ER+ BC as these drugs display satisfactory clinical efficacy and manageable toxicity profiles.<sup>10-12</sup> Nevertheless, complete estrogen deprivation induced by their administration leads to troublesome side effects (i.e. exacerbation of menopausal symptoms, osteoporosis, loss of bone density and musculoskeletal complaints). Alarmingly, despite initial therapeutic benefits, acquired resistance leading to disease relapse may arise in metastatic BC patients after prolonged AIs treatment. The molecular details of

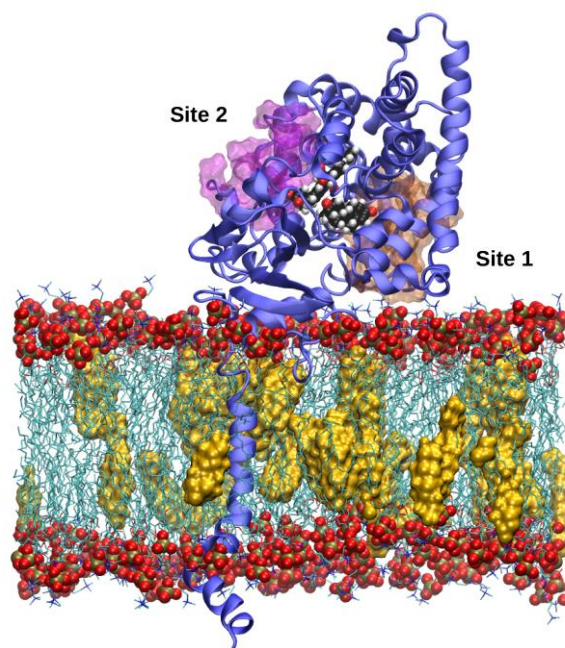
the complex scenario underlying resistance onset remains elusive,<sup>13</sup> even though large-scale genomic studies identified some aggressive ER $\alpha$  polymorphisms making the receptor intrinsically active even in the absence of estrogens.<sup>14</sup> As such, the development of novel therapeutic approaches able to modulate excessive estrogen biosynthesis with alternative molecular mechanisms represents both a current major clinical challenge to treat ER+ BC and a novel therapeutic opportunity.



**Figure 1.** Molecular structures of the endogenous substrate (ASD) of the enzyme aromatase (HA) along with the most clinically relevant aromatase inhibitors (AIs; LTZ, ANZ and EXE) and selective estrogen receptor modulators (SERMs; TAM and END).

Since the appearance of its first crystal structure,<sup>15</sup> HA has been object of intense investigations tackling several functional aspects,<sup>16-19</sup> and searching for novel competitive AIs.<sup>20-24</sup> Additionally, it was recently proposed that endoxifen (END), a tamoxifen metabolite belonging to SERMs, and letrozole (LTZ, Figure 1), the most prominent HA inhibitor, may exert a non-competitive/mixed inhibition mechanism.<sup>25,26</sup> The existence of this type of inhibition was further corroborated by kinetic studies on agricultural fungicides (bifonazole, imazalil, and flusilazole), and on industrial chemicals (triclocarban, an antimicrobial additive present in many industrial products).<sup>26,27</sup>

Exploiting non-competitive HA inhibition offers attractive advantages over its competitive counterpart as the former can reach a maximal inhibitory potency without completely abrogating estrogen production. This can expectedly result in the reduction of the side effects caused by conventional AIs and, possibly, in delaying/avoiding the onset of resistance. In this context, the *rational* search and design of mixed/non-competitive drug-candidates implies the non-trivial identification of allosteric sites.<sup>28</sup> In a previous study we have identified, using computational methods, three potential allosteric pockets,<sup>29</sup> two of them being involved in pivotal aspects of HA catalysis: (a) Site 1, lying along one of the possible substrate access channel to the catalytic site;<sup>30</sup> and (b) Site 2, also referred as the heme-proximal cavity and corresponding to the CPR binding site, which is believed to participate in the electron transfer from CPR to HA heme (Figure 2, HA residues composing Site 1 and 2 are reported in Table S1). Importantly, the existence of the latter site was confirmed by a recent crystal structure, trapping polyethylene glycol inside this cavity.<sup>31</sup>



**Figure 2.** Model of HA enzyme embedded in a POPC (phosphorous and oxygen atoms are shown as tan and red spheres)/CHL (yellow surface) membrane. Site 1 and 2 are shown as orange and magenta transparent surfaces, respectively, while heme and ASD are displayed in a van der Waals (vdw) representation.

With the aim of setting the stage for the discovery of new small-molecule therapeutic tools to tackle BC, here, we developed an integrated computational and experimental protocol based on: (i) structure-based virtual screening (VS) on libraries of commercial compounds; (ii) ensemble docking on distinct HA structures, targeting the two putative allosteric sites; (iii) cumulative 10  $\mu$ s-long force field (FF) based atomic-scale molecular dynamics (MD) simulations aimed at refining docking poses of selected compounds; (iv) spectrofluorometric and cell-based assays to establish the activity of the best-ranked molecules obtained from our *in silico* studies; (v) measurements of the inhibition kinetics; (vi) refinement of the binding pose, and estimation of the dissociation free energy barrier for the best drug-candidate via metadynamics (MTD) simulations.

As a result, five molecules inhibited HA in the low  $\mu$ M range via a non-active site-directed mechanism, strikingly providing the first compelling instance for the feasibility and exploitation of this alternative regulation mechanism of estrogen biosynthesis. Hence, our outcomes can lay the foundation for developing unexplored therapeutic strategies to treat ER+ BC, possibly able of delaying/avoiding the onset of resistance, which frequently develop under the pressure of current therapies.

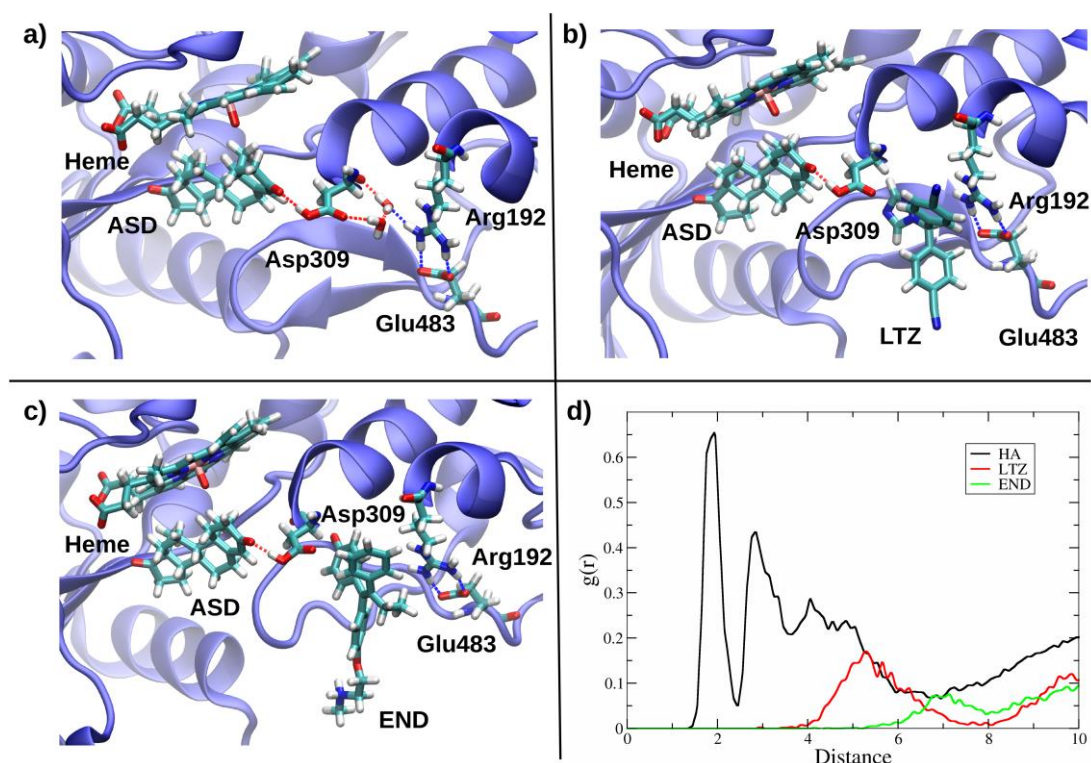
## 2. Results

### 2.1. Binding to the two putative allosteric pockets: the access channel cavity (Site 1)

A 400 ns MD simulation was initially performed on a HA model (PDB ID 3EQM)<sup>15</sup> embedded in a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine/cholesterol (POPC/CHL) membrane (Figure 2, RMSD plot is reported in Figure S1), with ASD bound to the active site. The distance between the center of mass of HA and the membrane and the tilt of the heme plane with respect to the membrane were consistent to previous studies (Table S2), confirming the reliability of our model.<sup>19,32,33</sup>

On the basis of recent evidence envisioning a mixed/non-competitive inhibition mechanism of LTZ,<sup>26</sup> we initially docked the latter into Site 1, using an induced fit protocol,<sup>34</sup> to assess whether this was a viable binding pocket. The best-ranked LTZ pose was then relaxed by 400 ns-long MD simulation

(see RMSD in Figure S1). Strikingly, after about 100 ns, LTZ increasingly penetrates into Site 1, displacing all the water molecules present within the cavity. In the absence of LTZ, Site 1, indeed, displays a network of water molecules lying between Asp309 and Arg192, (Figure 3a, b) as evidenced by radial distribution function ( $g(r)$ ), (Figure 3d) calculated along the MD trajectories,<sup>15</sup> and also observed in the enzyme crystal structure (pdb ID 3EQM). These water molecules are critical for HA function, as they supply the protons necessary for its catalysis,<sup>4,35,36</sup> and lay between Asp309 and Arg192, which mutagenesis studies identified as essential residues for enzymatic activity.<sup>37</sup> Upon LTZ binding to Site 1, the  $g(r)$  peaks are shifted to a larger distance (Figure 3d), clearly underlining that the drug displaces the waters molecules from Site 1. This hampers the delivery of protons to the substrate, one of the key requirements for estrogen production.



**Figure 3.** Representative structures of Site 1: the water network connecting Asp309 and Arg192 in the free enzyme (a) is displaced upon binding of LTZ (b) and END (c) to Site 1. Heme atoms, ASD, LTZ and END along with residues lining Site 1 are shown in licorice and colored by atom name, protein ribbons are shown in violet. d) Radial distribution function,  $g(r)$ , of the water molecules from the carboxyl oxygens of Asp309. Black, red and green lines refer to the apo, LTZ- and END-bound enzyme, respectively.

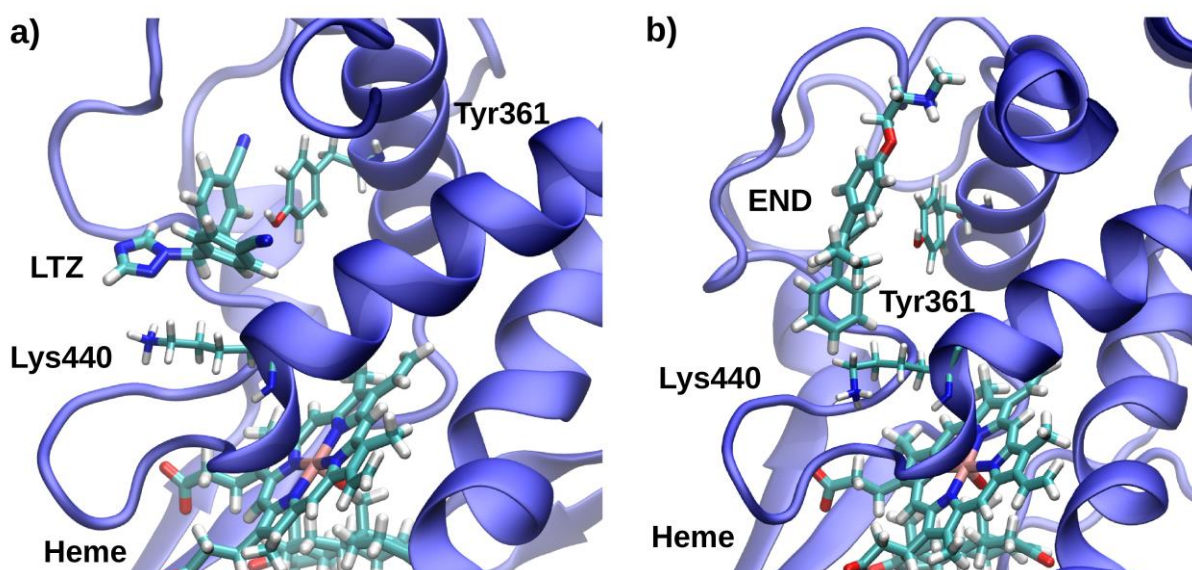
In order to inspect if the absence of the endogenous substrate, ASD, in the active site could affect LTZ binding pose, we also performed additional 200 ns of MD after ASD removal. As a result, LTZ stably sits within Site 1 retaining its binding pose (RMSD 0.88 Å), while the active site maintains its shape (Figure S2).

To further validate our findings, we also considered the binding of both E and Z endoxifen isomers to Site 1. In line with our previous study,<sup>29</sup> Z-END dissociates within 50 ns of MD simulation, while E-END (defined simply as END hereafter) firmly binds to this allosteric site for 400 ns-long MD simulation (Figure 3c) and, even more pronouncedly than LTZ, displaces the water molecules from the access channel (Figure 3d). Hence, upon binding to Site 1, END and LTZ may prevent/reduce catalysis via a similar mechanism, which consists in hampering the proton delivery necessary for estrogen biosynthesis.

Finally, since Site 1 lies along one of the substrate access/egress channel<sup>30</sup> and a two-steps substrate binding kinetic has been demonstrated for HA,<sup>38</sup> we additionally docked ASD to Site 1 and verified its stability by performing 200 ns-long MD simulation. Our simulation confirm that ASD may bind to Site 1 along its entrance/egress route (Figure S3).

## ***2.2. The heme-proximal cavity (Site 2)***

Following the same protocol adopted above, LTZ was docked to Site 2, where it remained stably bound for 400 ns of MD. In the best-ranked pose, LTZ lies between Lys440 and Tyr361 (Figure 4a). These are again two critical residues for estrogen biosynthesis, as evidenced by mutagenesis studies,<sup>37</sup> believed to stabilize the interactions between HA and its redox partner CPR. Hence, the binding of a drug at this site may hamper the formation of the CPR/HA adduct and/or affect the electrons flow necessary for catalysis. Similarly to LTZ, also END stably binds in between Lys440 and Tyr361 (see Figure 4b).



**Figure 4.** Representative snapshots extracted from the MD simulation of the binding of LTZ (a) and END (b) to Site 2. Heme atoms, ASD, LTZ and END along with residues lining Site 2 are shown in licorice and colored by atom name, protein ribbons are shown in violet

The binding free energy ( $\Delta G_b$ ), as calculated with Molecular Mechanics Generalized Born Surface Area (MM-GBSA),<sup>39-41</sup> of LTZ in the two allosteric sites reveals its higher affinity for Site 1 (Table 1), while END does not exhibit a clear preference for any of the two sites (Table 1).

### 2.3. Non-active site-directed inhibitors: identification and experimental validation

Virtual screening (VS) simulations were performed to identify drug-like molecules binding to both allosteric pockets, using three distinct target conformations (ensemble docking, see methods) in order to account for receptor flexibility. Among the best 100 compounds we selected, upon visual inspection, 42 molecules according to their ability to interact with critical residues involved in HA catalysis (i.e. Asp309, Arg192, Tyr361 and Lys440), and to maximize their structural diversity. The binding mode and the relative energy ranking of these molecules were then refined by performing 100 ns-long MD simulations of each drug/HA adduct. Twenty-one molecules spontaneously dissociated from HA within this simulation time, underlying the importance of accounting for protein dynamics whilst searching for non-active site-directed inhibitors.<sup>42,43</sup>



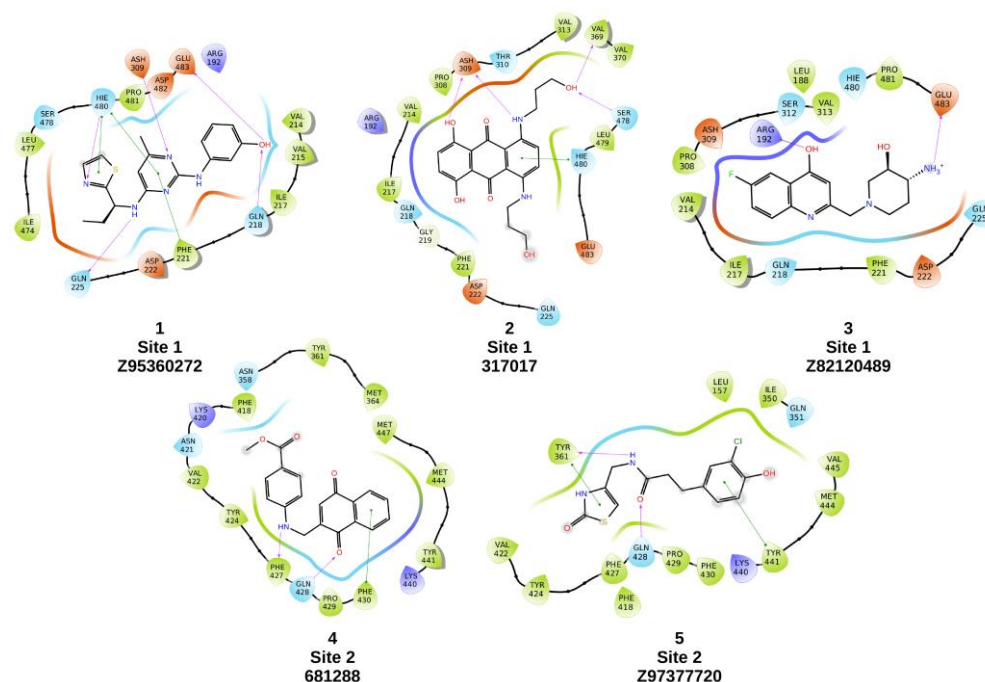
The other molecules, which remained stably bound to HA in this simulation time, were re-ranked by their relative  $\Delta G_b$ , computed with MM-GBSA (Table S3). Among them, the 18 best-ranked molecules were, then, tested as potential HA inhibitors, at initial concentrations of 100 and 200  $\mu\text{M}$  with an HA fluorimetric kit, using LTZ and END as controls. Six molecules (ZINC database: Z95360272, Z82120489, Z97377720, Drugbank: Taxifolin, NCI database: 317017, 681288) showed activity at both concentrations (Table S3). Taxifolin was excluded from further tests due to the similarity with the endogenous substrate ASD, which would most probably lead to a competitive mechanism of inhibition. Next, the remaining molecules were tested at different concentrations (ranging from 1 to 200  $\mu\text{M}$ ) to determine the  $\text{IC}_{50}$  value (i.e. the concentration able to inhibit HA catalytic activity by 50%) of each compound. All molecules showed  $\text{IC}_{50}$  values in the  $\mu\text{M}$  range (Table 1 and Figure S4). Consistent with its high inhibitory potency, LTZ was active in the nM range,<sup>26</sup> while, in contrast to previous studies, END showed activity only at a concentration of 200  $\mu\text{M}$ .<sup>25</sup>

**Table 1:** Compounds displaying inhibitory activity in fluorimetric assays. Database codes of the active drug-like molecules, binding free energies ( $\Delta G_b$ , kcal/mol) computed with the MM-GBSA method,  $\text{IC}_{50}$  fluorimetrically derived,  $\text{GI}_{50}$  calculated on ER+ (MCF-7) and ER- (MDA-MB-231) cell lines, and the Site in which the molecules were screened are reported from left to right columns, respectively.

Compound	Database Code	$\Delta G_b$ (kcal/mol)	$\text{IC}_{50}$ ( $\mu\text{M}$ ) on HA	$\text{GI}_{50}$ ( $\mu\text{M}$ ) on MCF-7	$\text{GI}_{50}$ ( $\mu\text{M}$ ) on MDA- MB-231	Site
LTZ	-	$-21.7 \pm 2.5$	$0.01^{26}$	$4.1 \pm 1.1$	$>10$	1
		$-17.1 \pm 2.3$				2
END	-	$-37.1 \pm 3.6$	$6.1^{25}$	-	-	1
		$-37.1 \pm 2.9$				2
1	ZINC-95360272	$-29.8 \pm 3.2$	$5.1 \pm 0.6$	$21.3 \pm 0.009$	$58.0 \pm 0.240$	1
2	NCI-317017	$-36.2 \pm 4.1$	$11.0 \pm 4.0$	$17.9 \pm 0.5$	$6.8 \pm 0.004$	1
3	ZINC-82120489	$-24.2 \pm 3.9$	$20.9 \pm 1.6$	$>200$	$>200$	1
4	NCI-681288	$-30.7 \pm 4.3$	$95.3 \pm 6.6$	$22.1 \pm 0.54$	$16.2 \pm 0.009$	2
5	ZINC-97377720	$-26.6 \pm 2.8$	$21.7 \pm 3.9$	$66.4 \pm 0.035$	$>200$	2

Among these compounds, three (**1**, **2** and **3**) molecules were identified by performing VS on Site 1, while the remaining two (**4** and **5**) on Site 2. Thus, the molecules targeting Site 1 establish interactions with Arg192, Asp309, Glu483 (H-bonds) and His480 ( $\pi$ - $\pi$  stacking), while those binding Site 2 establish interactions with Tyr361, Gln428 (H-bonds) and Phe430 and Tyr361 ( $\pi$ - $\pi$  stacking) (Figure 5). A summary of all residues involved in the intermolecular interactions for each allosteric Site is reported in Table S4.

All active molecules and LTZ were also tested for their ability to inhibit the growth of prototypical ER+ (MCF-7) and ER- (MDA-MB-231) BC cell lines (Table 1 and Figure S5). Compounds **1**, **2**, **4** and **5**, although to a different degree, were active in reducing the growth of MCF-7 cells, as indicated by the different Grow Inhibition activity ( $GI_{50}$ ) values (i.e. the test agent concentration able to inhibit growth by 50%). Moreover, consistently with the lack of ER expression, MDA-MB-231 cells were less sensitive than MCF-7 cells to **1** and completely resistant to **5** (Table 1 and Figure S5), which represent the best compounds for each site. These important results suggest the correct engagement of the desired target by the selected compounds.



**Figure 5.** Ligand interaction diagrams of the active drug-like molecules identified. H-bonds are represented as purple arrows, while stacking interactions are depicted in green. Site of binding and commercial code of the molecules is also reported on the bottom.

Enzymatic kinetics analyses were performed on these five active compounds to assess their type of inhibition mechanism. These molecules were tested at constant concentration (set at 33% of their IC<sub>50</sub> value) and at different concentrations of substrate, which ranged from 0.15 to 3.6 μM (Figure S6). The kinetic parameters were obtained both as the median values from a Cornish-Bowden direct linear analysis of the experimental data<sup>44</sup> and from Michaelis-Menten analysis (Table 2).

**Table 2:** Kinetic parameters of the inhibited reactions. \*:Median values from the ordered series of the intersection coordinates of all the possible couples of straight lines of equation:  $V_{max} = v_i + \frac{v_i}{[S_i]} K_m$ , obtained from each couple of experimental points  $\{v_i; S_i\}$ , according to the Cornish-Bowden direct linear method. §: values from the regression of experimental data to the Michaelis-Menten equation.

	$\widehat{V}_{max}^*$ (Fl.units msec <sup>-1</sup> )	$\overline{V}_{max}^{\S}$ (Fl.units msec <sup>-1</sup> )	$\widehat{K}_m^*$ (μM)	$\overline{K}_m^{\S}$ (μM)
<b>No inhibitor</b>	91.5	93.3 ± 1.3	1.92	2.73 ± 0.09
<b>1</b>	15.1	19.3 ± 0.4	1.28	2.51 ± 0.13
<b>2</b>	16.2	21.5 ± 0.3	1.13	2.31 ± 0.14
<b>3</b>	50.7	64.6 ± 1.5	1.30	2.78 ± 0.17
<b>4</b>	60.6	79.3 ± 0.3	2.16	3.93 ± 0.06
<b>5</b>	78.7	85.2 ± 0.3	3.46	3.54 ± 0.03

This analysis revealed that the maximum rate of the non-inhibited enzyme reaction is clearly no longer attainable in the presence of the inhibitors. This is particularly evident with **1** and **2**, where the maximum rate is about 16% of the non-inhibited one. The apparent values of Km are only slightly affected by inhibition with **1**, **2** and **3**, while a somewhat larger effect is observed with **4** and **5**, which are predicted to interact at Site 2. In this case the affinity for the substrate seems to decrease by less than two times. This finding may be in line with an allosteric communication between the proximal cavity and the active site recently identified in other CYP450s.<sup>45</sup>

Although preliminary, these measurements clearly support the occurrence of a mixed mechanism of inhibition for all tested molecules. At the same time, these results are not consistent with a competitive inhibition, supporting our initial working hypothesis for the discovery of non-competitive/mixed HA modulators.

#### ***2.4. Refinement of the binding poses and dissociation kinetics***

Target and ligand flexibility is an essential feature to identify allosteric pockets and drug-like molecules targeting them. In this context, the ligand dissociation kinetics is increasingly being recognized as one of the key traits in drug efficacy optimization studies.<sup>46-48</sup> Hence, we have employed MTD simulations, which by enhancing the exploration of the binding pocket and of the ligand dissociation route, can provide a refinement of the inhibitor binding pose, and furnish information about its dissociation free energy barriers ( $\Delta G_d^\ddagger$ ). These simulations were needed since LTZ and the identified inhibitors displayed calculated  $\Delta G_b$  values with opposite trend with respect to their measured  $IC_{50}$ s and in order to possibly infer drug efficacy by evaluating the relative unbinding kinetics of the newly discovered inhibitors with respect to LTZ, which has a proved clinical efficacy. To this end, MTD simulations were performed for the best lead identified (**1**) and for LTZ bound to Site 1. These disclosed that **1** assumes a slightly refined binding pose with respect to those obtained from combined docking and MD simulations (Figure S7). Moreover, the two compounds dissociate from Site 1 with a similar free energy cost ( $\Delta G_d^\ddagger$  of  $11.4 \pm 1.4$  and  $11.8 \pm 1.3$  kcal/mol, respectively). Hence, these results support the potential of our best lead molecule which, besides displaying a good binding affinity towards the target ( $\Delta G_b$ ), is characterized by a dissociation kinetic from its target site similar to that of a drug in clinical use.<sup>30</sup>

### **3. Discussion**

Despite the extended survival rate achieved for patients in the last decades, BC remains a frightening disease, being the most common female cancer and the second cause of cancer death in women worldwide. Approximately 70% of BC patients are affected by ER+ carcinomas for which one of most exploited therapeutic approach relies on abrogating estrogen biosynthesis via HA inhibition. Although AIs-based therapies show higher clinical efficacy and induce less severe side effects than SERMs, the development of acquired resistance, after prolonged treatments, remains the main concern for ER+ BC therapy. Acquired resistance to AIs (as well as to SERMs) is due to an intricate

and incompletely understood interplay of different mechanisms, which impact on multiple competing biological pathways.<sup>13</sup> One major route is represented by the occurrence of aggressive ER $\alpha$  somatic mutations, which are often developed in metastatic BC patients (up to 40%) under the evolutionary pressure of prolonged therapies relying on estrogen deprivation. These mutations make ER intrinsically active. As a result, BC becomes insensitive to estrogens suppression.

With the aim of exploring alternative en-route to novel therapeutic approaches against BC, here we provide a proof-of-principle study, which lays the foundation for novel regulatory strategies of estrogen biosynthesis potentially able to avoid the onset of resistance issues. This relies on a non-competitive/mixed inhibition of HA catalysis. Exploiting allostery to modulate estrogen biosynthesis offers major advantages with respect to conventional competitive drugs in that: (i) allosteric drugs can reach the maximal inhibitory power without complete inhibition of estrogen production, (as instead competitive inhibitors do) (ii) they may not be outcompeted at higher concentrations of natural substrates and, (iii) targeting allosteric sites, which are less conserved across protein families compared to classical active sites, may offer the possibility of developing selective drugs, limiting off-target interactions and side effects. As such, we suggest that our approach represents a viable and valuable strategy for developing alternatives to current therapeutic options.

By introducing a computational protocol that integrates ensemble docking, to account for receptor flexibility, cumulative 10  $\mu$ s MD simulations to consider allosteric site plasticity and adaptability to the screened drug-like compounds, binding and dissociation free energy calculations, we have identified new leads active in the  $\mu$ M range, fitting in two allosteric pockets. The best compounds identified from our combined *in silico* and experimental assays (**1-3**) target Site 1, displaying IC<sub>50</sub> values in the low  $\mu$ M range,<sup>1</sup> in line with the challenging and flexible nature of allosteric pockets, which usually are targeted by the ligands with lower affinity.<sup>45</sup> Importantly, enzymatic kinetics assays suggest that a mixed inhibition mechanism may be operative for all the identified active molecules.

---

<sup>1</sup> We remark that compound **1**, which is optically active, was tested as a racemic mixture, but according to our calculation only one enantiomer (1*S*) can tightly bind in Site 1, suggesting that its inhibitory power will be even higher than that measured by IC<sub>50</sub> and GI<sub>50</sub>. Compound **3** was purchased as a single diastereoisomer (3*R*, 4*R*).

Furthermore, our simulations rationalize their possible mechanism of action, pinpointing that, upon binding to Site 1 and 2, these compounds may hamper catalysis by interfering with the delivery of the protons, and electrons, respectively, necessary for estrogen production. Among those, molecules targeting Site 1 (**1-3**) have lower  $IC_{50}$ ,  $GI_{50}$  and larger  $\Delta G_b$ , pointing to this site as the most promising target candidate for future drug optimization studies. Remarkably, cell growth inhibition assays disclose that our best leads, **1** and **5**, targeting Site 1 and 2, respectively, are more effective in reducing the growth of MCF-7 ER+ cells with respect to MDA-MB-231 ER- ones.<sup>49</sup>

Our protocol underscores the necessity of accounting for protein dynamics via atomic-scale MD simulations in order to refine docking poses and eliminate false positives (i.e. highly ranked compounds by rigid docking algorithms/conventional scoring functions, which quickly dissociates from their target when solvent, receptor flexibility and finite temperature effects are accounted). Interestingly, the ligand dissociation rate is often used as an indicator of the residence time of a drug on its receptor, thus becoming a new paradigm in drug optimization studies aiming at increasing the desired biological effect.<sup>50</sup> By exploiting a MTD-based protocol, which provides an evaluation of ligand dissociation rate, by computing its dissociation free energy barrier ( $\Delta G_d^\ddagger$ ), we disclose that **1** and LTZ, that has a well-established clinical efficacy, have similar residence time at Site1.

The relevance of allostery has at times been considered in CYP450s, which besides participating to steroidogenesis, such as HA or CYP17A1, are also involved in the metabolism of exogenous compounds. While distinct allosteric sites types have been identified in this latter CYP450s category,<sup>51-53</sup> the impact on enzymatic activity induced upon effectors binding to those sites remains controversial.<sup>54</sup> An allosteric modulation of CYP450s catalysis has been, instead, suggested upon the binding of protein redox partner (CYPb<sub>5</sub>) to CYP17A1,<sup>55</sup> in addition to the serendipitous observation of non-competitive inhibition kinetics for HA exerted by drugs and industrial chemicals.<sup>27</sup> Despite these evidences support the hypothesis of exploiting allostery for CYP450s functional regulation, to the best of our knowledge no effective non-competitive/mixed inhibitor of any CYP450 has been hitherto identified in rational drug discovery studies.

Hence, our results, besides corroborating the existence of two functioning allosteric sites in HA, identify for the first-time drug-like mixed inhibitors of this enzyme with a fully rational approach, setting the stage for the development of novel regulatory strategies of estrogen biosynthesis to fight ER+ BC. This may, eventually, offer precious therapeutic alternatives depending on disease progression and on patients' genetic signatures, which may encode for the ER $\alpha$  isoforms refractory to conventional therapies.

## 4. Materials and Methods

### 4.1. *In silico* screening

The NCI (<https://cactus.nci.nih.gov/download/nci/>), drugbank<sup>56</sup> and the clean lead subset of the ZINC databases<sup>57</sup> were used as ligand libraries for VS studies. In order to take into account ionization and tautomerization states of the compounds, Schrodinger Suite 2017-1 Epik tool was used,<sup>58</sup> with a maximum number of 4 structures for each molecule. Compounds were filtered using the Schrodinger Ligfilter tool.<sup>58</sup> Ligands that violated Lipinski's rule of five, in order to eliminate compounds possessing poor absorption and permeation,<sup>59</sup> as well as compounds with more than 10 rotatable bonds, were discarded, since high ligand flexibility implies higher entropic contributions and reduce oral availability.<sup>60</sup> As well, Qikprop<sup>58</sup> was employed to predict LogP values of the compounds.

*In silico* screening of the above-mentioned libraries was performed on Site 1 and Site 2. A list of the residues composing these two sites and used for grid generation is reported in Table S1.<sup>29</sup> Accounting for receptor flexibility during *in silico* screening of ligands targeting allosteric pockets is highly recommended.<sup>28,61,62</sup> Hence, we performed ensemble docking simulations by selecting three structures of the target: (i) the HA crystal structure (pdb ID 3EQM),<sup>15</sup> (ii) the most representative protein structure as obtained from a cluster analysis of its MD simulation in the presence of a membrane mimic and (iii) the most representative structure obtained from the MD simulations of HA in the presence of LTZ bound either to Site 1 or to Site 2. The latter structures allow to account also for the induced fit effects imposed by LTZ, which is an already established mixed/non-competitive

inhibitor.<sup>26</sup> All water molecules were deleted before starting docking calculations. A van der Waals (vdW) radius scaling factor of 0.80 Å for protein and ligands atoms having a partial charge less than 0.15 was used in order to account for protein flexibility. A VS workflow<sup>58</sup> based on three subsequent steps of docking with increasing level of accuracy, was adopted using the Glide program.<sup>63</sup> Namely, (i) a fast high-throughput virtual screening (HTVS) was initially performed in order to efficiently select promising ligands among millions of compounds-large libraries; (ii) 5 % of the best ranked ligands are retained and a single precision (SP) docking calculation is performed; (iii) the top 5 % of the resulting compounds are screened using the extra precision (XP) protocol. This latter should eliminate false positives by using a more extensive sampling and more accurate scoring functions. The resulting molecules were sorted according to GlideScore scoring function, and, after visual inspection, the poses of the top-ranked 42 compounds were refined by performing classical MD simulations as detailed below.

An induced fit protocol<sup>34</sup> was used to dock LTZ and END, in Site 1 and Site 2, since they are suggested to be non-competitive/mixed inhibitors,<sup>25,26</sup> and the substrate, ASD, in Site 1. An initial SP docking was performed with a softened vdW radius scaling of 0.50 for ligands and protein atoms having a partial atomic charge of 0.15. Afterwards, 100 poses were submitted to Prime<sup>64</sup> for prediction of new side chains orientation for the residues within a shell of 6 Å from the ligand center of mass. Finally, a Glide re-docking was performed for HA/ligand adducts with score laying within 30 kcal/mol from the best one. The best ranked poses of the ligands inside Site 1 and Site 2 were subjected to MD simulations.

#### **4.2. Classical MD simulations**

An equilibrated HA model was taken from our previous work<sup>29</sup> (PDB ID 3EQM)<sup>15</sup> and, by using the CHARMMGUI webserver,<sup>65</sup> it was embedded in a POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) membrane containing 6 wt% of cholesterol (CHL) in order to mimic the endoplasmic reticulum membrane.<sup>66</sup> Physiological protonation states were calculated with the



webserver H++.<sup>67</sup> Asp309 was considered in its neutral form consistently with literature studies.<sup>3,68</sup> Parm99SB AMBER FF was employed for the protein,<sup>69,70</sup> lipid14 for the lipids,<sup>71</sup> Shahrokh *et al.* parameters for the heme moiety and Cys437.<sup>72</sup> Most simulations were done in presence of the substrate ASD for which the general Amber FF (GAFF) was used.<sup>73</sup> ESP charges<sup>74</sup> were calculated by performing geometry optimization of the substrates at Hartree-Fock level of theory using a 6-31G\* basis set with the Gaussian 09 software<sup>75</sup> and were later transformed in RESP charges with the Antechamber module of ambertools 16.<sup>76</sup> The system was solvated by water molecules (TIP3P model),<sup>77</sup> leading to a total of 131454 atoms. Topology, built with the ambertools 16, was later converted in a GROMACS format using the software acpype.<sup>78</sup> MD simulations were performed with GROMACS 5.0.4.<sup>79</sup> An integration time step of 2 fs was used and all covalent bonds involving hydrogen atoms constrained with the LINCS algorithm. Particle Mesh Ewald algorithm<sup>80</sup> was used in order to account electrostatic interactions. Simulations were performed in the isothermal-isobaric NPT ensemble, at a temperature of 300 K, under control of a velocity-rescaling thermostat.<sup>81</sup> Preliminary energy minimization was done with the steepest descend algorithm. An initial equilibration of the membrane was performed for 100 ns with the protein atoms harmonically restrained with a force constant of  $1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ , reaching a constant value ( $92 \times 92 \times 151 \text{ \AA}^3$ ) of the simulation box size. Constraints were then released, and the system was slowly thermalized to the target temperature of 300 K within 10 ns. Then, it was relaxed by performing a 400 ns MD simulation. In order to assess the stability of the docking poses obtained from the VS of the NCI, drugbank and ZINC databases the 42 best-ranked ligands were inserted into this equilibrated membrane and each one underwent 100 ns classical MD simulations.

### **4.3. Metadynamics (MTD) simulations**

In order to further refine the binding poses of LTZ and the compound with the best inhibitory activity (1) we have performed FF based MTDs simulations. In particular, MTDs runs of 50-100 ns were performed to refine the binding pose and study ligand dissociation with GROMACS 5.0.4 using the

PLUMED 2.0 plugin.<sup>82</sup> Two collective variables (CVs) were used: the first (CV1) describes the number of hydrophobic contacts between the ligands and Site 1, computed as a coordination number; the second (CV2) corresponds to the distance between the center of masses (COM) of the protein and the ligand. Gaussian hills having a height of 0.6 kJ/mol and widths of 0.016 and 0.095 (LTZ), 0.024 and 0.160 (**1**), were added respectively for CV1 and CV2 every 4 ps of MD. A harmonic wall was used to restrain the exploration of the FES on CV2. Three replicas of the MTD simulations were performed, starting from different frames as extracted from the equilibrated trajectory and the uncertainty of the dissociation free energy barriers ( $\Delta G_d^\ddagger$ ) were estimated from the standard deviation of the barriers obtained out of the three replicas, following previous simulations studies.<sup>4,83,84</sup>

#### **4.4. Analysis**

Cluster analysis, root mean square deviation (RMSD), radial distribution function ( $g(r)$ ) of the MD trajectories were done with the `g_cluster` tool, based on the Daura et. al algorithm,<sup>85</sup> `g_rmsd`, and `g_rdf` as implemented in the GROMACS 5.0.4 program. Molecular Mechanics Generalized Born Surface Area (MM-GBSA) free energy calculation were performed with the `MM_PBSA.py` tool of Amber 12 program, following a procedure successfully applied in previous studies,<sup>29,86</sup> keeping parameters at the default values.<sup>87</sup> The distance of the protein from the membrane, is obtained as the distance between the center of mass of HA and the center of mass of the membrane. The heme tilt angle, is represented by the angle between the heme plane, defined by heme nitrogen atoms, and the z-axis. Visualization of the MD trajectories and images were done with the VMD program.<sup>88</sup>

#### **4.5. HA inhibition and kinetics assays**

Inhibition of HA was quantified by the Aromatase Inhibitor Screening Kit (BioVision Inc., San Francisco, USA), using a fluorogenic substrate that is converted by HA into a highly fluorescent metabolite. Briefly, after the reconstitution of the reagents, a standard curve was generated by diluting the fluorescent standard. Test compounds were dissolved in DMSO at a final concentration of  $\leq$

0.25% (v/v), after having verified that such concentration of solvent does not affect the enzyme activity in significant way. Each mother solution was diluted in aromatase assay buffer to obtain a range of concentrations for generating a multi-point dose-response curve. The reaction was prepared by adding Aromatase mix (containing Recombinant Human Aromatase (2X), Aromatase assay buffer and NADPH-generating system) to test compounds, inhibitor control, background control and positive control (1  $\mu$ M LTZ). The reaction mixture was preincubated at 37°C for 10 min to allow test compounds to interact with HA, then, reaction initiated after the addition of 30  $\mu$ l of Aromatase Substrate/NADP<sup>+</sup> mixture (containing buffer, aromatase substrate and  $\beta$ -NADP<sup>+</sup> 100X stock). Assays were conducted in 96-well microtiter plates (Corning Incorporated, Corning, ME, USA) in a final reaction volume of 100  $\mu$ l/well. Sample fluorescence was measured using a TECAN Ultra microplate reader (Tecan Trading AG, Switzerland) at dual wavelengths of 488/527 nm for 60 min. Results were expressed as relative fluorescence units (RFUs). Experiments were performed in triplicate and the average values were used to construct the dose-response curves. The percentage of inhibition was calculated as the ratio between the RFUs of control and test compound, according to the manufacturer's instructions.

The 20 molecules selected on the basis of our VS protocol were purchased from Enamine, Chembridge, and Sigma-Aldrich or donated by the NCI (Table S5). Tests were initially performed at 200 and 100  $\mu$ M concentrations of the inhibitor to identify potentially active molecules. All chiral compounds were tested as racemic mixtures. The five compounds that showed  $\geq 40\%$  inhibition at the concentration of 100  $\mu$ M were then tested at different concentrations (200, 100, 50, 10, 1 and 0.1  $\mu$ M). For each compound, the concentration able to inhibit HA activity by 50% (IC<sub>50</sub>) was calculated by nonlinear regression of the experimental data to a tetraparametric logistic curve (SigmaPlot 13.0 - Systat Software Inc.).

To obtain a preliminary analysis of HA inhibition mechanism, we performed several kinetic runs at constant test compound (set at 33% of their IC<sub>50</sub> values) and at different substrate concentrations ranging from 0.25 to 6 times the approximate value of K<sub>m</sub> declared by the supplier of the screening

kit. The substrate and inhibitors concentrations were chosen within the limitations imposed by compound solubility; 1% acetonitrile was used in all the measures. Under such conditions, the % inhibition decreases as the substrate concentration increases in the case of competitive inhibition, while it remains constant if the inhibition is noncompetitive and increases if the inhibition is uncompetitive.<sup>44,89</sup> It is important to remark that, among these five molecules, compounds **2** and **4**, possessing a quinone moiety, are reported as potential PAINS (Pan Assay Interference Compounds), and consequentially could interfere with biochemical high throughput screenings.<sup>90</sup>

#### ***4.6. Cell growth inhibition assay***

The human breast cancer cell lines MCF-7 (ER+) and MDA-MB-231 (ER-) were obtained from the American Type Culture Collection (Rockville, MD, USA). Tumor cells were seeded in triplicate in 12-well plates (50000 cells/well) and 24 h later exposed for 72 h to increasing concentrations (from 0.01 to 10  $\mu$ M, 1 to 200  $\mu$ M and 10 to 200  $\mu$ M for LTZ, compounds **1** - **4** and **5**, respectively). Tumor cells were harvested using trypsin and counted with a cell counter (Beckman Coulter, S.p.A., Milan, Italy). Three independent experiments (each in triplicate) were carried out and the average values were used to construct dose-response curves and calculate the concentration of compound able to inhibit cell growth by 50% (GI<sub>50</sub>).

#### **Acknowledgements**

AM thanks the Italian Cancer Research Association (AIRC: My first AIRC grant no 17134) for financial support. We thank CINECA, the Italian supercomputing center, for computational resources via the ISCRA B grant 'MAESTRO'.

#### **Conflicts of interest**

The authors declare no conflicts of interest.

## References:

1. Simpson, E. R.; Mahendroo, M. S.; Means, G. D.; Kilgore, M. W.; Hinshelwood, M. M.; Graham-Lorence, S.; Amarneh, B.; Ito, Y.; Fisher, C. R.; Michael, M. D.; et al. Aromatase cytochrome P450, the enzyme responsible for estrogen biosynthesis. *Endocr. Rev.* **1994**, *15*, 342-55.
2. Spinello, A.; Ritacco, I.; Magistrato, A. The Catalytic Mechanism of Steroidogenic Cytochromes P450 from All-Atom Simulations: Entwinement with Membrane Environment, Redox Partners, and Post-Transcriptional Regulation. *Catalysts* **2019**, *9*, 81.
3. Sgrignani, J.; Magistrato, A. Influence of the membrane lipophilic environment on the structure and on the substrate access/egress routes of the human aromatase enzyme. A computational study. *J. Chem. Inf. Model.* **2012**, *52*, 1595-606.
4. Spinello, A.; Pavlin, M.; Casalino, L.; Magistrato, A. A Dehydrogenase Dual Hydrogen Abstraction Mechanism Promotes Estrogen Biosynthesis: Can We Expand the Functional Annotation of the Aromatase Enzyme? *Chem. Eur. J.* **2018**, *24*, 10840-10849.
5. Wang, Z. Y.; Yin, L. Estrogen receptor alpha-36 (ER-alpha36): A new player in human breast cancer. *Mol. Cell Endocrinol.* **2015**, *418 Pt 3*, 193-206.
6. Liang, J.; Shang, Y. Estrogen and cancer. *Annu. Rev. Physiol.* **2013**, *75*, 225-40.
7. Lewis, J. S.; Jordan, V. C. Selective estrogen receptor modulators (SERMs): mechanisms of anticarcinogenesis and drug resistance. *Mutat. Res.* **2005**, *591*, 247-63.
8. McDonnell, D. P.; Wardell, S. E.; Norris, J. D. Oral Selective Estrogen Receptor Downregulators (SERDs), a Breakthrough Endocrine Therapy for Breast Cancer. *J. Med. Chem.* **2015**, *58*, 4883-7.
9. Ghosh, D.; Lo, J.; Egbuta, C. Recent Progress in the Discovery of Next Generation Inhibitors of Aromatase from the Structure-Function Perspective. *J. Med. Chem.* **2016**, *59*, 5131-48.
10. Bonnetterre, J.; Buzdar, A.; Nabholz, J. M.; Robertson, J. F.; Thurlimann, B.; von Euler, M.; Sahmoud, T.; Webster, A.; Steinberg, M.; Arimidex Writing, C.; Investigators Committee, M. Anastrozole is superior to tamoxifen as first-line therapy in hormone receptor positive advanced breast carcinoma. *Cancer* **2001**, *92*, 2247-58.
11. Mouridsen, H.; Gershanovich, M.; Sun, Y.; Perez-Carrion, R.; Boni, C.; Monnier, A.; Apffelstaedt, J.; Smith, R.; Sleeboom, H. P.; Jaenicke, F.; Pluzanska, A.; Dank, M.; Becquart, D.; Bapsy, P. P.; Salminen, E.; Snyder, R.; Chaudri-Ross, H.; Lang, R.; Wyld, P.; Bhatnagar, A. Phase III study of letrozole versus tamoxifen as first-line therapy of advanced breast cancer in postmenopausal women: analysis of survival and update of efficacy from the International Letrozole Breast Cancer Group. *J. Clin. Oncol.* **2003**, *21*, 2101-9.
12. Howell, A.; Cuzick, J.; Baum, M.; Buzdar, A.; Dowsett, M.; Forbes, J. F.; Hocht-Boes, G.; Houghton, J.; Locker, G. Y.; Tobias, J. S.; Group, A. T. Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. *Lancet* **2005**, *365*, 60-2.
13. Augusto, T. V.; Correia-da-Silva, G.; Rodrigues, C. M. P.; Teixeira, N.; Amaral, C. Acquired resistance to aromatase inhibitors: where we stand! *Endocr. Relat. Cancer* **2018**, *25*, R283-R301.
14. Pavlin, M.; Spinello, A.; Pennati, M.; Zaffaroni, N.; Gobbi, S.; Bisi, A.; Colombo, G.; Magistrato, A. A Computational Assay of Estrogen Receptor alpha Antagonists Reveals the Key Common Structural Traits of Drugs Effectively Fighting Refractory Breast Cancers. *Sci. Rep.* **2018**, *8*, 649.
15. Ghosh, D.; Griswold, J.; Erman, M.; Pangborn, W. Structural basis for androgen specificity and oestrogen synthesis in human aromatase. *Nature* **2009**, *457*, 219-23.
16. Hackett, J. C.; Brueggemeier, R. W.; Hadad, C. M. The final catalytic step of cytochrome P450 aromatase: A density functional theory study. *J. Am. Chem. Soc.* **2005**, *127*, 5224-5237.
17. Sgrignani, J.; Iannuzzi, M.; Magistrato, A. Role of Water in the Puzzling Mechanism of the Final Aromatization Step Promoted by the Human Aromatase Enzyme. Insights from QM/MM MD Simulations. *J. Chem. Inf. Model.* **2015**, *55*, 2218-2226.

18. Jiang, W.; Ghosh, D. Motion and flexibility in human cytochrome p450 aromatase. *PLoS One* **2012**, *7*, e32565.
19. Park, J.; Czapla, L.; Amaro, R. E. Molecular simulations of aromatase reveal new insights into the mechanism of ligand binding. *J. Chem. Inf. Model.* **2013**, *53*, 2047-56.
20. Caporuscio, F.; Rastelli, G.; Imbriano, C.; Del Rio, A. Structure-based design of potent aromatase inhibitors by high-throughput docking. *J. Med. Chem.* **2011**, *54*, 4006-17.
21. Ghosh, D.; Lo, J.; Morton, D.; Valette, D.; Xi, J.; Griswold, J.; Hubbell, S.; Egbuta, C.; Jiang, W.; An, J.; Davies, H. M. Novel aromatase inhibitors by structure-guided design. *J. Med. Chem.* **2012**, *55*, 8464-76.
22. Miller, W. R.; Larionov, A. A. Understanding the mechanisms of aromatase inhibitor resistance. *Breast Cancer Res.* **2012**, *14*, 201.
23. Favia, A. D.; Nicolotti, O.; Stefanachi, A.; Leonetti, F.; Carotti, A. Computational methods for the design of potent aromatase inhibitors. *Expert Opin. Drug Discov.* **2013**, *8*, 395-409.
24. Sgrignani, J.; Cavalli, A.; Colombo, G.; Magistrato, A. Enzymatic and Inhibition Mechanism of Human Aromatase (CYP19A1) Enzyme. A Computational Perspective from QM/MM and Classical Molecular Dynamics Simulations. *Mini Rev. Med. Chem.* **2016**, *16*, 1112-24.
25. Lu, W. J.; Desta, Z.; Flockhart, D. A. Tamoxifen metabolites as active inhibitors of aromatase in the treatment of breast cancer. *Breast Cancer Res. Treat.* **2012**, *131*, 473-81.
26. Egbuta, C.; Lo, J.; Ghosh, D. Mechanism of inhibition of estrogen biosynthesis by azole fungicides. *Endocrinology* **2014**, *155*, 4622-8.
27. Rampogu, S.; Son, M.; Park, C.; Kim, H. H.; Suh, J. K.; Lee, K. W. Sulfonanilide Derivatives in Identifying Novel Aromatase Inhibitors by Applying Docking, Virtual Screening, and MD Simulations Studies. *Biomed. Res. Int.* **2017**, *2017*, 2105610.
28. Wagner, J. R.; Lee, C. T.; Durrant, J. D.; Malmstrom, R. D.; Feher, V. A.; Amaro, R. E. Emerging Computational Methods for the Rational Discovery of Allosteric Drugs. *Chem. Rev.* **2016**, *116*, 6370-90.
29. Sgrignani, J.; Bon, M.; Colombo, G.; Magistrato, A. Computational approaches elucidate the allosteric mechanism of human aromatase inhibition: a novel possible route to Small-molecule regulation of CYP450s activities? *J. Chem. Inf. Model.* **2014**, *54*, 2856-68.
30. Magistrato, A.; Sgrignani, J.; Krause, R.; Cavalli, A. Single or Multiple Access Channels to the CYP450s Active Site? An Answer from Free Energy Simulations of the Human Aromatase Enzyme. *J. Phys. Chem. Lett.* **2017**, *8*, 2036-2042.
31. Ghosh, D.; Egbuta, C.; Lo, J. Testosterone complex and non-steroidal ligands of human aromatase. *J. Steroid. Biochem. Mol. Biol.* **2018**, *181*, 11-19.
32. Yu, X.; Cojocar, V.; Mustafa, G.; Salo-Ahen, O. M.; Lepesheva, G. I.; Wade, R. C. Dynamics of CYP51: implications for function and inhibitor design. *J. Mol. Recognit.* **2015**, *28*, 59-73.
33. Ohta, Y.; Kawato, S.; Tagashira, H.; Takemori, S.; Kominami, S. Dynamic structures of adrenocortical cytochrome P-450 in proteoliposomes and microsomes: protein rotation study. *Biochemistry* **1992**, *31*, 12680-7.
34. Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **2006**, *49*, 534-53.
35. Shaik, S.; Kumar, D.; de Visser, S. P.; Altun, A.; Thiel, W. Theoretical perspective on the structure and mechanism of cytochrome P450 enzymes. *Chem. Rev.* **2005**, *105*, 2279-328.
36. Taraphder, S.; Hummer, G. Protein side-chain motion and hydration in proton-transfer pathways. Results for cytochrome p450cam. *J. Am. Chem. Soc.* **2003**, *125*, 3931-40.
37. Lo, J.; Di Nardo, G.; Griswold, J.; Egbuta, C.; Jiang, W.; Gilardi, G.; Ghosh, D. Structural basis for the functional roles of critical residues in human cytochrome p450 aromatase. *Biochemistry* **2013**, *52*, 5821-9.
38. Sohl, C. D.; Guengerich, F. P. Kinetic analysis of the three-step steroid aromatase reaction of human cytochrome P450 19A1. *J. Biol. Chem.* **2010**, *285*, 17734-43.

39. Gohlke, H.; Case, D. A. Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *J. Comput. Chem.* **2004**, *25*, 238-50.
40. Genheden, S.; Ryde, U. How to obtain statistically converged MM/GBSA results. *J. Comput. Chem.* **2010**, *31*, 837-46.
41. Lammi, C.; Zanoni, C.; Aiello, G.; Arnoldi, A.; Grazioso, G. Lupin Peptides Modulate the Protein-Protein Interaction of PCSK9 with the Low Density Lipoprotein Receptor in HepG2 Cells. *Sci. Rep.* **2016**, *6*, 29931.
42. Sattin, S.; Tao, J.; Vettoretti, G.; Moroni, E.; Pennati, M.; Lopercolo, A.; Morelli, L.; Bugatti, A.; Zuehlke, A.; Moses, M.; Prince, T.; Kijima, T.; Beebe, K.; Rusnati, M.; Neckers, L.; Zaffaroni, N.; Agard, D. A.; Bernardi, A.; Colombo, G. Activation of Hsp90 Enzymatic Activity and Conformational Dynamics through Rationally Designed Allosteric Ligands. *Chem. Eur. J.* **2015**, *21*, 13598-608.
43. Moroni, E.; Zhao, H.; Blagg, B. S.; Colombo, G. Exploiting conformational dynamics in drug discovery: design of C-terminal inhibitors of Hsp90 with improved activities. *J. Chem. Inf. Model.* **2014**, *54*, 195-208.
44. Cornish-Bowden, A. (2004). *Fundamentals of Enzyme Kinetics* (London: Portland Press Ltd.).
45. Dubey, K. D.; Shaik, S. Choreography of the Reductase and P450BM3 Domains Toward Electron Transfer Is Instigated by the Substrate. *J. Am. Chem. Soc.* **2018**, *140*, 683-690.
46. De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* **2016**, *59*, 4035-61.
47. Pan, A. C.; Borhani, D. W.; Dror, R. O.; Shaw, D. E. Molecular determinants of drug-receptor binding kinetics. *Drug Discov. Today* **2013**, *18*, 667-673.
48. Franco, D.; Vargiu, A. V.; Magistrato, A. Ru[(bpy)(2)(dppz)](2)(+) and Rh[(bpy)(2)(chrysi)](3)(+) targeting double strand DNA: the shape of the intercalating ligand tunes the free energy landscape of deintercalation. *Inorg. Chem.* **2014**, *53*, 7999-8008.
49. Hammoudeh, D. I.; Follis, A. V.; Prochownik, E. V.; Metallo, S. J. Multiple independent binding sites for small-molecule inhibitors on the oncoprotein c-Myc. *J. Am. Chem. Soc.* **2009**, *131*, 7390-401.
50. Bernetti, M.; Cavalli, A.; Mollica, L. Protein-ligand (un)binding kinetics as a new paradigm for drug discovery at the crossroad between experiments and modelling. *Medchemcomm* **2017**, *8*, 534-550.
51. Polic, V.; Auclair, K. Allosteric Activation of Cytochrome P450 3A4 via Progesterone Bioconjugation. *Bioconjug. Chem.* **2017**, *28*, 885-889.
52. Williams, P. A.; Cosme, J.; Ward, A.; Angove, H. C.; Matak Vinkovic, D.; Jhoti, H. Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* **2003**, *424*, 464-8.
53. Denisov, I. G.; Baylon, J. L.; Grinkova, Y. V.; Tajkhorshid, E.; Sligar, S. G. Drug-Drug Interactions between Atorvastatin and Dronedarone Mediated by Monomeric CYP3A4. *Biochemistry* **2018**, *57*, 805-816.
54. Hlavica, P. Challenges in assignment of allosteric effects in cytochrome P450-catalyzed substrate oxidations to structural dynamics in the hemoprotein architecture. *J. Inorg. Biochem.* **2017**, *167*, 100-115.
55. Akhtar, M. K.; Kelly, S. L.; Kaderbhai, M. A. Cytochrome b(5) modulation of 17{alpha} hydroxylase and 17-20 lyase (CYP17) activities in steroidogenesis. *J. Endocrinol.* **2005**, *187*, 267-74.
56. Wishart, D. S.; Knox, C.; Guo, A. C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* **2006**, *34*, D668-72.
57. Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: a free tool to discover chemistry for biology. *J. Chem. Inf. Model.* **2012**, *52*, 1757-68.
58. Small-Molecule Drug Discovery Suite 2017-1, Schrodinger, LLC, New York, NY, 2017.

59. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3-26.
60. Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* **2002**, *45*, 2615-23.
61. Teague, S. J. Implications of protein flexibility for drug discovery. *Nat. Rev. Drug Discov.* **2003**, *2*, 527-41.
62. Laine, E.; Goncalves, C.; Karst, J. C.; Lesnard, A.; Rault, S.; Tang, W. J.; Malliavin, T. E.; Ladant, D.; Blondel, A. Use of allosterity to identify inhibitors of calmodulin-induced activation of *Bacillus anthracis* edema factor. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 11277-11282.
63. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739-49.
64. Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J.; Honig, B.; Shaw, D. E.; Friesner, R. A. A hierarchical approach to all-atom protein loop prediction. *Proteins* **2004**, *55*, 351-67.
65. Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859-65.
66. Lipowsky, R., Sackmann, E. (1995). *Structure and Dynamics of Membranes from Cells to Vesicles*. (Elsevier), pp. 1-63.
67. Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **2012**, *40*, W537-41.
68. Di Nardo, G.; Breitner, M.; Bandino, A.; Ghosh, D.; Jennings, G. K.; Hackett, J. C.; Gilardi, G. Evidence for an elevated aspartate pK(a) in the active site of human aromatase. *J. Biol. Chem.* **2015**, *290*, 1186-96.
69. Wickstrom, L.; Okur, A.; Simmerling, C. Evaluating the performance of the ff99SB force field based on NMR scalar coupling data. *Biophys. J.* **2009**, *97*, 853-6.
70. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950-8.
71. Dickson, C. J.; Madej, B. D.; Skjevik, A. A.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C. Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.* **2014**, *10*, 865-879.
72. Shahrokh, K.; Orendt, A.; Yost, G. S.; Cheatham, T. E. Quantum Mechanically Derived AMBER-Compatible Heme Parameters for Various States of the Cytochrome P450 Catalytic Cycle. *J. Comput. Chem.* **2012**, *33*, 119-133.
73. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157-74.
74. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - the Resp Model. *J. Phys. Chem.* **1993**, *97*, 10269-10280.
75. M. J. Frisch, G. W. T., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W.



- Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, *Gaussian 09, Revision A.02*, (Gaussian, Inc., Wallingford CT, 2009).
76. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **2006**, *25*, 247-60.
77. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926-935.
78. Sousa da Silva, A. W.; Vranken, W. F. ACPYPE - AnteChamber PYthon Parser interface. *BMC Res. Notes* **2012**, *5*, 367.
79. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701-18.
80. Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089-10092.
81. Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
82. Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604-613.
83. Sgrignani, J.; Magistrato, A. QM/MM MD Simulations on the Enzymatic Pathway of the Human Flap Endonuclease (hFEN1) Elucidating Common Cleavage Pathways to RNase H Enzymes. *ACS Catal.* **2015**, *5*, 3864-3875.
84. Bisha, I.; Laio, A.; Magistrato, A.; Giorgetti, A.; Sgrignani, J. A Candidate Ion-Retaining State in the Inward-Facing Conformation of Sodium/Galactose Symporter: Clues from Atomistic Simulations. *J. Chem. Theory Comput.* **2013**, *9*, 1240-1246.
85. Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide folding: When simulation meets experiment. *Angew. Chem. Int. Ed.* **1999**, *38*, 236-240.
86. Spinello, A.; Vecile, E.; Abbate, A.; Dobrina, A.; Magistrato, A. How Can Interleukin-1 Receptor Antagonist Modulate Distinct Cell Death Pathways? *J. Chem. Inf. Model.* **2019**, *59*, 351-359.
87. D.A. Case, T. A. D., T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman; *AMBER 12*, University of California: San Francisco, 2012.
88. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph. Model.* **1996**, *14*, 33-38.
89. Copeland, R. A. (2005). *Evaluation of Enzyme Inhibitors in Drug Discovery* (Hoboken: Wiley Interscience).
90. Baell, J. B.; Holloway, G. A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* **2010**, *53*, 2719-40.